# THE EVOLUTIONARY ORIGINS OF COGNITION: UNDERSTANDING THE EARLY EVOLUTION OF BIOLOGICAL CONTROL SYSTEMS AND GENERAL INTELLIGENCE

Ву

Anselmo Carvalho Pontes

### A DISSERTATION

Submitted to Michigan State University in partial fulfillment of the requirements for the degree of

Computer Science - Doctor of Philosophy Ecology, Evolutionary Biology and Behavior-Dual Major

#### ABSTRACT

### THE EVOLUTIONARY ORIGINS OF COGNITION: UNDERSTANDING THE EARLY EVOLUTION OF BIOLOGICAL CONTROL SYSTEMS AND GENERAL INTELLIGENCE

By

#### Anselmo Carvalho Pontes

In the last century, we have made great strides towards understanding natural cognition and recreating it artificially. However, most cognitive research is still guided by an inadequate theoretical framework that equates cognition to a computer system executing a data processing task. Cognition, whether natural or artificial, is not a data processing system; it is a control system.

At cognition's core is a value system that allows it to evaluate current conditions and decide among two or more courses of action. Memory, learning, planning, and deliberation, rather than being essential cognitive abilities, are features that evolved over time to support the primary task of deciding "what to do next".

I used digital evolution to recreate the early stages in the evolution of natural cognition, including the ability to learn. Interestingly, I found cognition evolves in a predictable manner, with more complex abilities evolving in stages, by building upon previous simpler ones. I initially investigated the evolution of dynamic foraging behaviors among the first animals known to have a central nervous system, Ediacaran microbial mat miners. I then followed this up by evolving more complex forms of learning. I soon encountered practical limitations of the current methods, including exponential demand of computational resources and genetic representations that were not conducive to further scaling. This type of complexity barrier has been a recurrent issue in digital evolution. Nature, however, is not limited in the same ways; through evolution, it has created a language to express robust, modular, and flexible control systems of arbitrary complexity and apparently open-ended evolvability.

The essential features of this language can be captured in a digital evolution platform. As an early demonstration of this, I evolved biologically plausible regulatory systems for virtual cyanobacteria. These systems regulate the cells' growth, photosynthesis and replication given the daily light cycle, the cell's energy reserves, and levels of stress. Although simple, this experimental system displays dynamics and decision-making mechanisms akin to biology, with promising potential for open-ended evolution of cognition towards general intelligence.

Copyright by ANSELMO CARVALHO PONTES 2021 In memory of my mother, Elfa.

#### ACKNOWLEDGEMENTS

Thank you to Laura Grabowski and David Bryson, who helped me set up the first Avida experiments, and to Andrew Mitchell, who created the tool to plot the behavior arenas and the digital organisms' paths. Many thanks to my collaborators: Nick Panchy, who spent many hours with me discussing the fine details of gene regulation and cell metabolism; Shin-Han Shu, who suggested I used a photosynthetic cell as a model for the evolution of gene regulation, and let me take Nick away from his work; Fred Dyer, for the numerous hours he spent working on paper drafts and educating me on animal behavior and cognition; Charles Ofria, for getting me unstuck many times, and giving me clever suggestions about how to solve problems and make our research projects better; Ian Whalen and Andrew Mitchell who worked diligently and creatively on our projects, with great results; Cliff Bohm for replicating our results in record time, and providing constructive criticism; Robert Mobley and Ali Tehrani for their time, ideas, and support.

I am especially thankful to David Arnosti, William Henry, Shelagh Ferguson-Miller, Kay Holekamp, Mark Reimers, Heather Eisthen, Chris Adami, and Charles Ofria, who welcomed me into their classes, supported me in my objectives, and from whom I learned so much.

I am equally thankful to the Evolving Intelligence group, especially Rob Pennock, Fred Dyer, and Chris Adami, a wonderful assembly of like minds, who gave me great feedback on my research, pointed me to interesting directions and taught me a great deal about our field.

vi

I would like to give a special thanks to Erik Goodman and Wolfgang Banzhaf for the time they spent listening to my ideas and providing feedback and support.

I am also very grateful to Luis Zaman, Josh Nahum, Anne Sonnenschein, Jory Schossau and Bill Punch for their time, and suggestions on great research and resources that were important to my work.

Thank you very much to Annat Harber for reviewing the manuscript of Chapter 2 and making numerous suggestions to improve it. I also appreciate your enthusiasm, support, and advice, which were crucial to this project.

Thank you to Ricardo Chagas, Scott Wagner, Alex Smith (UWEC) and Michael R. Weil all of whom were instrumental along the way.

Thank you to my lab mates, past and present, Emily Dolson, Anya Vostinar, Rose Canino-Koning, Alex Lalejini, Mathew Rupp, Michael Wiser, Luis Zaman, Jay Bundy, Aaron Wagner, and Heather Goldsby for your help over the years.

Just as appreciated is the help, support, and advice from my graduate committee members, Charles Ofria, Fred Dyer, Chris Adami, David Arnosti and Wolfgang Banzhaf.

Special gratitude to Charles Ofria for welcoming me at MSU, being my advisor, and enabling me to realize my life-long dream of discovering wonderful things about the world.

Finally, my deepest gratitude to my wife, Amy Skalmusky, for tireless and unwavering support, and for sharing my dream.

vii

# TABLE OF CONTENTS

LIST OF TABLESx	i
LIST OF FIGURESxi	i
Chapter 1: Introduction 1	L
1.1 A Promising Idea Became a Sand Trap2	2
1.1.1 The Rise of Information Processing Theory (IPT)	2
1.1.2 IPT Reveals its Shortcomings	1
1.1.3 Alternative Theories5	5
1.2 What Is Cognition?	5
1.2.1 Cognition Controls Behavior	5
1.2.2 What Is Behavior?6	5
1.2.3 Cognition Is an Agent's Decision-Making System	3
1.3 Rational Agency and Decision-Making	)
1.3.1 What Is a Rational Agent?	)
1.3.2 Rational Agents Versus Clockwork10	)
1.3.3 The Tryptophan System in Escherichia coli13	3
1.3.4 Decision-Making in Abstract Terms15	5
1.3.5 The Value System and Its Role in Decision-Making16	5
1.3.6 Available Information and Its Role in Decision-Making17	7
1.3.7 Valuation-Associated Actions and Inaction18	3
1.3.8 Caveats of Rational Agency19	)
1.3.9 Self-Governance and Degrees of Autonomy20	)
1.3.10 Degrees of Intelligence across Different Dimensions	)
1.3.11 Collective Agency and Multicognition21	L
1.3.12 Coming Full Circle23	3
1.4 Control Systems, and Why Cognition Is Not Information Processing24	1
1.4.1 What Are Control Systems?24	1
1.4.2 Control System Techniques and Objectives25	5
1.4.3 Implementing Control Systems in Computers25	5
1.4.4 The Data Processing Algorithm and IPT27	7
1.4.5 Natural Versus Artificial Control Systems28	3
1.4.6 Understanding Control Systems Illuminates Cognition	)
1.4.7 Why Cognition Is Not Information Processing	2
1.4.8 Changing the Cognition Discussion: Instead of Process Information, Make Decisions 33	3
1.5 Cognition is Intrinsic and Necessary for Life	1
1.5.1 What is Life?	1
1.5.2 Life Requires Cognition	5
1.5.3 Cognition Is a Precursor of Life	7
1.5.4 Life Is a Type of Rational Agency	)

1.6 Path Forward	40
Chapter 2: The Evolutionary Origin of Associative Learning	44
2.1 Introduction	44
2.2 Experimental System	50
2.3 The Behavioral Task	52
2.3.1 Experimental Conditions	56
2.4 Results	58
2.4.1 Repeated Evolution of Adaptive Behaviors: Error Recovery, Imprinting, and Re	eversal
Learning	58
2.4.2 Early trail predictability produces behavioral building blocks for learning	63
2.4.3 Learning May Not Generalize to Novel Environments	67
2.4.4 Cue Reversals During Evolution Foster Ability to Relearn During Lifetime	68
2.4.5 The Stepwise Evolution of Learning	70
2.4.6 Learning Can Evolve Suddenly	74
2.5 Discussion and Conclusions	75
2.5.1 Emergence of Learning Depends on the Prior Evolution of Reflexive Behaviors	75
2.5.2 Stepwise and Modular Evolution of Complex Behaviors	76
2.5.3 Why Learning Was Rare	78
2.5.4 The Scientific Value of an Open-Ended Evolutionary Model	79
2.5.5 Early Evolution of an Intrinsic Value System	80
2.5.6 Reversal Learning Seems No More Complex than Initial Learning	81
2.5.7 How Evolution Continues to Shape Associative Learning	82
2.5.8 Implications for Artificial Intelligence	83
2.5.9 Implications for the Evolution of Behavior	84
Chapter 3: Evolution of Patch Harvesting, an Insight into Early Bilaterian Cognition	85
3.1 Introduction	85
3.1.1 Background	87
3.1.2 Previous Research	87
3.1.3 Recent Evidence	88
3.1.4 Our Experiments	89
3.1.5 Our Findings	91
3.2 Methods	92
3.3 Results	96
3.3.1 Evolved Behaviors Resemble the Fossil Record	
3.3.2 Evolved Behaviors Fall into Four Stereotypical Strategies	97
3.3.3 Evolved Strategies Depended on Patch Structure	100
3.3.4 In-Depth Results from a Single Environment, Rectangular with Edges	102
3.3.5 The Spiraling Strategy Evolved Memory Usage	103
3.3.6 Lineage Analysis Shows Different Strategies Evolving from One Another	105
3.4 Discussion and Conclusions	107
3.4.1 Complex Trails Require Healthy Patches	107
3.4.2 Patch Boundaries May Have Guided Mat Mining Behavior	108

3.4.3 Peculiar Spiraling Behavior Balances Exploration and Exploitation	109
Chapter 4: Beyond Associative Learning, the Early Evolution of Configural Learni	ng 111
4.1 Introduction	111
4.1.1 Background and Motivation	111
4.1.2 Our Experiments	112
4.1.3 Our Findings	113
4.2 Methods	113
4.3 Results	116
4.4 Discussion	120
4.4.1 Extra Cognitive Abilities May Contribute to Adaptation	
4.4.2 We Are Approaching a Complexity Barrier	121
Chapter 5: Evolution of Allosteric Regulation in Cyanobacteria	123
5.1 Introduction	
5.1.1 A New Digital Evolution Platform	124
5.1.2 Evolution of Allosteric Regulation Experiment	125
5.1.3 Findings	126
5.2 Methods	126
5.2.1 Cell and Population Organization in Elfa	
5.2.2 Ancestral Cells	127
5.2.3 Ligands in Elfa	129
5.2.4 Cellular Costs and Constraints	130
5.2.5 Experimental parameters	131
5.2.6 Analyses	131
5.3 Results	132
5.4 Discussion and Conclusions	139
5.4.1 Elfa as an open-ended digital evolution platform	140
APPENDICES	142
Appendix A: Supplementary Material for Chapter 2	143
Appendix B: Supplementary Material for Chapter 4	
REFERENCES	

### LIST OF TABLES

Table 2.1: Environments for experiment 1
Table 2.2: Behavioral strategies found in all experiments60
Table 2.3: Experiment 1: summary of results
Table 3.1: Environments in the order they were used in the experiments. Each environmentconsisted of four arenas with different patch configurations. Each organism experienced onlyone arena in its lifetime
Table 3.2: Environmental Interaction instructions used in the experiment
Table 3.3: Behavioral strategies found across all treatments. For each environment, we chose the three populations with highest overall task quality across the 200 replicates and analyzed the navigation strategy of their predominant organisms. Although there was a great deal of variation, all behaviors could be classified into four major strategies
Table 3.4: Strategies that evolved among the three best performing populations from each      environment.      .101
Table A.1: Avida instructions mnemonic references 145
Table A.2: Preliminary experiment summary of results. Performance and strategies of the organisms with AMTQ equal or higher than 25%, organized by environment*
Table A.3: Experiment 2 and follow-up experiment summary of results. Comparison between the two conditions, standard Turing-complete instruction set, and non-Turing-complete, minimal memory instruction set

### LIST OF FIGURES

Figure 1.1: Steam loom centrifugal governor, Museum of Science and Industry, Manchester, l	JK.
	.26
Figure 2.1. Sample Arona and Nutrient Trail. Shown is one of four virtual aronas from an	
Figure 2.1: Sample Arena and Nutrient Trail. Shown is one of four virtual arenas from an	
environment. Each virtual arena contained a single trail of nutrients laid out in a unique	
configuration. At the beginning of its life, each organism was placed alone at the start of the	
trail (green circle) in a randomly selected arena and oriented in the direction of the next	
nutrient	53

Figure 2.2: Two top-performing strategies in experiment 1. Shown are the paths of the final predominant organisms from two different replicates that evolved in the nutrient cued environment in experiment 1. Both were tested in the same trail configuration to facilitate comparison. In the left panel, an organism using an error recovery strategy achieved a task quality score of 81% of the maximum. Starting from the green circle, it moved straight while sensing forward cues but always tried to turn right (45 degrees) when sensing a turn cue. If turning right led the organism into an empty cell, it would retreat to the previous position and turn toward the left (90 degrees). It continued to repeat this behavior at every turn cue without ever learning from its error. In the right panel, an organism from a separate replicate using a generalizable imprinting strategy achieved a task quality score of 98% of the maximum. It also tried to turn right when sensing a turn cue. However, it stepped off the path only once at the first left turn. It learned the correct cue-response association and navigated the remainder of the trail without error.

Figure 2.5: Evolutionary history: 10 Lineages. Shown is the evolution of task quality over time in each of the 10 lineages that were ultimately capable of serial relearning from experiment 2. As they transitioned to a new strategy, some lineages had great gains in task quality, while others had more gradual ones. All lineages, however, went through occasional periods of fitness loss. Different task quality ranges often corresponded to specific behavioral strategies. Range 1 corresponded to path predicting, range 2 corresponded to hybrid strategies that included searching, range 3 corresponded to error recovery, and ranges 4 and 5 corresponded to imprinting and relearning.

Figure 3.4: Illustration of the edge-reflecting (left) and edge-hugging (right) variations of the pattern cycling strategy
Figure 3.5: Example of the reactive meandering strategy in the environment irregular with holes and edges
Figure 3.6: Example of the plowing strategy in the environment rectangular with hole and edges
Figure 3.7: Example of the spiraling strategy in the environment disconnected patches without

Figure 5.1: The 18 populations that evolved allosteric regulation could be categorized into three groups according to which of their proteins evolved allosteric regulation. The first group (four populations) evolved growth factor regulation. The second group (10 populations) evolved RNA polymerase (RNAp) regulation. The last group (four populations) evolved both growth factor and RNAp regulation. All final populations had higher fitness than the ancestral one (green dotted line at 0.9 generations per day). In addition, most populations with regulation (14/18) had higher fitness than the best non-regulated final population (orange dashed line at 3.4 generations per day).

Figure 5.2: Gene regulatory network of the ancestral cell with no regulation. On the top are the five ligands available to the cell: ROS, fat reserves (Res), sugar reserves (Sug), irradiance (Light), and cAMP. The ancestral cell does not sense the concentrations of any of these ligands...... 135

Figure 5.3: Gene regulatory network of one of the cells with regulated growth factors. Here, the growth factor could bind three different ligands with different affinities, and with different effects. The fat reserves ligand (Res) would bind weakly and had an agonistic effect. The irradiance ligand (Light) would also bind weakly but would have a reverse agonistic effect. Finally, the cAMP ligand would bind strongly and also have a strong reverse agonistic effect. In practice, high fat reserves would promote a slight increase in growth factor production, while

Figure 5.4: Gene regulatory network of one of the cells with regulated RNAp. Here, the RNAp could bind two different ligands with different affinities, and with different effects. The ROS ligand would bind strongly and also cause a strong agonistic effect, while the sugar reserves ligand (Sug) would bind weakly and have a weak reverse agonistic effect. In practice, the presence of ROS, even in small amounts, would cause the cell to increase the expression rate of all its genes, including the growth factor, causing it to replicate faster. Replication increased the ratio of surface area to volume in the daughter cells, temporarily alleviating the source of stress. High sugar reserves (indicating active photosynthesis), on the other hand, would slightly reduce the expression rate of all genes, and contribute to slow down replication and build energy reserves.

Figure A.9: "Oscillator" organism. This organism moved back and forth between its star	t
position (green circle) and its final position (red circle)	

Figure A.17: Change in the behavioral algorithm from Error recoverer 2 to First learner due to a single mutation. The mutation connected the error-recovery module to the memory-storing module. Previously, memory-storing was executed only once, right after the organism was initialized on the trail (left sequence). After the mutation, every time it made a wrong turn and recovered, the organism stored the cue that led it off the trail in memory (right sequence)...171

Figure A.19: Distribution of Average Maximum Task Quality (AMTQ) per condition for experiment 2 and follow-up experiment. Each violin plot represents the distribution of AMTQ

across replicates for a given condition. The difference between the two conditions was not	•
significant (Kruskal–Wallis test, H = 0.11008, df = 1, p = 0.7401).	174

# **Chapter 1: Introduction**

Science has progressively revealed the nature of the world around us and within us, but scientific advances have been uneven. One subject that remains stubbornly resistant to our efforts is cognition. We do not know how it works or how to measure it, we are not sure who or what has it, and we even have trouble defining it. If we could understand cognition better, it would change the way we see ourselves and the living world around us. If we understood cognition to the point where we could recreate it artificially and with high degrees of competence, we could harness one of the most powerful forces of all and transform both technology and our society.

Currently, most scientific efforts related to cognition are based on a model that equates the brain to a type of computer system. However, this paradigm is outdated, and although there are alternative views, they have not yet gained traction. Our flawed understanding of cognition is stunting the field of artificial intelligence (AI), which relies on theoretical guidance from research on natural cognition.

In this chapter, I start by discussing the theory of cognition that is currently dominant, including its history and how it has affected AI research. Later, I discuss the nature of cognition in the context of agency and decision-making. I then provide a new perspective on the subject that I believe is key to advancing both our understanding of cognition and our research on AI. I relate the various concepts back to living organisms that were produced through an evolutionary

process and discuss approaches to creating AI with levels of autonomy and intelligence similar to what we see in nature.

In subsequent chapters, I explore the evolution of behavioral control systems that exhibit key cognitive abilities before finally discussing results from a system of my own. This system uses an approach more directly inspired by nature to produce versatile control systems with greater complexity than current evolutionary techniques permit.

# **1.1 A Promising Idea Became a Sand Trap**

### 1.1.1 The Rise of Information Processing Theory (IPT)

In the years following the Second World War, there was much excitement about the newly invented electronic digital computer and its potential for revolutionary applications. Early computers were called "electronic brains", not only in the popular press, but by scientists as well. Amid this computer age enthusiasm, the field of artificial intelligence (AI) was launched in 1956 [1], the same year that a new theory equating human cognition to a computer system started to take shape<sup>1</sup> [2].

This theory, which became known as Information Processing Theory (IPT), posits that human cognition is an instantiation of a Turing machine, a general-purpose symbol manipulation system. IPT holds that sensory data (stimuli), received as input, is converted into symbols,

<sup>&</sup>lt;sup>1</sup> In the summer of 1956, Allen Newell and Herbert A. Simon, two of the founders of Information Processing Theory (IPT), attended the historical workshop at Dartmouth College where the field of AI was named. There, they presented their Logic Theorist system, the first application of IPT [3].

stored in memory, and processed according to some rule, which then generates an output (behavior)<sup>2</sup> [2], [4].

Al rapidly adopted IPT and they shared some impressive early results, such as computer programs that proved theorems and processed natural language, and robots that navigated inside a room and manipulated objects [2], [5]. Buoyed by these achievements and a good dose of computer-age enthusiasm, IPT soon replaced Behaviorism as the dominant theory of cognition [2].

Originally, IPT considered cognition to be a uniquely human capacity – at most shared with other primates – since we were supposedly the only creatures capable of symbolic representation [2]. Later, animal behavior researchers extended IPT to non-primate animals [6] by relaxing the concept of representation<sup>3</sup>, going beyond explicit symbols, to include broadly the encoding of information in neural signals [8]–[11], [7]. However, this broader definition still implied that cognition was exclusive to organisms with a nervous system.

<sup>&</sup>lt;sup>2</sup> Information Processing Theory has many variations, and some authors refer to information processing as a set of related theories [2]. Despite their differences, the various flavors of information processing all share the view of cognition as a symbol manipulation process, following a well-defined input–process–output sequence, akin to a digital computer system executing a data processing task.

<sup>&</sup>lt;sup>3</sup> Gallistel [7] (p. 4 and 7) defines representation as "signals, symbols and the operations on them", where symbols and signals are functionally homomorphic to aspects of the world. That is, "the signals and symbols carry information about properties of the experienced world" in such a way that the brain, operating on those signals and symbols, can "anticipate behaviorally relevant states of the world". However, "a symbolic memory mechanism has not so far been identified ... The absence of a symbolic memory mechanism is a problem, because a mechanism functionally equivalent to the tape in Turing's abstract conception of a general-purpose computing machine is essential to computation and representation."

### 1.1.2 IPT Reveals its Shortcomings

Despite IPT's popularity, it eventually became clear that humans are not the general-purpose symbol manipulators that the theory predicted. Experiments demonstrated that humans do not rely as much on representation and often use heuristics to solve problems [12], [13]. IPT was also criticized for its one-way, serial processing, input–output model that strictly separated perception from action, when the two are often intrinsically coupled, such as when our eyes scan a scene for us to build an image [14], [15]. Moreover, IPT neglected the role of feedback and required an unrealistically accurate internal model of the world to achieve real-time motor control [5], [16].

Meanwhile, Al's early promising results turned out to be brittle, as they did not scale well to more complex versions of the same problem, nor did they generalize to related problems [5]. Since then, AI has mostly set aside its original goal of recreating human intelligence and has instead settled for domain-specific targets, a narrow range of tasks where norms can be articulated clearly and existing knowledge can be tapped to prime the system [1], [17].

There is also an ironic circularity in the relationship between AI and IPT. While AI researchers look to IPT to understand how cognition works in order to emulate the process in a computer system, IPT claims that cognition works just like ... a computer system. In spite of these shortcomings, IPT is still the most influential theory of cognition in AI, permeating most rulebased and connectionist approaches [15], [18].

### **1.1.3 Alternative Theories**

Alternative theories of cognition have been proposed where cognition is not akin to a data processing system and does not rely on symbolic representation. The most prominent ones, including Perceptual Control Theory [19], Dynamicism [20], and 4E Cognition [21], use a controlsystem paradigm instead of the sequential input–process–output algorithm. None, however, have yet achieved widespread acceptance. Most importantly, even in the light of these theories, cognition remains a mysterious phenomenon.

While AI is trapped in this theory quicksand, with no clear agreement on what cognition really is, recreating general intelligence<sup>4</sup> will likely remain out of reach. After all, how can we build what we do not understand?<sup>5</sup>

This dissertation attempts to understand cognition and find a path to artificially create it.

<sup>&</sup>lt;sup>4</sup> General intelligence, as opposed to domain-specific intelligence, refers to the multifaceted and versatile intelligence that natural organisms possess, which gives them autonomy and helps them adapt to changes in their environments. It is a broader term than artificial general intelligence (AGI), which refers specifically to human-like intelligence. It also differs from the term general AI, which loosely refers to AI systems that perform well across multiple domains, without specific retraining for each one.

<sup>&</sup>lt;sup>5</sup> This is a common quip among AI researchers and developers: how can we perform a task if we do not even know what the task is?

# 1.2 What Is Cognition?

### **1.2.1 Cognition Controls Behavior**

We routinely evaluate the cognitive abilities of organisms by observing their behavior. Cognition gives an organism the ability to adapt its behavior according to the circumstances. This is a point on which all theories of cognition generally agree: cognition is responsible for controlling behavior [19], [22], [20], [23], [21], [10]. What is surprising, however, is that there is no agreement on what qualifies as behavior.

### 1.2.2 What Is Behavior?

The extent of this disagreement was brought to light by a 2009 study [24] in which the authors reviewed the literature on behavior and surveyed members of three behavior-focused scientific societies and found that existing definitions of behavior often conflict with one another, are too vague, or exclude whole groups of organisms<sup>6</sup>. The authors also found that many researchers rely on their own intuitive definitions of behavior, which can be inconsistent and often contradict those from the literature.

In an attempt to reach consensus, the authors proposed yet another definition: "Behaviour is the internally coordinated responses (actions or inactions) of whole living organisms

<sup>&</sup>lt;sup>6</sup> According to some of these definitions, only animals are capable of behavior, even though there are many clear examples of cognitive abilities in organisms as varied as plants and protists [25]–[30].

(individuals or groups) to internal and/or external stimuli, excluding responses more easily understood as developmental changes."

As often happens in such compromise, this new definition is quite hedged and still open to interpretation. Furthermore, by downplaying behaviors that rely on physiological or developmental processes, the definition seems biased towards animals.

The reckoning and discussions brought about by the study spurred another round of papers proposing yet more definitions of behavior and ways to classify it. One paper, however, stands out by casting the phenomenon in a new light and addressing many of the limitations of the above definition<sup>7</sup> [31]. It proposes a definition of behavior that is grounded in the concept of agency, which is particularly insightful for the study of cognition: behavior is the "observable consequences of the choices a living entity makes in response to external or internal stimuli."

What the authors mean is that all living organisms should be understood as agents [32], entities capable of initiating their own actions and whose choice of action is influenced by their own state and the state of their environment<sup>8</sup>. A simple example of a choice is switching a molecular

<sup>&</sup>lt;sup>7</sup> The authors of this study are two plant scientists and a philosopher, all with an interest in cognition. They were motivated to propose a less-biased definition of behavior that applied to all living entities.

<sup>&</sup>lt;sup>8</sup> More precisely, the agent's estimate of the state of the environment.

pathway due to an environmental signal<sup>9</sup> [31]. The consequences of such a choice can be as overt as movement or as covert as the production of a chemical compound<sup>10</sup>.

Additionally, behavior and agency are not exclusive to single, whole organisms. Certain parts of multicellular organisms, such as individual cells, are also agents and capable of behavior. Likewise, a group of organisms, such as eusocial insects, can work together as a collective agent [31], [33].

### 1.2.3 Cognition Is an Agent's Decision-Making System

By viewing living organisms as agents, and behavior as a byproduct of agency and its intrinsic action-selection process, we begin to get a clearer picture of cognition. If cognition is what controls behavior, then cognition must be an agent's decision-making system. Moreover, any system making the decisions for another is a control system [34], so it follows that cognition is an agent's control system.

The idea that cognition is a control system is not new. Since at least the foundation of Cybernetics in the 1940s [35], several researchers based their theories in this assumption [19], [23], [36], [37]. Even Allen Newell, one of the founders of Information Processing Theory (IPT),

<sup>&</sup>lt;sup>9</sup> Clearly, choosing to do nothing when an alternative exists is also a choice.

<sup>&</sup>lt;sup>10</sup> The observability requirement in the definition is redundant and can be safely ignored. It was likely included for compatibility with historical views of behavior. Observability changes with the availability of resources and technology. Moreover, behaviors are classified in the literature as overt or covert, where the latter may be difficult or even impossible to observe.

understood as much. In his final book, "Unified Theories of Cognition" [22] (p. 43), Newell wrote:

I want to take the mind to be the control system that guides the behaving organism in its complex interactions with the dynamic real world. ... The mind then is simply the name we give to the control system that has evolved within the organism to carry out the interactions to the benefit of that organism or, ultimately, for the survival of its species.

Although Newell and others intuitively understood that cognition is a control system, they never actually demonstrated it. As I will show in the next sections, this is not merely a hypothesis nor a metaphor. The decision-making mechanism required for agency constitutes a control system, and this is what we call cognition<sup>11</sup>. Realizing this has many implications for the study of behavior and cognition in general and, more importantly, for AI, due to the vast differences between a control system and a data processing system. By shifting our perspective, we can finally free ourselves from the sand trap of IPT.

# **1.3 Rational Agency and Decision-Making**

# 1.3.1 What Is a Rational Agent?

In philosophy, an agent typically refers to a human being, an entity whose choices of action are guided by their own beliefs and preferences [33], [38]. Humans, however, are only one instance

<sup>&</sup>lt;sup>11</sup> The words *cognition* and *intelligence* are often used interchangeably. Throughout this dissertation, I will draw a distinction between them. I will use *cognition* to refer to a binary quality that an entity can either possess or not possess, while I will use *intelligence* to refer to a quality that cognitive entities possess in different degrees. I do, however, use the term artificial intelligence (AI) as it is commonly understood.

of a larger class of rational agents, or intentional systems [39]–[41], which can be natural, artificial, and even exobiological.

An agent is said to be rational if it always attempts to improve its own good, according to its own value system (its preferences), given its state and estimate of current conditions (its beliefs). Any entity that always make decisions randomly, without any sensitivity to changes in current conditions, is therefore neither cognitive nor a rational agent<sup>12</sup>.

All rational agents are cognitive, self-governing, and use information and energy. As noted above, however, they do not need to be living [39], [40]. A refrigerator, a Roomba vacuum cleaner, and a self-driving car are rational agents, just like bees, oak trees, and *E. coli* bacteria.

### **1.3.2** Rational Agents Versus Clockwork

It is important to distinguish agents<sup>13</sup> from the close-resembling but non-cognitive entities called clockwork<sup>14</sup>. Clockwork systems also use energy and can be quite complex, but they lack the fundamental capacity of adaptation. They are insensitive to changes in conditions and incapable of self-adjusting. They do not use information or make decisions. Some examples of

<sup>&</sup>lt;sup>12</sup> A rational agent may choose to behave randomly occasionally. To be able to produce random, unpredictable behavior at will is an asset in some circumstances, such as when being pursued by a predator or in certain social situations.

<sup>&</sup>lt;sup>13</sup> Throughout the rest of this dissertation, all references to "agents" should be understood to mean specifically "rational agents".

<sup>&</sup>lt;sup>14</sup> Although I do not agree with Boulding's system categorization overall [42], one of its categories named "clockworks" (p. 202) seems to describe the class of systems I am referring to here.

clockwork include a mechanical clock, an electric water pump, a Ford model T, and a typical protocell [43].

The addition of a control system (cognition), as simple as it may be, can turn a clockwork system into a rational agent. A control system provides the capacity to sense the state of at least one variable and to use that information to make the appropriate course of action.

Take, for example, an electric water pump that fills a swimming pool. This pump has no automation and must be switched on or off by a person. It is clockwork. The pump, however, can be easily automated and turned into a water-level maintaining device – a rational agent – albeit a very simple one. One way to do this is by placing a float in the pool, attached to an articulated arm (like a toilet float), so that the arm, depending on its angle, flips the pump's switch on or off. As the water level gets lower, the float dips into the pool until the arm reaches an angle that turns the pump on. The float then rises with the water level until the arm reaches the angle that turns the pump off.

This solution is admittedly clunky: there are many ways that this control system could be made smarter and more elegant. Nevertheless, the ensemble pump, float, articulated arm, and switch together make up a self-governing system, a rational agent.

The upgraded system uses the estimated water level to inform its choice of action. For example, in the state where the pump is off, the system continuously decides between doing nothing and turning the pump on. The threshold at which it prefers to turn the pump on is determined by its value system, an intrinsic set of biases due to design and manufacture (such

as the length of the arm, the position of the electrical switch, the displacement of the float, etc.). In more anthropomorphic language, the system can be described as valuing a water level above a certain threshold. When the system believes that the water level is low, it acts to raise this until it believes that the water has reached its preferred level.

It is easy to see how cognition, at its simplest level, can be built up from basic physics. We took a clockwork mechanism (electric pump and switch) and interposed it with another (float and articulated arm), in such a way that the original causal chain (electricity powering the pump) became conditional on a measured variable (pool water level), thus creating cognition from non-cognitive parts.

There is no upper limit for how complex cognition can be, and as it becomes more complex, it also becomes more opaque. For example, the automated water pump could be further elaborated by adding sensors and degrees of freedom. In addition to maintaining a minimum water level, the system could be made to maintain the water below a maximum level by controlling the pool's drain. The water level thresholds themselves could also be made adjustable and placed under the system's control, allowing the system to decide on different water levels depending on the time of the day or the weather, for example. As the control system becomes more complex, it may be beneficial to account for correlations among variables and regularities in the environment. Therefore, the system could be upgraded with memory to store past states and events. It could also be made to learn associations, make predictions, communicate with other systems, set goals, and make plans to achieve these goals.

These changes may improve the water pump, but they are not necessary for cognition. The ability to store memories, learn, make predictions, and so on are supporting features that often evolve, or are added by design, if they contribute to the agent's fitness (or performance). In Chapter 2, we demonstrate one such evolutionary process that produced agents with various types of cognitive abilities, including associative learning.

#### **1.3.3** The Tryptophan System in Escherichia coli

A biological analogue to the automated water pump is the tryptophan system in *Escherichia coli*. Living organisms' control systems are vastly more complex than our previous example. However, complex control systems (natural or artificial) are typically modular and made of simpler control systems that are arranged hierarchically, influencing one another, but allowing for relatively independent analysis of the parts. The tryptophan system, taken separately, is also a level-maintaining system based on proteins that interact with signaling chemicals to make decisions. It maintains the concentration of tryptophan in the cell above a minimum level by either allowing tryptophan to be produced when its concentration is low or blocking its production when tryptophan concentration is high.

*E. coli* depends on the amino acid tryptophan to make essential proteins. Shortage of tryptophan slows down protein production, which hinders most cellular activity, including growth and reproduction, and in extreme cases can lead to death. *E. coli* can obtain tryptophan from the environment, such as when it is inside an animal's gut, but it can also synthesize its own from simpler molecules, although with a cost. Synthesizing tryptophan requires spending time, energy, and other resources not only on the production of tryptophan, but also on the

production of the tryptophan synthesis machinery. These are resources that could be allocated to other cellular processes, and this ultimately impacts *E. coli*'s fitness. Therefore, *E. coli* should only synthesize tryptophan when its concentration inside the cell is low.

*E. coli* senses and controls tryptophan concentration by means of a repressor protein complex that the cell continuously produces in small amounts. Each of these complexes has pockets that can bind tryptophan and cause the complex to change conformation. The binding is temporary and depends on the concentration of tryptophan in the cell. When the concentration of tryptophan in the cell is low, most of the complexes will be unbound and inert, allowing the tryptophan synthesis machinery to be produced. As the tryptophan concentration gets higher, more complexes will bind tryptophan and become active, thus blocking the production of the tryptophan synthesis machinery, causing the synthesis of tryptophan to taper off.

All control systems have a cost. In this example, *E. coli* must produce the repressor protein complex continuously as well as maintain a gene for it. However, the cost of the tryptophan system is more than offset by the savings gained by producing tryptophan only when needed. Typically, control (also known as regulation) is only employed where its cost is offset by the savings or other benefits it provides. In nature, this is one reason why organisms evolve different cognitive abilities depending on their environment. In the following chapters, we perform digital evolution experiments and describe this phenomenon in more detail.

### 1.3.4 Decision-Making in Abstract Terms

*E. coli* makes decisions, such as those regarding tryptophan synthesis, in the context of its internal state. This includes the cell's energy level, whether it is obtaining tryptophan from the environment or synthesizing its own, the current number of repressor protein complexes, the number of copies of the tryptophan synthesis machinery, how many ribosomes it has to make new proteins, etc. The internal state narrows down the decisions the cell must make, as well as the information that is available and the actions that are relevant or even possible. In turn, the cell's decisions affect its internal state and set the stage for future decisions. Just like with *E.coli*, any rational agent's internal state is the result of its past decisions and experience.

In abstract terms, decision-making is the evaluation of current conditions, in the context of the agent's state, and the selection of a course of action associated with this valuation.

Decisions can be prompted by changes of state, previous decisions, or simply the passage of time. In both the automated pump system and the Tryptophan system, decisions are made continuously. An agent's decisions, however, can also happen discretely, or even as singular events, and – depending on an agent's complexity – multiple decisions can occur simultaneously.

Decision-making does not require deliberation, which may or may not precede the event of decision-making. Some agents have complex models of the world and can simulate the outcome of alternative actions to predict future conditions. Decision-making, however, does not require prediction about the future, only information about current conditions. Even agents

that are capable of predicting the future do not do this for every decision. More complex deliberation processes take longer and consume more time and resources; as such, they are reserved for only a subset of the decisions that an agent makes. Decisions are often made based on heuristics, which are faster than deliberation.

Goal setting, memory, and learning are also supporting – rather than core – features of cognition. They are not required for decision-making and may or may not be present in a particular rational agent.

#### 1.3.5 The Value System and Its Role in Decision-Making

Central to decision-making is the agent's value system, a set of biases used to evaluate current conditions and select a course of action. All agents have an intrinsic set of values, which vary among individuals. An agent's values can change over time and, depending on the agent, also be learned. Artificial agents can also have values that are explicitly supplied.

Because a value system guides an agent's decisions, an agent will tend toward certain endstates, or frequently-visited states, which we can interpret as the agent's goals. Thus, by observing an agent's behavior over time, we can deduce the agent's values and goals, even if the agent is not capable of explicit goal setting, which is a more complex cognitive ability and not essential to cognition.

In evolving agents, intrinsic values are under continuous selection, just as any other physical or behavioral trait, and those that contribute towards fitness tend to be selected over those that do not. We will discuss this in more detail in Chapter 2.

### 1.3.6 Available Information and Its Role in Decision-Making

Decision-making requires evaluating information in real-time (or its digital equivalent). Time constrains how much information is available for a decision, as well as how long that information remains relevant. Given that agents make decisions with limited time and based on partial information, decisions are rarely optimal. Time, however, can also reveal data patterns both over time and across sensors that convey additional information and create opportunity for learning. It is important to note that information extracted from sensors are estimates. The state of a variable can never be known for certain, only inferred from measurements.

Information is a fundamental resource for agents, just like energy. Agents select which signals to tend to, search for, and pursue, all the while evaluating competing signals, ignoring signals, pooling signals, filtering signals, applying error-correction to signals, etc. This helps the agent to obtain the best estimates about the state of the world, as well as about the agent itself, and therefore decide what to attend to. There may be competing needs and not enough means to meet all the demands simultaneously.

When an agent responds to information from sensors, it responds to semantic information (meaning), not syntactic information (sensor data). Semantic information must be extracted from the sensor data stream before it can be evaluated during decision-making. The amount of semantic information that the data stream conveys does not need to correlate with the amount of syntactic information. For example, in agents with multiple sensors it is common that the data stream provides a continuous flow of sensor data, that the agent could not possibly respond to, nor it should. Only salient features of the data stream that are relevant to the

current state of the agent are evaluated. The salient feature (carrying semantic information) could be a pattern in the data, or specific values, or even the absence of the data stream.

Sensory data can be noisy and can have limited resolution and accuracy. Data can also be sparse, or there can be too much of it. Agents employ statistical treatments, within their capacity, to filter, integrate, and find correlations that help them infer the state of the variables of interest. The aim is to get the best possible inference from the available information.

In the case of the automated water pump, the size and shape of the float, and the inertia of the arm, combine to average the water level over an area and over time. The system is thus applying a statistical filter to the information. In the tryptophan system, the numerous copies of the tryptophan sensing complex and their sensitivity to tryptophan also produces an averaging effect. This is an example of how the statistical treatments that the cell uses (such as sensor copy number and affinity) are under evolutionary influence.

It is worth noting that, even in agents that are capable of predicting the future, they extrapolate from present information. They must first produce an estimate of current conditions in order to make a prediction. In terms of evolutionary origins, estimation must have come before prediction, since the first is required for the latter.

### **1.3.7 Valuation-Associated Actions and Inaction**

For every evaluation that an agent makes of the current conditions, there is an associated action. This could be inaction – the decision to do nothing – or it could be as simple as a change
of state. It could also be more complex: an agent could deploy an action consisting of a sequence of steps.

Actions can be innate or learned, although not all agents are capable of learning. In addition, associations between valuations and actions can either be innate or learned. In either case, upon reaching a valuation, the agent triggers the associated action.

### **1.3.8 Caveats of Rational Agency**

Both rational agency and decision-making have caveats due to their existence in physical reality. Some decisions that a rational agent makes may not be rational, while others may be rational but self-defeating. Agents can suffer wear and tear or degradation of their sensors. Agents can be affected by noise, fatigue, or toxins, or else suffer from external manipulation. Even digital agents face limitations of computational resources. In addition, agents can change over time due to experience or stage of development; this in turn impacts decision-making.

Factors such as environmental mismatch, novel situations, ambiguous cues, an inept cognitive system, self-defeating or detrimental values, and overwhelming or conflicting demands can also lead to poor decision-making.

Decision-making can occasionally be random in order to break ties, confuse adversaries, increase beneficial variation, etc. Nevertheless, as noted in Section 1.3.1, an entity that always makes decisions randomly is not a rational agent.

In addition, some actions may seem to initiate spontaneously [44]–[46]. Many of the actions that we assume to be spontaneous may be due to hidden chains of events, such as changes of

state, which include rational decisions. Some actions, however, may occasionally be triggered without the agent's authorization, due to degradation, external interference, or random processes (noise).

### **1.3.9 Self-Governance and Degrees of Autonomy**

All rational agents are self-governing since they are ultimately the ones that authorize their own actions. Rational agents can, however, have differing degrees of autonomy [33]. Humandesigned agents typically depend on externally-supplied values and goals (thresholds, tolerances, setpoints, etc.). They are therefore less autonomous than evolved agents, whose values – and goals derived from them – are entirely their own<sup>15</sup>.

The autonomy of human-designed agents may also be restricted to particular tasks and to particular periods of time, such as an automobile's cruise control. While the cruise control is engaged, it autonomously maintains the speed. The speed target, however, is externally supplied, and the agent has no control over the steering.

### 1.3.10 Degrees of Intelligence across Different Dimensions

Although all rational agents are cognitive, they vary in their intelligence, in terms of how effective their cognition is. An agent with high intelligence is one that makes the best decisions given constraints faced, such as limited time for decisions, incomplete information, and

<sup>&</sup>lt;sup>15</sup> One of the goals of AI is to make machines less dependent on human supervision, i.e., more autonomous. However, we also want to retain control over them, and make sure that their behavior conforms to our standards.

competing demands, notwithstanding that decisions are rarely optimal and organisms often rely on heuristics that give good enough solutions.

The intelligence of an agent – how effective its cognition is – can be measured across multiple dimensions and time scales. One perspective is to measure how well an agent is adapted to its own environment. This can be further divided into how competent the agent is at a specific type of task versus how competent the agent is at multiple types of tasks [47]. An agent could be considered very intelligent if it solves a hard task, in which case the agent would be said to have high capability. If this hard problem were rare, however, then another agent that is better at more common, easier problems – in other words, an agent with high generality – may be considered the more intelligent agent given a longer time scale.

Another perspective is to consider a higher order of generality in which an agent's ability to perform multiple novel tasks is assessed. This would be a measure of the agent's ability to adapt to novel environments. After all, an agent that is good at solving familiar problems may nevertheless struggle to solve a new type of problem. An agent can thus be intelligent in its own environment but poor at generalizing to other environments.

These two perspectives for measuring intelligence are of interest for AI research. The higher order of generality is of particular importance to the goal of achieving general intelligence.

### 1.3.11 Collective Agency and Multicognition

As noted in Section 1.2.3, individual agents can work together as a collective agent. The coupling or integration of agents can vary from occasional and facultative, as in the case of

whales coming together to hunt, to permanent and obligatory, as in the case of a honeybee colony or a multicellular organism such as a plant or animal.

There are both natural and artificial collective agents. An autonomous vehicle, for example, is made of up various cooperating subsystems, each of which is itself an agent. It is among natural agents, however, that we see a greater spectrum of organizations and the highest levels of complexity.

To take the example of a honeybee colony, here there is not just one but multiple orders of collective agency. First, a eukaryotic cell is a collective agent because it is made up of, among other things, mitochondria and ribosomes; these, although non-living, are bona-fide rational agents. Each individual bee is also a collective agent since, as is the case with any multicellular organism, it is made up of cells working together. The bee colony is the third order of collective agency.

When the connection among constitutive agents becomes tighter, and their integration becomes permanent, the collective agent can transition into a new higher-order individual agent with its own higher-order cognition. After the transition, much of the cognitive machinery remains subsumed at the level of the constitutive agents and obscured from analysis at the higher level. For example, a lot of processing that happens in neurons happens at the level of molecular circuits. Memory itself may even be stored intracellularly at that level.

This transitioning of a collective agent into a new higher-order agent is one of the main reasons why cognition appears so mysterious and impenetrable; it muddles the problem of what

cognition is. To understand cognition in humans or other animals, we must understand that this is a higher-order phenomenon: it is *multicognition*. We cannot understand the cognition of the supra-entity without understanding that many of its components are hidden in the simpler cognitive entities of which it is made. Our cognition, rather than dwelling in our brain, can be thought of as a brain of brains, distributed through our entire body, with trillions of interconnected cognitive units acting more or less as one.

## **1.3.12 Coming Full Circle**

Amusingly, it is worth noting that the word *intelligence* derives from the Latin *inter-legere*, meaning "to choose between" [25]. In our centuries-long quest to understand intelligence, it seems we are coming full circle and rediscovering what our forebears knew all along: intelligence means making good decisions.

# 1.4 Control Systems, and Why Cognition Is Not Information Processing

### 1.4.1 What Are Control Systems?

Any system that controls another system or process is a control system. A control system can be artificial (human engineered), such as a thermostat that maintains a consistent temperature in a room, or it can be natural (produced by evolution) such as the human thermoregulatory system that maintains our body temperature even as outside conditions change. A control system can be simple or complex, and it can be implemented in many different substrates, including mechanical, electronic, biochemical, and digital.

It is important to note that the terms "control", "regulate", "manage", and "govern" – when applied to systems – are all synonyms. Molecular biologists, for example, talk about gene regulation. Here "regulation" is just another word for "control".

Artificial control systems have existed for thousands of years, but until the Industrial Revolution they were largely curiosities or else had niche applications. It was during the Industrial Revolution that they became widely adopted and brought substantial gains in productivity; for example, Watt's centrifugal governor was used to control the speed of a steam engine. The Industrial Revolution also saw the beginning of control theory: James Maxwell's "On Governors" was published in 1868.

Control systems are fundamental to a lot of modern technology, including automatic appliances, factory robots, modern aircrafts, and space probes. Despite this, and although they have long been used by engineers, control systems largely remain a hidden technology [48]. A

better understanding of control systems across the wider science community would greatly benefit numerous areas of research, helping to solve many current and future challenges.

### **1.4.2 Control System Techniques and Objectives**

The most famous control system technique is the negative feedback loop, where the system tries to maintain a controlled variable close to a setpoint value by acting to minimize the error between the current value and the setpoint. This is the technique implemented in Watt's centrifugal governor (fig. 1) and in natural homeostatic systems. It is, however, only one of the many techniques that control systems can employ. Others include positive feedback, feed forward, open loop, etc.

For a control system, just as important as pursuing a target is the manner in which it achieves this. While a control system must respond quickly to changes in conditions, it must also minimize overshoots, oscillations, and hysteresis. It must be resilient to noise and perturbation too, and in case of partial failure of its components, it should degrade its performance rather than lose control completely (graceful degradation). Finally, it should minimize wear and tear of its components and the use of energy by avoiding superfluous control activity.

#### **1.4.3 Implementing Control Systems in Computers**

Control systems can be implemented using digital computers and, indeed, most modern ones are: computer-based control systems are cheap, practical, and powerful. Control systems as a

class are not, however, instantiations of a Turing machine. Some control systems are discrete,

but many are analog and continuous and can only be approximated using a digital computer.



**Figure 1.1**: Steam loom centrifugal governor, Museum of Science and Industry, Manchester, UK. Special types of algorithms are used when control systems are implemented using digital computers. Typically, these algorithms simulate real-time concurrent execution by using techniques such as time-slicing, where computer time is divided into short intervals and distributed among many different tasks that ideally would be performed simultaneously and continuously. Such tasks include the evaluation of current conditions to decide the next course of action, the deployment of action sequences, the polling of data from sensors, and the refreshing of outputs to actuators. The concurrent tasks are executed in a loop (which is continually on) for the duration of the mission and often for the lifetime of the system.

### 1.4.4 The Data Processing Algorithm and IPT

In contrast to the algorithms used to implement (real-time) control systems in computers, the algorithm that serves as the model for Information Processing Theory (IPT) is the standard data processing algorithm, as used in tasks such as payroll, word processing, and weather forecasting. This algorithm implements function mapping (in the mathematical sense) [22], which takes in some data and, through the application of a set of rules, outputs the corresponding result.

The adoption of the data processing algorithm as a model for IPT was likely influenced by several factors. This was the algorithm used in the early computer applications, and it remains the most common type of algorithm used today. It is still the standard algorithm taught in computer science education. Another factor is that in the 1940s and 1950s, control systems were simple and mostly mechanical or electromechanical devices; it was not until the late 1950s that digital computers became commercially available for control applications [49]. In addition, control algorithms have traditionally been researched within engineering and not computer science departments.

It is important to note that data processing applications, unlike control systems, do not have their own timing but instead follow the timing of data arrival. They start processing data as soon as it is received, and they produce the output as fast as possible. IPT suggests our cognition functions as a series of isolated transactions, each starting with an input and ending with an output [22]. Cognition supposedly waits for a stimulus, maps this stimulus to the appropriate behavior to execute, and then waits for the next stimulus, in a stimulus–action

loop. As a consequence of this flawed understanding of cognition, most cutting-edge AI systems – such as convolutional neural networks – are designed to operate in this way, mapping input to output and iterating if necessary.

#### **1.4.5 Natural Versus Artificial Control Systems**

Many natural control systems have been characterized over the years and, despite their unique substrates, they function much like artificial ones [50]–[53]. For example, in humans, the subsystem responsible for thermoregulation uses negative feedback control [54], while the subsystem responsible for birth contractions uses positive feedback control [55]. In *E. coli*, the subsystem responsible for chemotaxis implements a proportional–integral (PI) controller [50], which is the same type of control strategy used in calcium homeostasis in cows [56]. "Negative feedback control," "positive feedback control," and "PI controller" are engineering terms that describe effective control strategies that we developed independently; we only found out later that nature also uses them.

A variation of the PI control strategy called proportional—integral—derivative (PID) is likely the strategy most commonly used in industry today. This type of controller is not capable of making predictions about the future nor capable of learning. A more advanced control method is Model Predictive Control (MPC). MPC is capable of making predictions about the future of the dependent variables; it can also learn the model that describes the behavior of the dependent variables. One of its applications is to control flight paths of autonomous aircrafts. It will likely be many years before we discover whether bees and birds use a similar control strategy for flying. Nevertheless, there is an interesting parallel between how more sophisticated control

technology has incorporated prediction and learning in order to solve more complex problems, and how natural cognition may have evolved similar capabilities as organisms faced more complex behavioral control challenges [57].

There are a number of important differences between artificial control systems and natural ones. When an artificial control system is designed, the aim is for it to perform as optimally as possible while operating within well-defined constraints. Optimal performance is highly dependent on minimizing noise. When redundancy is required, the most common method to achieve this is to deploy additional copies of the same system. Furthermore, artificial control systems are designed to be easy to maintain.

By contrast, natural control systems must primarily be evolvable. Whereas a random change to an artificial control system would likely cause it to stop working, natural control systems must be robust to mutations: they must be resilient to negative mutations while at the same time being amenable to evolving adaptations. Natural control systems also tolerate higher levels of noise and may even exploit it [58], [59], allowing them to function in a wider range of conditions. They are also often redundant, but their redundancy is typically not due to multiple copies of the same system but rather due to alternative systems that overlap. Finally, they must have great generality to operate in unstructured, dynamic environments. As a result of these features, natural control systems are rarely optimal, but they are often resistant to failure, which also facilitates them being reused by evolution in new contexts.

Of course, not all artificial control systems are designed: some are evolved. Notable examples are the control systems of finless rockets and legged robots [60], [61]. Evolved control systems

have a lot more in common with natural systems. For example, they are more robust to mutations and more autonomous, since they evolve their own values and goals. Those artificial systems that have so far been evolved, however, are nowhere near as complex as natural systems.

Level of complexity is probably the most obvious difference between natural and artificial control systems. The most complex artificial control systems today – such as those responsible for automated factories, powerplants, and spacecrafts – are still much simpler than the control system responsible for a whole living organism, such as a bacterium, not to mention the vast distributed multicognitive systems of mammals. At the end of this chapter, we will discuss a path forward for producing more autonomous and intelligent – and inevitably more complex – control systems.

### 1.4.6 Understanding Control Systems Illuminates Cognition

The acceptance that cognition is a control system gives us a unified language to discuss both natural and artificial cognition. It also lets us use an existing framework to analyze, benchmark, construct, and deconstruct cognition.

For example, it is clear that mechanisms such as memory, learning, prediction, and planning are not fundamental for rational control, and thus not necessary for cognition. They may be valuable additions in certain circumstances, but their evolution or engineering depends on the cost-benefit tradeoff. Conversely, a value system, internal states, and the evaluation of current information to decide a course of action are fundamental requirements for rational control.

As discussed previously, the control system of more complex agents can be subdivided into subsystems to facilitate analysis. This, however, abstracts away their interconnections, and it is important to remember that cognition is the control system of the whole agent. For example, despite humans' highly derived form of cognition (multicognition) with its many specialized subsystems, the experience of psychological stress affects the organism at every level, from intellectual performance to digestion, to growth and development, and even to inheritable DNA methylation patterns that prime the next generation.

Something that is not apparent when we study simple control systems, but which becomes clear when we analyze more complex ones – natural or artificial – is that agents are not driven by sensor input, i.e. agents are not stimulus-response systems. An agent has its own motive force: the need to reevaluate current conditions and "decide what to do next" [62] (p. 233). For this, information is a resource, not a prompt.

A complex control system designed using best practices typically polls sensors at regular intervals to extract semantic information. It only responds to this information if it is relevant to the current state of the system. In other words, even if a sensor input is updated continually, the new values will not affect the behavior of the system until the system chooses to read the sensor and finds information that is relevant. This is similar to what occurs in nature, where organisms are awash with sensory stimuli but, at their own timing, selectively respond to some stimuli and not others, depending on the context.

Put simply, rational agents do not necessarily react to stimuli: they respond to information.

### 1.4.7 Why Cognition Is Not Information Processing

Using the words "information" and "process" when referring to cognition may seem innocuous. After all, it makes sense that cognition uses information and that it is a type of process. However, the expression "information processing" has a very specific meaning, synonymous with data processing. When used in reference to cognition, it stands for IPT's assumption that cognition is an instantiation of a Turing machine executing a data processing algorithm. This assumption, however, is false: cognition is a control system and control systems are not instantiations of a Turing machine. As discussed in Section 1.5.3, control systems can be implemented in digital computers as reasonable and practical approximations using a real-time control algorithm. The data processing algorithm is not a good model for a control system, even as an approximation.

Data processing follows a unidirectional chain of events that starts with data input and ends with the output of the corresponding result. Control systems also have inputs and outputs, but the chain of events does not need to flow in any particular direction. In addition, it is common to have inputs without correlated outputs and outputs without correlated inputs (past or present). There can also be internal chains of events that change the state of the system due to the passage of time and which do not involve inputs or outputs. In control systems, the output is not a transformation of the input: input and output may influence each other but they do not need to be connected. Control systems can also be highly recursive, in which case the system's decisions affect not only inputs and outputs but also the system itself and how it operates.

While data processing systems are driven by data (syntactic information), control systems are not. The receipt of data is what triggers a data processing algorithm to start working. Once it has produced the corresponding output, it either terminates or waits for more data. Control systems, on the other hand, operate continuously, and the receipt of new data, such as an updated sensor measurement, does not necessarily cause the system to respond. Instead, control systems regularly scan their sensors to extract semantic information and make it available for evaluation; they can continue to make decisions even in the absence of new data.

Although the data processing algorithm is not a good model for a control system, it should be noted that data processing can be part of a control system in a supporting role, such as in the application of statistical filters to data or in the storage and retrieval of data from memory. Also worth noting is that complex cognitive systems like ours can manipulate symbols and perform data processing tasks, but within limitations. The cognitive mechanisms that support these tasks perform very differently to the data processing algorithm.

# **1.4.8 Changing the Cognition Discussion: Instead of Process Information, Make** Decisions

The idea that cognition is information processing is deeply entrenched, including in the language that we use when discussing behavior, cognitive phenomena, and its mechanisms. It is important that we adjust our language to allow us to leave behind this flawed understanding of cognition. Instead of saying "process information," we should say "make decisions."

For example, instead of "mycorrhizal networks process information," we should say "mycorrhizal networks make decisions." Instead of "uncovering the information processing mechanisms of the cell," we should say "uncovering the decision-making mechanisms of the cell." Instead of "how information is processed in the cerebellum," we should say "how decisions are made in the cerebellum."

Other expressions such as "control the organism," "evaluate and make decisions," "coordinate its actions," and "sense and respond" are also good alternatives.

# 1.5 Cognition is Intrinsic and Necessary for Life

# 1.5.1 What is Life?

Of the definitions that we have discussed so far, life is the most elusive. However, there seems

to be a consensus that living systems possess a few basic properties [63]–[65]:

- Enclosure, meaning that the living system is separate from the environment;
- Self-maintenance<sup>16</sup>, meaning that the living system can maintain its integrity despite wear and tear, and through developmental and adaptation processes;
- Self-sustenance, meaning that the living system can obtain resources, such as energy, information, and materials, from the environment;
- Self-reproduction, meaning that the living system can produce additional living systems like itself.

It is possible to define life as possessing only a subset of these properties, however, this risks

including systems that are less viable or less interesting, as living. Likewise, it is possible to

<sup>&</sup>lt;sup>16</sup> Enclosure and self-maintenance can be summed as autopoiesis [66], [67].

define life with additional properties beyond these four, but that risks biasing our definition toward systems that closely resemble biological Earth life.

For example, a version of life where organisms could not self-replicate, would likely be fragile, prone to extinction, and vulnerable to changing environmental conditions. Reproduction provides opportunity for renewal, multiplication, perpetuation, and Darwinian evolution. Without reproduction, organisms would accumulate damage until they could no longer survive, since even self-repair mechanisms can suffer damage. In the absence of a process that created organisms anew, this version of life would be vulnerable to extinction. Moreover, without undergoing Darwinian evolution, these organisms would have limited ability to adapt and increase in complexity, and no ability to speciate. Therefore, this version of life would likely be very simple and dependent on a benign environment for survival.

Alternatively, a version of life where organisms could not self-maintain, would likely also be fragile, simple, and vulnerable to extinction. Without self-maintenance, organisms would be prone to degradation and need to replicate quickly in order to avoid extinction. This dynamic would select against size and complexity. Therefore, these organisms would likely remain small, simple, and dependent on a benign environment until they evolved mechanisms for selfmaintenance.

Finally, a definition of life with additional properties, for example that life must be a chemical system [68], would unnecessarily rule out life that we may create using a mechanical, electronic, or digital substrate.

### **1.5.2 Life Requires Cognition**

Assuming that the set of life defining properties above is sound, it becomes evident that life must also be capable of regulating its behavior based on current conditions; in other words, it must be cognitive. Without cognition, living systems would be clockwork, meaning they would need to self-maintain, self-sustain, and self-replicate without any form of regulation. Since these clockwork systems could not adjust their behavior as conditions changed, they would require an unrealistically stable, homogeneous, and benign environment to survive. Moreover, they would remain vulnerable to random events such as disrepair and mutations. If such systems could exist at all, they would likely be very simple.

In practice, a perfectly stable and homogeneous environment for life is not possible. Even the most consistent environment, such as an underground reservoir where free energy comes from radioactive decay, would be disrupted when organisms started replicating. The presence of a variable number of organisms and their potential interactions would be sufficient to make the environment heterogeneous, non-stable, and less predictable.

Moreover, regulation allows adaptation not only to the external environment but also to internal changes. For example, regulation and coordination among separate processes in a cell is essential for survival. In order to replicate, a cell must grow in size at the same rate that it duplicates its genetic material, otherwise, at cell division, there could be more of one than the other, causing a protein or genetic imbalance. In a clockwork cell, all of these processes would need to happen without sensing or regulation, depending on perfect sequencing and synchronization. This is an unrealistic proposition considering the stochasticity of the chemical

environment at the scale of the cellular components. Such a cell would be sensitive to errors, and difficult to evolve. Mutations could quickly lead to imbalances for which the cell would have no means to compensate.

Regulation also allows living systems to stabilize favorable steady state conditions far from equilibrium that would not be possible without active control. These nonequilibrium steady states enable living systems to achieve yields and accuracy levels that that are critical to life, such as in the error correction during DNA replication [69] while minimizing energy expenditure. In addition, regulation allows living systems to maximize the extraction of energy and other resources from the environment for a given energy budget, which directly affects their fitness. They do this by using information to make decisions, such as when cyanobacteria adjust their growth rate according to light availability (Chapter 5). Therefore, we can say that living systems convert information into free energy [70], [71], which clockwork systems cannot do.

A clockwork version of life would be brittle, less viable and unlikely to exist in practice. Therefore, we can assume that cognition must be part of life from the outset.

## 1.5.3 Cognition Is a Precursor of Life

There are some compelling theories of how life originated on Earth, such as Eigen and Schuster's hypercycles, Kauffman's autocatalytic sets, and Deacon's Autogen. Although none may turn out to fully explain the origins of life, their common premise, that life originated from self-replicating chemical systems undergoing pre-biotic Darwinian evolution, is likely correct.

In physical reality, self-replication is sufficient for Darwinian evolution. Replication implies inheritance, whether the system constructs a copy of itself using an explicit genetic template, or the system serves as a template for itself and divides by fission. In addition, any replicating system is subject to copy errors that produce variation. Even error-correcting copying processes in a computer are subject to errors, such as when cosmic rays change memory values. Finally, any self-replicating system is subject to natural selection, since space and resources, even if vast, are finite.

At any point in the evolution of these self-replicating chemical systems, they could have benefited from evolving self-regulation. Self-regulation could have stabilized favorable nonequilibrium conditions, shielding them from environmental variation and allowing them to replicate faster. Over time, these more stable self-regulated systems could have outlasted nonregulated ones and also become more complex.

Regulation would have been relatively straight-forward to evolve. All that is required is that one or more of the molecular components in the self-replicating chemical system acquires allostery, i.e., it becomes sensitive to another chemical from the environment, and changes behavior in its presence. Allostery is relatively common among macromolecules (biological or not), which generally form pockets that can temporarily bind smaller molecules. If this happens to an enzyme, for example, the binding may increase or decrease the enzyme's function.

If a newly evolved allostery turned out to be beneficial, such as when the binding chemical signaled a change in conditions that the system could now respond adaptively, it could be selected during evolution. Of course, allostery is not the only form of regulation that self-

replicating chemical systems could have evolved, but it is the simplest and one that could evolve commonly. (I explore this in experiments in Chapter 5)

Therefore, a proto-life chemical system evolving allostery and later giving origin to the first cell is immensely more plausible than a fully formed clockwork cell arising, and then evolving regulation. Therefore, it is safe to assume that cognition appeared before life, and that it is intrinsic to it. In fact, cognition is one of the properties that enables life. As Jacques Monod said, allostery is "the second secret of life", the first being the genetic code [72].

### 1.5.4 Life Is a Type of Rational Agency

A simpler way of defining life is as a type of rational agency, where most actions are selfdirected: self-maintenance, self-sustenance, and self-reproduction. This contrasts with manmade rational agents, which direct most of their actions externally, such as the pool pump system keeping the level of the water.

Life's self-directed actions create recursive processes where the system's actions affect the mechanisms that produce those actions, and so forth. Additionally, many of these processes are operating in parallel and are dynamically interacting, often at a microscopic level and at different time scales. This intricacy makes it difficult to analyze and understand life's underlying processes, causing them to appear mysterious and special.

Life is also an extreme case of autonomy and intelligence. While man-made rational agents need to be supervised, living beings are independent. They have their own values, are

responsible for their own decisions, and have been selected by evolution to be highly competent within their environments.

While we would like to create machines that reach levels of autonomy and intelligence similar to natural organisms, we do not need to recreate life to achieve this. We should, however, attempt to emulate life's genetic representation and the sensing and regulatory mechanisms it expresses, since they have the track record of evolving control systems with the levels of complexity and competence that we seek.

## 1.6 Path Forward

Although cognition is a simple concept, the demands of natural life have led to the evolution of cognitive systems with extreme levels of autonomy, intelligence, and complexity. Even the most advanced man-made control systems, such as those from automated factories, autonomous vehicles, and space probes, seem simplistic in comparison to the cognitive system of a lowly bacterium. If we succeed in creating a control system with the levels of intelligence and autonomy of an insect, or even a mammal, it will be by far the most complex system we have ever created.

However, current engineering techniques may not be up to the task. Manually written software reaches a complexity barrier somewhere between 10<sup>7</sup> and 10<sup>8</sup> lines of code [73]. State-of-the-art ANNs get exponentially harder to train and deploy as they become more complex, the largest current ones costing upwards of millions of dollars in computer time to train [74], indicating that they, too, may be approaching a complexity barrier. Another potentially more

vexing problem, whether we are writing the software or training an ANN, is that people are not good at anticipating all the possible scenarios that an autonomous system will encounter during its operation, which can lead to hard failure in a multitude of unforeseen cases.

Digital evolution would appear to be the perfect way to solve this problem. However, current digital evolution approaches have struggled with a low complexity barrier. If we could overcome the technical limitations, evolution would allow us to create cognition in an open-ended manner, which has the potential to be more generalizable than anything we design or train. Moreover, evolution allows us to make progress even when we do not fully understand the problem, as is the case with complex cognition. Evolution is an inventive process; all it requires is for us to be able to evaluate the quality of a provided solution, even if we cannot envision a high-quality solution ourselves.

In the following chapters, I apply current evolutionary techniques to the problem of evolving cognition and demonstrate a potential solution to the complexity barrier in the form of a new evolutionary platform based on a genetic representation modeled after those found in nature.

In Chapter 2, we used the evolutionary platform Avida to investigate the evolution of navigation behavior and associative learning. During these experiments we observed that behaviors evolve in a predictable sequence, where more complex behaviors build upon simpler ones. Avida proved well-suited for the evolution of behavioral control systems, thanks to its bacteria-like circular genome. Avida's genome allows the control program to be executed continuously without contrived entry and exit points. In addition, when instructions for input and output

evolve, they can be used asynchronously at the request of the control program, informing program's decisions instead of prompting them.

In Chapter 3, we used Avida to investigate the evolution of patch harvesting behaviors. The earliest animals with a nervous system in the fossil record are microbial mat miners from the late Ediacaran. Their fossilized burrows are found all over the world and have intrigued scientists for over a century. Although our study was exploratory, we were able to replicate the harvesting patterns seen in the fossil record, propose evolutionary relationships among different behaviors, and identify an additional selection pressure that may have driven the evolution of these animals' behavior.

In Chapter 4, we extended the study on the evolution of navigation and associative learning to evolve a more complex learning ability, configural learning. This ability allowed an agent to distinguish between combinations of cues and individual cues and respond differently in each case. Although we succeeded in evolving the expected behavior, we also seem to have approached a practical limit on the complexity of the control systems that we can evolve with the current version of Avida.

In Chapter 5, I demonstrate Elfa, a new digital evolution platform inspired on cellular regulatory and signaling mechanisms, by investigating the evolution of allosteric regulation in simulated cyanobacteria. Ancestral cells express three core genes at low and constant rates; a growth factor gene, a photosynthetic complex gene and an RNA polymerase gene. Cells also contain five types of ligands whose concentration signals the amount of energy reserves the cell contains, the intensity of the environmental light, starvation, and reactive oxygen species

stress. I found that while most populations evolved a higher rate of gene expression without any regulation, some populations evolved allosteric regulation of different genes, which adjusted their behavior to the light cycle, allowing them to harvest more energy, and replicate faster than any of the non-regulated populations.

# **Chapter 2: The Evolutionary Origin of Associative**

# Learning

Authors: Anselmo C. Pontes, Robert B. Mobley, Charles Ofria, Christoph Adami, and Fred C. Dyer

This chapter is adapted from Am. Nat. 2020. Vol. 195, pp. E1–E19 [75]. © 2019 by The University of Chicago. CC BY-NC 4.0. DOI: 10.1086/706252

# 2.1 Introduction

Associative learning has long been considered fundamental to the adaptability of behavior and development of knowledge about the world [76]. It is also widely assumed that associative learning emerged as animal behavior evolved greater complexity and may have provided new avenues for this complexity to increase [77]–[81]. The general fitness advantage of learning in living organisms seems clear: learning enables an organism to adapt its behavior during its lifetime without requiring genetic changes across generations (as with evolution), and, unlike other forms of behavioral plasticity that occur during development, learning can result in very rapid rather than gradual behavioral modifications [82], [83]. Most research on the evolution of learning has focused on the adaptive specialization of learning—how the speed of learning, biases to learn certain things better than others, and capacity to store learned information correlate with the reliance on learning in an organism's natural environment [84]–[89]. Little is known, however, about the historical question of what selection pressures and evolutionary

precursors facilitated the emergence of learning from ancestors incapable of doing so or about the processes that allowed more complex forms of learning to evolve from simpler ones [90], [91].

Most people assume that complex behavior evolves in response to complex challenges; however, the evolution of behavioral complexity need not entail the emergence of learning [23]. Rather, learning evolves under specific environmental dynamics: where conditions that are relevant to the organism's fitness change on the timescale of generations but remain relatively stable within an individual's lifetime [85]. Furthermore, in the particular case of the evolution of associative learning, there must also be learnable cues that reliably correlate with the state of the environment [92]. In this situation, organisms may benefit if they use those cues to track current conditions and map them to appropriate responses. Since the environment and cues may change between generations, the mapping cannot be encoded genetically and must be learned during the organism's lifetime.

Researchers have explored the factors of environmental dynamics and cue availability that are necessary for the evolution of associative learning using both mathematical and empirical approaches [85], [92]. However, it is still an open question whether these factors are sufficient for associative learning to emerge during the evolution of an organism's behavioral repertoire. Here, we propose:

**HYPOTHESIS 1:** The initial evolution of associative learning depends on the scaffolding provided by the prior evolution of a repertoire of instinctual behaviors that exploit stable environmental patterns.

Skinner and others speculated that complex behavioral traits do not evolve independently of each other but build on preexisting ones according to a characteristic evolutionary sequence that starts with simple movement, then sensing, followed by tropisms and reflexes, and finally learning [57], [93]. Similarly, it has been suggested that different forms of learning are not independent but evolve from one another in a specific sequence, where more complex forms build on the mechanisms of simpler ones and subsume them [94]–[97]. For example, associative learning would have evolved from sensitization [94]–[97], a simpler, nonassociative, form of learning where an organism increases its response to a repeated stimulus [98]. Therefore, we propose:

**HYPOTHESIS 2:** Complex behaviors, including learning, do not arise and function independently from one another. Instead, as more complex cognitive processes arise, they do so in a modular and stepwise manner, where early instinctual behaviors (such as moving and sensing) are co-opted and integrated into increasingly more complex ones (such as error recovery or path prediction) before eventually reaching associative learning.

It has also been speculated that the emergence of associative learning required only minor modifications in preexisting memory mechanisms [80], [96], [97], enabling it to evolve in parallel in different species [80]. Thus, we propose:

**HYPOTHESIS 3:** Associative learning can arise suddenly, as a result of small modifications in preexisting cognitive mechanisms, as opposed to arising gradually and independently by accumulating incremental changes under selection.

Finally, given the expectation that environmental characteristics, such as stability and cue availability, shape the type of learning that evolves [85], [89], [92], we investigate an additional hypothesis on the flexibility of the associative learning mechanism that evolves in a particular environment. We propose:

**HYPOTHESIS 4:** Organisms that evolve associative learning will not be able to change established associations (e.g., reversal learning) unless such changes were necessary for success during evolution.

Our research focuses on a definition of associative learning that emphasizes its consequences for behavior rather than the mechanisms by which it works. We think this approach is justified because associative learning is traditionally defined in operational rather than mechanistic terms-for example, as "a behavioral modification, dependent on reinforcement, involving new associations between different sensory stimuli, or between sensory stimuli and responses" [80]- and may not even be a unitary behavioral trait with consistent properties across species. For example, it is by no means clear that associative learning involves distinct mechanisms from those underlying simpler, nonassociative forms of learning, such as habituation and sensitization. In *Drosophila*, mutants incapable of associative learning also show reduced habituation and sensitization [99], and in *Aplysia*, sensitization and associative learning share many of the same molecular elements [100]. It is also not clear whether there is only one way of implementing associative learning mechanistically. All animals in which associative learning has been well established have a central nervous system (i.e., brains) - although many animal groups have not yet been tested [80] - but having a brain is not necessary for associative

learning: plants are capable of it [101], and single-cell organisms may be as well [102], [103]. These observations suggest that associative learning has evolved independently, acquiring different properties in different lineages [80], [88], [90]. Hence, they also justify the assumption that we can study the evolution of associative learning as a phenotypic attribute of behavior that is independent of a particular mechanistic implementation.

Major challenges arise in studying the evolutionary origin of learning. One challenge is the utter lack of fossil evidence, especially from periods as remote as the Precambrian, when associative learning behavior is believed to have first evolved [80]. Another is the difficulty of performing phylogenetic comparisons to study the origin, as opposed to the adaptive function, of behavioral traits. Although phylogenies are valuable to infer ancestral character states, sequences, and timing of evolution of traits, this approach is virtually silent on the selective forces and mechanisms involved [104] and may suggest patterns of evolution that could result from multiple different processes [105]. In addition, associative learning is such a widespread and likely ancient behavior that it is particularly challenging to reconstruct an accurate phylogeny because of the lack of out-groups and because its origin presumably predates the rapid adaptive radiation of the Cambrian explosion [98]. The ubiquity of associative learning behavior among extant species is also a challenge for experimental evolution, which has been very successful in studying the adaptive modification of existing learning mechanisms in animals but can reveal little about the origins and early evolution of learning [91].

To overcome these limitations, here we study the origins of learning behavior in populations of self-replicating computer programs that undergo open-ended evolution in a virtual

environment [106]. These digital organisms are selected for their ability to cope with behavioral challenges in which associative learning may confer a fitness advantage; specifically, the environment provides alternative courses of action and cues that reliably correlate with the correct action, although these cues vary across generations [92].

This approach allows ample opportunities for a wide range of behaviors to evolve and enables the discovery of evolutionary principles that are potentially independent of the cognitive machinery that is undergoing evolution. We emphasize that digital evolution is not a simulation of evolution but rather an instantiation of it [107]: although digital organisms are evaluated in simulated environments, their behavioral control algorithm undergoes actual Darwinian evolution. Specifically, (i) organisms reproduce and pass on their evolved traits, including their behavioral algorithm, to their offspring; (ii) inheritance is subject to mutations, producing variation; (iii) individual fitness depends on an organism's performance at specific behavioral tasks and determines the outcome of the competition for space in a size-limited population. This approach enables true experimental study of evolutionary history across multiple replicate lineages evolving under different conditions, providing insights not only on the outcomes of evolution but also on the transitions that occur in different lines of descent. Digital evolution has a proven track record of expanding evolutionary theory [108]–[110] with supporting evidence often collected later in biological systems [111]. Previous studies in Avida have also demonstrated the evolution of instinctive navigation, such as gradient ascent and trail-following behavior [112], including the genetically encoded use of memory to dictate subsequent behavior [106]. Here, we extend this work beyond reflexive behaviors to study the evolution of

associative learning where each individual organism must discover a mapping between environmental cues and the optimal response.

Our results support the aforementioned hypotheses and, moreover, provide a rich picture of the circumstances that favor—or disfavor—the evolution of learning, including the critical role played by historical contingency. Learning is a rare outcome of evolution in our system, not because of any intrinsic difficulty in the underlying computation but rather because oftentimes lineages evolve highly flexible behavioral strategies over which learning does not provide a strong selective advantage. When learning does evolve, it emerges via an almost stereotypical sequence, as proposed by Skinner and others [57], [93]. Finally, we find that the evolution of behavior is inseparable from the evolution of an intrinsic value system, the innate gauge of an organism's experiences that provides positive or negative feedback on its actions.

## 2.2 Experimental System

We used the Avida digital evolution platform for all of our experiments [113], [114]. Avida is a linear genetic programming platform, meaning that each organism's genome consists of an ordered sequence of computer instructions in a machine-like language. Instructions are simple, self-contained operations, such as adding two values, storing a value in memory, or skipping to another instruction if one value is greater than another. During evolution, random mutations occur that can insert, remove, or replace instructions in offspring. Note that any sequence of Avida instructions can be executed; as such, mutations will always produce valid programs even if their functionality may be meaningless.

In addition to the instructions for arithmetical and logical operations described above, we used a single instruction that caused the organism to reproduce as well as a set of instructions that acted as simple sensors and effectors to interact with the environment (described in the next section). Using Avida provided key benefits for the experimental study of evolution. For example, the set of instructions we used formed a Turing-complete programming language that, in theory, can represent any algorithm—including any behavioral control algorithm—given the necessary sensors and effectors. In addition, it is easy to analyze an Avida organism to dissect and study the behavioral control algorithms that evolve. Furthermore, we can archive all ancestors and their evolutionary lines of descent to examine the evolutionary transitions that occurred along any lineage, allowing us to study patterns that reveal how one set of behaviors might potentiate another.

An Avida organism is defined by a sequence of instructions (its genome), and each particular sequence defines its genotype. In our experiments, each population was seeded with a "naive" organism that lacked any instruction for behavioral control other than the one necessary to reproduce. Such an organism's genome consisted of a sequence of null instructions that acted as placeholders for future behavioral "genes" and a single "reproduce" instruction. To reproduce, an organism had to execute a minimum number of instructions, that is, spend a minimum amount of time in the environment in order to mature. At the same time, an organism also had an upper limit in the number of instructions it could execute before it tried to reproduce, essentially creating a maximum age. If an organism failed to reproduce by the time this limit was reached, it was eliminated from the population. Reproduction was asexual and resulted in the production of two offspring, both inheriting a copy of the parent's genome.

However, only one of the offspring was subject to mutation, while the other remained identical to the parent and essentially replaced it.

Populations were capped at 3,600 organisms. Once that limit was reached, every organism that was born resulted in an existing one being randomly removed. Organisms did not interact with each other in the environment; however, the age limit and the competition for space in the size-limited population created a strong selection pressure for fast reproduction.

How well an organism performed the behavioral task determined the rate at which its offspring's instructions were executed and consequently how quickly they could reproduce. Therefore, the better an organism performed on the behavioral task, the faster its offspring executed their behavioral algorithm and reproduced. Thus, behaviors evolved in this digital system in a purely Darwinian fashion.

## 2.3 The Behavioral Task

Bees, ants, and other insects are known to use local and distant landmarks for navigation [106], [115]–[118]. For example, experiments have shown that bees can learn visual cue associations to successfully navigate complex mazes [119], [120]. Inspired by these experiments, the behavioral task that we presented to evolving Avida organisms consisted of navigating a trail of nutrients in a virtual arena (fig. 2.1), where nutrients provided cues that indicated the direction to follow — if organisms evolved the ability to sense and use them.



**Figure 2.1**: Sample Arena and Nutrient Trail. Shown is one of four virtual arenas from an environment. Each virtual arena contained a single trail of nutrients laid out in a unique configuration. At the beginning of its life, each organism was placed alone at the start of the trail (green circle) in a randomly selected arena and oriented in the direction of the next nutrient.

An organism's task was to complete as much of the trail as possible and then reproduce before the end of its life. The system kept track of the organism's cumulative performance by counting the number of new nutrient locations it visited and subtracting the number of empty (i.e., offtrail) locations encountered. This count was then divided by the total number of nutrients in the trail to compute the organism's "task quality," which ranged from 0 to 1 (negative values were set to zero). Nutrient locations were counted only on the first visit; subsequent visits to the same nutrient location would not affect an organisms' task quality. However, visits to empty locations were always deducted. The organism had no sensory feedback about its task quality (i.e., cumulative performance), similar to the way a natural organism cannot sense its own fitness. Nevertheless, our organisms are limited relative to natural ones that may be able to measure payoffs of their foraging decisions by the rate of some physiological condition, such as gut fullness [121].

Each environment in our evolutionary experiments consisted of four virtual arenas, each with a different trail configuration (fig. 2.1). Every time an organism was born, it was randomly assigned to one of the four arenas, placed at the beginning of the trail on a nutrient location, and oriented in the direction of the next nutrient. The use of four trail configurations reduced the likelihood of an organism evolving a rigid control algorithm tailored to a single nutrient trail (genetically hardwiring a sequence of actions) instead of a flexible control algorithm that captures the principles of trail navigation.

Each of our experiments consisted of between 50 and 900 replicates. At the end of an experiment, we selected the predominant (most abundant) genotype from each replicate's final population for behavioral analysis. Given the large population size, the predominant genotype typically represented dozens of organisms, implying that they, on average, outperformed the rest of the population on all four trail configurations. Indeed, in these experiments we found that the predominant genotype typically had the highest task quality scores on each of the four trails; thus, we measured its performance by computing the population's average maximum task quality (AMTQ) scores across all trail configurations.

An organism's interaction with the environment depended on sensor and effector instructions acquired through mutation and maintained during evolution. These instructions conferred the abilities to sense the nutrient content of the current location ("sense current"), rotate right by
45 degrees ("rotate right"), rotate left by 45 degrees ("rotate left"), take one step ahead ("move ahead"), and take one step back while facing forward ("move back").

The execution of a sense current instruction provided feedback, in the form of an integer, about the nutrient content of the location the organism occupied. Empty locations and nutrients, in both straight portions of the trail and at turn points, were each sensed as different values. Therefore, the numeric value of the nutrients could cue the organism to the direction of the trail once they evolved the ability to interpret the sensed values correctly. There were four types of cues: right turn (45 degrees), left turn (45 degrees), forward, and empty location (fig. 2.1). Nutrients that indicated forward (forward cue) and empty location were always represented by the integers 0 and -1, respectively. Meanwhile, nutrients that indicated turns (turn cues) were each assigned a distinct random number between 1 and 100 every time an organism started on a trail, and this assignment persisted only during the organism's lifetime. Since forward cues and empty locations had persistent values between generations, organisms could evolve to use them to predict optimal future moves. However, the environmental uncertainty represented by randomized turn cues created an additional challenge for the organism: it not only had to move and follow the trail, but it also had to identify the direction represented by the turn cues. The optimal way to overcome this challenge was for an organism, within its lifetime, to associate either the right turn cue or the left turn cue with the correct action, thus identifying the opposite turn cue by exclusion.

Although one cannot predict the course of evolution, if associative learning evolved in our experiments, we expected to recognize it by observing the path of the organism along the trail.

When placed on a new trail and allowed a period of exposure to the different turn cues, an organism capable of associative learning should be able to consistently turn to the correct direction every time it encounters a turn cue, something that would not be possible if the organism were using heuristics or choosing randomly.

#### **2.3.1 Experimental Conditions**

#### 2.3.1.1 Experiment 1

In experiment 1, we tested four different environments, each with four possible trail configurations (table 2.1; figs. A.1-A.4, Appendix A). In three of the environments, the trails of nutrients started with a simple (and presumably predictable) pattern (table 2.1). In the fourth environment, which served as a control, the trails provided nutrients in an unpredictable pattern - that is, each of the first two turns had equal probability of being to the right or to the left. This setup allowed us to test our first three hypotheses (as presented above). We performed 50 evolutionary replicates for each of the four conditions listed in table 2.1. See Section A.1 of Appendix A for additional details on methods.

**Table 2.1:** Environments for experiment 1

	F	Control environment		
	One fixed turn	Two fixed turns	Nutrient cued	Random start
Trail start pattern	The first turn was always to the right in all four trails (fig. A.1).	The first turn was always to the left and the second was always to the right in all four trails (fig. A.2).	The direction of the first and second turns was random in all four trails but could be predicted by counting the number of forward cues preceding the first turn (an odd number meant left, while even meant right; fig. A.3).	The direction of the first and second turns was random, and the number of forward cues preceding the first turn was the same in all four trails (fig. A.4).

**Note**: Each environment contained four different trail configurations. An organism experienced only one trail configuration in its lifetime. See section A.1 of Appendix A for images of each environment.

#### 2.3.1.2 Experiment 2

In experiment 2, we applied an additional selection pressure aimed at the evolution of reversal learning. We used only the nutrient cued environment but reversed the turn cues at approximately the 85% mark of each trail (fig. A.5). In a complementary experiment, we tested different cue reversal positions ranging between 10% to 90% in 2.5% increments and found that it did not affect the results significantly (Appendix A, sec. A.4; fig. A.8). Therefore, we report only the results for the 85% mark.

We performed 900 evolutionary replicates in experiment 2. The reason for the larger number of replicates than in experiment 1 was to generate a sufficient number of phenotypes for lineage studies, especially to explore the ancestry of the rare organisms that evolved reversal learning. A lineage study consists of singling out the final predominant organism of a population and reconstructing its line of descent, testing every ancestral genotype on the behavioral task to

uncover how the behavior evolved over time. Although this experiment was designed to test hypothesis 4, it also enabled us to obtain additional evidence relevant to hypothesis 3. See Section A.1 of Appendix A for additional methodological details.

Raw data, code, and a video associated with this research are available in the Dryad Digital Repository: (http://doi:10.5061/dryad.f45gh6s;(Pontes et al. 2020). The custom version of Avida used in this study is available at the following location:

https://github.com/mercere99/Avida-AssociativeMemory [114].

### 2.4 Results

# 2.4.1 Repeated Evolution of Adaptive Behaviors: Error Recovery, Imprinting, and Reversal Learning

Our experiments resulted in the evolution of organisms capable of adapting to unpredictable environments by using a variety of strategies, including associative learning. We also observed the evolution of flexible strategies that did not rely on learning (table 2.2). We called the most successful nonlearning strategy "error recovery," in which an organism attempted to follow the nutrient trail and, on stepping off the trail, performed the necessary actions to return to it but did not modify its future behavior based on the error. A particularly notable result was the repeated evolution of associative learning, including both a rigid form that we called "imprinting", and a more flexible form that we called "relearning" (described in table 2.2).

We also found recurrent patterns in the behavioral strategies that evolved. Organisms from different evolutionary replicates, which inevitably had genotypes producing distinct behavioral

control algorithms, generated a consistent set of behavioral phenotypes. We analyzed more than 300 out of 1,100 replicates across all experimental conditions and found, notably, that they all fell into five easily recognizable categories, including relearning, imprinting, and error recovery (previously mentioned), plus "searching" and "path predicting" (see table 2.2). We found some hybrids of these strategies as well.

	Behavioral Strategy	Description	Typical AMTQ
Learning behaviors	Relearning	A flexible and generalizable strategy, based on instrumental conditioning, that allowed organisms to navigate any trail configuration regardless of the starting pattern. Organisms using this strategy were able to re-form the cue-response association multiple times, even when the cues were reversed (fig. A.15).	95% - 97% (experiment 2; 97% - 99% in non-cue- reversal environment)
	Imprinting	A somewhat rigid strategy, based on instrumental conditioning, where organisms made the cue-response association early, and only once, in their lifetimes (fig. 2.2). These organisms were not able to relearn cues that were changed or reversed. This strategy was further classified into two subtypes: "generalizable" and "nongeneralizable." The generalizable subtype enabled organisms to imprint and navigate on any trail configuration regardless of the starting pattern, while the nongeneralizable subtype limited organisms to imprint on trails with the same starting pattern in which the organisms evolved.	97% - 99% (experiment 1; 87% - 92% in cue reversal environment)
Reflexive behaviors	Error recovery	A flexible and generalizable strategy that allowed organisms to navigate any trail configuration regardless of the starting pattern. These organisms did not discriminate between right and left turn cues. Instead, they reacted to turn cues by turning in one direction (each organism had its own default direction) and, if this direction led an organism off the trail, returning and trying the other direction (fig. 2.2).	78% - 84% (experiments 1 and 2)
	Searching	A strategy that appeared only in hybrid combinations with others, especially error recovery and path predicting. It was triggered if the organism stepped off the trail and typically involved performing several moving and turning steps to try to find another segment of the trail and rejoin it (fig. A.7).	48% - 60% (experiments 1 and 2)
	Path predicting	A strategy where organisms encoded behavioral sequences in their genomes that matched the initial portion of different trails (fig. A.12). This strategy enabled them to successfully navigate the first few segments of any trail but not the entire trail.	7% - 14% (experiments 1 and 2)

## **Table 2.2:** Behavioral strategies found in all experiments

Table 2.2 (cont'd).

**Note**: We analyzed more than 300 organisms across all experimental conditions. Although the details of their behavior differed among experiments, all behavioral phenotypes could be classified into five strategies or hybrids of two or more. The performance results for each behavioral strategy fell into a typical performance range, measured by average maximum task quality (AMTQ), and ranked from highest to lowest.



**Figure 2.2:** Two top-performing strategies in experiment 1. Shown are the paths of the final predominant organisms from two different replicates that evolved in the nutrient cued environment in experiment 1. Both were tested in the same trail configuration to facilitate comparison. In the left panel, an organism using an error recovery strategy achieved a task quality score of 81% of the maximum. Starting from the green circle, it moved straight while sensing forward cues but always tried to turn right (45 degrees) when sensing a turn cue. If turning right led the organism into an empty cell, it would retreat to the previous position and turn toward the left (90 degrees). It continued to repeat this behavior at every turn cue without ever learning from its error. In the right panel, an organism from a separate replicate using a generalizable imprinting strategy achieved a task quality score of 98% of the maximum. It also tried to turn right when sensing a turn cue. However, it stepped off the path only once at the first left turn. It learned the correct cue-response association and navigated the remainder of the trail without error.

The specific type of associative learning that evolved in our experiments was "instrumental conditioning", in which an organism forms an association between a stimulus and a behavior from its repertoire [80]. Organisms that performed imprinting formed an association early in their lives that was used for future decisions but could never be modified. Organisms that performed relearning also formed associations between cues and actions early in their lives but were able to form new associations if the cues changed, regardless of whether they were swapped or replaced with novel ones. Additionally, we identified environmental factors and historical constraints that strongly influence whether associative learning evolves.

The ability to relearn when cues are swapped is called "reversal learning," a learning ability that is sometimes regarded as cognitively complex [123]–[125]. A typical organism capable of reversal learning followed the trail of nutrients until it encountered a turn cue. Since the integers representing turn cues were randomly assigned for each generation, the organism then attempted to turn 45 degrees in a default direction and move forward one step. If this step led to a nutrient-containing location, the organism continued to follow the trail. However, if the organism turned in the "wrong" direction and found itself on an empty location, it engaged in a corrective reaction by taking one step back and turning 45 degrees twice (90 degrees) in the opposite direction (as if recoiling and turning away). The organism then made the association between the turn cue and correct action, such that the turn cue alone was sufficient to trigger the correct action in subsequent encounters. If the turn cues remained consistent, the organism navigated the remainder of the trail without error. Alternatively, if the turn cues changed further along the trail (including cue reversals) the organism again exhibited the corrective reaction and updated its association to the new cue, resuming the navigation

without further error. Cues could be changed any number of times with the organism always relearning the turn cue and navigating without error until the cue changed again or it reached the end of the trail (fig. A.15).

This serial reversal learning behavior evolved repeatedly, although it was a rare outcome, evolving in only 10 out of 900 replicates in experiment 2, where we specifically selected for reversal learning (and not at all in the 200 replicates in experiment 1 where reversal learning was not actively selected). Nevertheless, many replicates that did not result in the evolution of reversal learning still produced organisms that were able to efficiently navigate the entire trail using either imprinting or error recovery.

#### 2.4.2 Early trail predictability produces behavioral building blocks for learning

Although all environments could promote the evolution of simple controlled movement, not all could lead to the evolution of learning. All environments were constructed in a way that could potentially select for behavioral biases, such as moving along the trail of nutrients and avoiding empty locations, that contributed to the overall task performance. Indeed, the very first behaviors to evolve were simple forms of controlled movement, such as oscillatory behavior (moving back and forth) and moving to an edge of the path segment and stopping (see Appendix A, sec. A.4.1 for an example).

In addition, all environments provided organisms with the features thought necessary for learning to evolve: frequent choices of actions (move straight, turn right, or turn left) and cues that change each generation but reliably indicate the best choice within a generation [91], [92],

[126]. However, while these features were present in all environments, they proved insufficient to evolve learning. Specifically, no replicates in the random start environment produced organisms capable of learning (or even error recovery). In fact, none of the organisms from this environment were able to navigate past the first turn, and their task quality remained at or below 4% of the maximum across all 50 replicates. The environments in which learning did evolve (i.e., one fixed turn, two fixed turns, and nutrient cued) all had a property that the random start environment lacked: trails providing a high initial degree of predictability across generations that enable organisms to evolve behavioral building blocks and navigate the trail reflexively before evolving learning [57]. These building blocks include moving repeatedly, sensing the current cue, distinguishing the different cues and reacting to them, turning to either side, retreating to the trail when an empty location is sensed, storing a cue in memory, and comparing the current cue with the one in memory. This result supports hypothesis 1.

The predictable-start environments (one fixed turn, two fixed turns, and nutrient cued; table 2.1; figs. A.1-A.3) were the only ones to evolve complex behaviors, including learning. These environments also produced a wider range of navigational strategies and organisms that reached substantially higher task quality than any organism in the random start environment (fig. 2.3; table 2.3). The nutrient cued environment produced the largest proportion of organisms that could navigate the entire trail, followed by the one fixed turn and the two fixed turns environments. These organisms used imprinting, error recovery or a hybrid strategy (table 2.3). The organisms that achieved at least 25% AMTQ but did not complete the trail used the

same strategies but performed more slowly, or simply reproduced before reaching the end of the trail (table 2.3).



**Figure 2.3:** Distribution of average maximum task quality (AMTQ) per environment for experiment 1. Each violin plot represents the distribution of AMTQ across replicates for a given environment. Only the environments that started with a predictable pattern (one fixed turn, two fixed turns, and nutrient cued) evolved organisms that could finish the trail. They also produced a wider range of navigational strategies and organisms that reached much higher task quality than the control environment (random start). Table 2.3: Experiment 1: summary of results

		Prec	Control environment		
		One fixed turn	Two fixed turns	Nutrient cued	Random start
Replicates in which organisms finished the trail	Proportion of replicates	18/50	13/50	23/50	0/50
	Strategies evolved	Imprinting (3)	Imprinting (5)	Imprinting (3*)	
	(no. replicates)	Error recovery (15)	Error recovery (7) Hybrid of path predicting and error recovery (1)	Error recovery (20)	NA
	Highest AMTQ observed (strategy)	99% (imprinting)	99.7% (imprinting)	99% (imprinting)	NA
	Proportion of replicates	9/50	4/50	4/50	0/50
Replicates in which organisms did not finish the trail (AMTQ ≥ 25%)	Strategies evolved (no. replicates)	Imprinting (1) Error recovery (8)	Imprinting (2) Error recovery (1) Hybrid of error recovery and searching (1)	Error recovery (1) Path predicting (1) Hybrid of error recovery and searching (1)	NA
23701				Hybrid of path predicting, searching, and imprinting (1)	

**Note**: Shown are the performance and strategies of the organisms with average maximum task quality (AMTQ) equal to or higher than 25%, organized by environment. We examined only a sample of organisms that had less than 25% AMTQ. Those that were examined displayed previously described strategies and did not travel far on the trail. NA = not applicable.

\* Two of these organisms performed a generalizable version of the imprinting strategy that allowed them to navigate any trail configuration independently of the starting pattern.

#### 2.4.3 Learning May Not Generalize to Novel Environments

Both imprinting and error recovery were successful strategies in experiment 1, but they differed in how well organisms could generalize to novel trails. Organisms that used error recovery did not depend on the pattern at the start of the trail for their navigation and could finish any trail configuration that we tested (fig. 2.2).

In contrast, most of the organisms that used imprinting depended on the specific start pattern from the environment in which they had evolved to form the cue association. When tested in trails with a different start pattern, these organisms were not able to navigate far and scored poorly in task quality. However, two replicates in the nutrient cued environment evolved a generalizable version of imprinting that allowed the organisms to navigate any trail configuration independently of the starting pattern. These organisms began navigating the trail and, when sensing a turn cue, turned to a default direction. On making their first wrong turn and stepping off the trail, these organisms used error recovery to step back onto the trail and turn to the other direction. At this point, they imprinted the turn cue that led them astray and used the learned association to navigate the remainder of the trail (fig. 2.2). However, when tested in trails containing cue reversals or replacements, these organisms were not capable of coping with such changes and made wrong turns and stepped off the trail. They then resorted to using error recovery to get back on the trail and continue navigating until the end. This result led us to propose hypothesis 4, namely, that the environment has to present cue reversals along the trail to foster the evolution of more "complex" learning abilities, such as relearning and reversal learning [123]–[125], a hypothesis that we tested as part of experiment 2.

#### 2.4.4 Cue Reversals During Evolution Foster Ability to Relearn During Lifetime

In experiment 2, we used only the nutrient cued environment because it was the only one where generalizable imprinting evolved in experiment 1. At approximately the 85% mark of each trail we swapped (reversed) the values associated with the turning cues, requiring the organism to learn to turn in the opposite direction of the one it learned at the beginning of the trail. We named this condition the "cue reversal" environment (fig. A.5). In a complementary experiment, we tested varying the cue reversal position between the 10% and the 90% mark, without any significant effect on the results (Appendix A, sec.A.4; fig. A.8).

The results support hypothesis 4, although, as in previous experiments, the evolution of a complex learning ability proved to be a rare occurrence. Of 900 replicates, only 18 evolved the capacity for any form of reversal learning (fig. 2.4). In 10 of these 18 replicates, organisms also evolved the capacity for serial reversal learning, even though their ancestors had only experienced a single cue reversal in their lifetimes. In a serial reversal learning trial, the agent is confronted by a repeated reversal of a two-symbol combination. Organisms from the 10 replicates that could perform this task exhibited behavior that generalized to any trail configuration. In the other eight replicates, organisms evolved at least some capacity for reversal learning and relearning. However, their behavior had limitations, such as (*i*) being able to learn certain pairs of cues and not others, (*ii*) generalizing their behavior to some novel trail configurations and not others, or (*iii*) having a "short memory", which is "forgetting" the association after a while and needing to learn it anew on making a wrong turn. These

limitations led to failures in staying on the trail, and in these cases organisms got lost or stuck outside the trail or resorted to navigating by error recovery or searching.



**Figure 2.4:** Distribution of average maximum task quality across 900 replicates. The performance histogram of all final predominant organism in experiment 2 reveals a marked grouping by behavioral strategy. Organisms in groups 1 and 2 did not finish the trails, while those in groups 3, 4 and 5 did. Group 1 consisted mainly of organisms that navigated by path predicting and its hybrids. Group 2 consisted mainly of organisms that navigated by error recovery, imprinting, and their hybrids. Group 3 consisted mainly of organisms that navigated by more effective forms of error recovery. Group 4 consisted mainly of organisms that employed imprinting hybrids. Group 5 consisted mainly of organisms. Those of groups 4 and 5 were assessed from all organisms.

As in experiment 1, the fittest organisms (based on task quality) that evolved in experiment 2

were those that used learning strategies. The organisms capable of relearning scored as high as

97% of the maximum and were the fittest overall. Their behavior was similar to the

generalizable imprinting that evolved in experiment 1, in that they made the association

between the cue and the correct action on stepping off the trail, but they were also capable of

relearning if a cue reversal led them off the trail. Intriguingly, these organisms could also relearn when tested in environments where an initial pair of turn cues was replaced by a completely different pair as well as when they were reversed or changed multiple times along the trail, even though we did not specifically select for this form of flexibility.

The next fittest organisms that were capable of learning employed various hybrid strategies involving imprinting, error recovery, and path predicting to reach task quality scores as high as 93% of the maximum. Although incapable of relearning per se (i.e., replacing a cue association with another), they were able to form temporary associations (short-term imprinting). This "short memory" gave the organisms the opportunity to form a new association after the previous one had extinguished. This hybrid strategy turned out to score higher in task quality than imprinting or error recovery alone. For additional results and a "bestiary" of evolved behaviors, see sections A.3 and A.4 of Appendix A.

#### 2.4.5 The Stepwise Evolution of Learning

We found a discernible pattern in the evolutionary trajectories of the organisms that evolved learning strategies (relearning and imprinting). Despite the organisms having evolved completely independently, these lineages passed through a characteristic sequence of phenotypic stages corresponding to two or more of the categories we described in table 2.2.

We analyzed the ancestral lineages of all of the final predominant organisms that evolved imprinting in experiment 1 and ten of the final predominant organisms capable of relearning in experiment 2. Starting from a sessile common ancestor, all lineages first evolved the capacity for moving, then sensing, followed by reflexive navigation and then learning, a result that supports hypothesis 2. In addition, error recovery preceded the evolution of associative learning in all of the lineages where the final predominant organism made the cue-response association by stepping off the trail (generalizable imprinting, and relearning). In lineages where imprinting evolved directly from path predicting, the final predominant organisms were not capable of error recovery, and their behavior did not generalize to other trail configurations (nongeneralizable imprinting; figs. 2.5,2.6).



**Figure 2.5:** Evolutionary history: 10 Lineages. Shown is the evolution of task quality over time in each of the 10 lineages that were ultimately capable of serial relearning from experiment 2. As they transitioned to a new strategy, some lineages had great gains in task quality, while others had more gradual ones. All lineages, however, went through occasional periods of fitness loss. Different task quality ranges often corresponded to specific behavioral strategies. Range 1 corresponded to path predicting, range 2 corresponded to hybrid strategies that included searching, range 3 corresponded to error recovery, and ranges 4 and 5 corresponded to imprinting and relearning.



**Figure 2.6:** Commonly observed evolutionary sequences. Shown are the evolutionary trajectories of the 11 lineages that evolved associative learning in experiment 1, and the 10 lineages that evolved serial relearning in experiment 2. Behaviors evolved in a characteristic sequence of phenotypic stages. Starting from a naive and sessile common ancestor, all 21 lineages evolved the capacity for moving, then sensing, followed by reflexive navigation and then learning. The numbers next to the arrows indicate how many lineages followed a particular pathway, with thicker lines indicating more common evolutionary pathways in relation to alternatives.

#### 2.4.6 Learning Can Evolve Suddenly

Finally, we found that during evolution, the transitions from one strategy to another could occur abruptly, often as a result of a single mutation. This is not to say that a single mutation was sufficient to produce a new strategy but rather that the new strategy often evolved silently over a great number of generations until one or a few mutations triggered the transition in behavior, a result that supports hypothesis 3. Sometimes this evolutionary transition would give the organism a large fitness advantage, and its descendants would sweep the population. For example, the transition between error recovery and associative learning (imprinting or relearning) always occurred in one generation (we never observed any instance of a behavior that would be intermediary, such as a simpler form of learning). In one of the lineages we analyzed, the transition from error recovery to relearning raised the AMTQ from 81% to 98% in a single generation. This strategy transition was triggered by a single mutation that changed the flow of the algorithm so that after an error recovery event, the value of the currently sensed cue would be stored in memory (figs. A.16, A.17). The remainder of the error recovery process stayed intact and was subsumed by the newly acquired relearning capacity. Other components of the relearning algorithm, such as the module for storing the cue in memory, had already been part of the ancestor for many generations but were not used or did not affect the organism's task quality. This result represents a clear case of historical contingency, where one or more modules had to be in place before new mutations could lead to a fitness gain [127], [128]. See section A.4.1 of Appendix A for figures and phenotypic descriptions of the major evolutionary transitions in this lineage.

### **2.5 Discussion and Conclusions**

#### 2.5.1 Emergence of Learning Depends on the Prior Evolution of Reflexive Behaviors

Most studies of the evolution of learning have focused on the selection pressures that may act to increase or decrease an organism's reliance on learning [86], [91], [92]. Our study complements and extends this work by examining how learning may have first arisen. As Dunlap demonstrated [91], [92], [126], learning is favored in environments that present alternative courses of action, where the best action cannot be predicted at the beginning of an organism's life but environmental cues exist that reliably correlate with the best action. However, we found that although all of the environments possessed those presumably necessary qualities, they were not sufficient for learning to arise, as evidenced by results from the random start environment. Instead, as hypothesis 1 predicts, for organisms to initially evolve the capacity for learning, they must first accumulate simple behavioral building blocks to cope with the environment reflexively. In the cases presented here, these building blocks include an ability to move, to sense different cues, and to perform a range of actions (move forward, turn left or right, step back) in response to different cues. Crucially, generalizable learning (generalizable imprinting, and relearning; table 2.2) arose only in lineages that first evolved a reflexive ability to correct for missteps and return to a trail of resources (error recovery; table 2.2). With these reflexive behaviors in place, associative learning can then evolve because it confers an advantage by enabling an organism to modulate its behavior based on experience. Moreover, we find that reflexive and learning behaviors are shaped by different characteristics of the environment - the former by regularities that are stable across

generations, and the latter by patterns that vary across generations but persist for periods within an organism's lifetime. The most flexible learning ability – relearning new cue associations multiple times during an organism's life – depends on specific selection for it (i.e., swapping cues within the individual's lifetime as in experiment 2), as we proposed in hypothesis 4.

#### 2.5.2 Stepwise and Modular Evolution of Complex Behaviors

Across many replicate evolutionary runs in several experimental conditions, we found an almost stereotypical historical sequence leading to the ability to learn. Furthermore, our results are consistent with the idea that behavioral control algorithms evolve modularly [129], where more complex behaviors evolve by building on simpler ones and sharing their mechanisms. For example, learning mechanisms incorporated previously evolved error recovery behavior (figs. A.16, A.17). This result supports our hypothesis 2, originally proposed by Skinner and others for natural organisms [57], [90], [93], that learning abilities evolve by building on previously evolved reflexive behaviors.

However, in contrast to Skinner's model, we found that not all intermediate modules have an immediate survival value. Such is the case with the previously mentioned organism that evolved relearning from error recovery in a sudden transition triggered by a single mutation (figs. A.16, A.17), and whose error recovery ancestors had already acquired the capacity to store the cue in memory but never used this ability and, therefore, gained no fitness benefit. Only when a mutation connected the memory-storing with the error recovery module did the organism acquire the capacity to learn, thus gaining fitness.

It is important to clarify that no single Avida instruction or even specific set of instructions could bestow learning on an arbitrary nonlearning organism. All associative learning algorithms we observed were assemblies of many instructions that had to be executed in the proper order for learning behavior to manifest (see sample learning organisms; fig. A.18). That a single mutation could activate this behavior in an offspring only demonstrates that the remainder of the mechanism was already in place either as part of the existing behaviors or as neutral instructions [109], [130].

In the eleven lineages leading to imprinting in Experiment 1, and in the ten lineages leading to serial relearning in experiment 2 (fig. 2.5), we routinely found that complex abilities evolved from simpler ones in sudden transitions triggered by just a few mutations. This finding supports our hypothesis 3, that learning may arise through minor modification of existing mechanisms, and also lends credence to the proposition that something similar could have happened among natural organisms leading up to the Cambrian explosion [80].

More generally, these sharp transitions in phenotype are a consequence of the modular evolution of behavior. Modularity inherently reduces the requirements for evolving a new trait if it can build on existing ones, increasing phenotypic complexity with relatively modest genetic modifications [131]–[133], which can build up silently and, once completed, cause a sudden shift in phenotype.

#### 2.5.3 Why Learning Was Rare

Despite striking regularities during the course of evolution, associative learning was actually a rare outcome even in environments that fostered it (7% of lineages in experiment 1, 2% of lineages in experiment 2). Our results suggest some possible explanations. First, as mentioned above, complex behaviors can be hard to evolve, in part because they may depend on the preexistence of reusable intermediary modules - including features without survival value - and are therefore subject to the stochasticity of historical contingencies in general. Another possible explanation is that a reflexive strategy involving error recovery may already confer high fitness, such that the fitness gain associated with a learning strategy may not be enough for learning to arise and spread in the population. Across evolutionary replicates, we found organisms that could solve the problem in surprisingly different ways and obtain high levels of fitness, even in these simple environments. Furthermore, there can be implicit costs on more complex algorithms, including greater mutational fragility and longer processing times. Even making more mistakes on the trail, a shorter, sufficiently faster algorithm could reproduce more quickly and thus outcompete more complex algorithms that made fewer mistakes but executed too slowly. Surprisingly, we found in a follow-up experiment (Appendix A, sec.A.5) that the amount of computational memory available to an organism is not a constraint on the evolution of learning in our system as long as the minimum amount necessary to solve the task is provided. We performed a version of experiment 2, where we reduced the amount of memory available in the organism's CPU from 26 integers to 2, which is the minimum necessary to solve the learning task, but did not see a significant difference in the frequency of evolution of the relearning strategy or in the average task quality and distribution of task quality in the

final population compared with Experiment 2 (fig. A.19; table A.3). Overall, the same conditions that explain the rarity of solutions involving learning were also responsible for the variety of solutions and evolutionary paths we observed, better resembling natural evolution, where learning typically entails some kind of cost and is not always adaptive [23], [91], [93], [126], [134].

Interestingly, the stepwise succession of behaviors observed in our lineage studies, in conjunction with the diversity of final strategies from different replicates, are reminiscent of how behaviors appeared on trace fossils from the Ediacaran and early Cambrian, becoming more complex and diverse over time [135].

#### 2.5.4 The Scientific Value of an Open-Ended Evolutionary Model

In comparison with prior studies of the evolution of learning using computational methods [136]–[138], ours is striking in the open-endedness of the evolutionary process, which parallels that of biological evolution. Avida [113] employs relatively neutral genetic building blocks consisting primarily of algebraic and logic instructions, which do not constrain or favor the evolution of any particular behavioral algorithm. Thus, we were able to explore a large solution space and gain insights into the evolutionary dynamics that are also likely to occur in natural open-ended systems, even though nature uses very different building blocks. In Avida, the sheer number of potential solutions creates evolutionary dynamics and patterns that are not possible to observe using simpler digital evolution systems. For example, Izquierdo's groundbreaking work on the evolution of associative learning using neural networks consisted of evolving only the connections between preexisting neurons [137], [138]. Although many

insights were gained from that experiment, the limited number of potential solutions also led to a smaller diversity of outcomes. In fact, simply using neural networks, which are intrinsically designed to form associations, means that fewer mutational steps are needed from a starting point to evolve appropriate connections compared with the enormous search space in Avida. In our experiments, the behavioral algorithms evolved from scratch, using the most basic computer programming language elements, from ancestors incapable of sensation, movement, or navigation of any kind.

Furthermore, in our system even a basic behavioral building block, such as the move back instruction, could evolve from an assembly of simpler actions, as we demonstrate in a preliminary experiment (Appendix A, sec. A.2). We performed a version of experiment 1 without the move back instruction, and even without it many organisms evolved the capacity to navigate the entire trail using either imprinting or error recovery. These organisms evolved behavior functionally equivalent to the move back instruction by assembling other instructions from the basic instruction set.

#### 2.5.5 Early Evolution of an Intrinsic Value System

An unexpected outcome of this study was that it provided insights into the evolution of motivational mechanisms, which are thought to be integral to adaptive decision-making [93], [139]–[143]. Some of the earliest building blocks to evolve across all of our experiments were those responsible for evaluating experiences. In our system, evaluations were implicit features of the evolved controller and not distinct modules for deciding "good" or "bad." They were also essential to behavior control, since organisms could not sense their own task quality scores to

determine whether an action was beneficial or harmful. Early in evolution, values started as arbitrary biases, such as moving constantly or favoring turning one way or another, but biases that proved adaptive (e.g., preferring continuous movement while avoiding empty or previously visited locations) would fix, excluding less fit alternative biases. Over time, an intrinsic value system evolved that ensured appropriate behavior in response to specific inputs, and when associative learning arose, this value system provided reinforcement for behavior conditioning.

We can thus reinterpret the associative learning mechanism we have observed in the light of a value system: when an organism capable of learning senses an empty location, it displays the avoidance behavior because, in effect, it *negatively values the experience*. It associates this negative experience with the cue that led it to the empty location, and from then on, experiencing the cue alone is sufficient to activate the avoidance behavior.

#### 2.5.6 Reversal Learning Seems No More Complex than Initial Learning

Reversal learning is often deemed more challenging cognitively than initial learning [123]–[125]. However, in our experiments organisms that evolved the ability for reversal learning showed no difference in the capacity or speed of learning between the initial and subsequent learning events. Thus, once reversal learning evolves, it does not seem cognitively more complex than associative learning itself, at least in this system. Our point is not to undercut the study of how serial reversal learning in animals becomes faster with experience and correlates with cognitive flexibility [144]–[146] but to call for a refinement of the ideas around what is required for reversal learning to occur.

#### 2.5.7 How Evolution Continues to Shape Associative Learning

In a follow-up analysis (Appendix A, sec. A.6), we looked into how evolution continued to shape learning after it appeared in a lineage. In the lineages that produced associative learning in experiments 1 and 2, we found that learning ability would become attuned to the environment of evolution in a variety of ways. For example, (i) ancestral organisms that could learn some cue combinations but not others would eventually give rise to descendants that could learn any cue combination; (ii) ancestral organisms that required multiple exposures to learn the cueresponse association gave rise to descendants that required fewer exposures and, ultimately, final organisms that required only a single exposure; and (iii) in environments without cue reversals, ancestral organisms that could re-form associations multiple times gave rise to final organisms that could imprint only once.

These adaptations are consistent with the literature on preparedness and other so-called constraints on learning [84], [88], [89], [147], as well as the literature on sensitive periods of plasticity [148]–[150]. The key themes are that evolution produces learning mechanisms that are optimized for the needs of an animal in the environment where it evolved, and since learning is costly, evolution will often restrict the periods of an animal's life when it is most capable of learning (sensitive periods). An example of learning optimization is when an animal that relies on odors for foraging can learn more quickly to associate odors with good or bad foods than visual cues with the same foods [151]. The phenomenon of sensitive periods for learning is illustrated by filial imprinting in birds, where a chick learns who its mother is early in life and that association does not change [149].

Consistent with this literature, the imprinting strategy was adaptive in experiment 1, where there were no cue reversals. In that environment, ancestral organisms that were capable of reforming the cue association multiple times eventually gave rise to organisms that could form the association only once, presumably becoming more efficient. The sensitive period for learning in those lineages became restricted to the beginning of an organism's life. Meanwhile in experiment 2, where the environment contained cue reversals, the ability to re-form the cue associations (relearning and short-term imprinting) was adaptive, and the sensitive period for learning lasted an organism's entire life.

Although our experiments were not designed to investigate these topics, the patterns we found suggest a future area of study in which Avida is used to systematically explore how the evolutionary environment can constrain and optimize learning abilities.

#### 2.5.8 Implications for Artificial Intelligence

The insights of this study are relevant to the field of Artificial Intelligence, where lifetime learning has long been a challenge. We demonstrated that adaptive autonomous agents, capable of learning and navigation, can be produced by evolutionary methods, using biologically consistent scenarios where the environment fosters the evolution of learning and decision making, instead of traditional methods based on human design, which are difficult to scale up and to apply to novel tasks. One of our future goals is to extend this study and test whether we can evolve more complex forms of learning, such as contextual learning and rule learning (the learning of rules and concepts), and see whether their evolution follows the same sequence suggested in the literature [57], [90], [93]–[97]. We could test this hypothesis by

introducing additional cue types and requiring the organism to perform additional tasks in more intricate trails.

#### 2.5.9 Implications for the Evolution of Behavior

Finally, we believe that the evolution of learning in a digital environment would be useful to investigate the effect that learning behavior has on evolvability and rate of adaptive evolution. Some researchers have proposed that learning increases evolvability, since behavioral flexibility shields organisms from some selective pressures, allowing the population to maintain its diversity to cope with future selective events [81], [83]. Others have proposed that learning could either drive evolution by helping organisms adapt to different niches, where they would experience different selective pressures leading to change, or inhibit it by protecting them from selective pressures, leading to stasis [79]. It has even been suggested that the emergence of learning drove the diversification of complex behavior during the Cambrian explosion [80]. Overall, we agree with the remarkable assertion by B. F. Skinner [57] (p. 220) that understanding "the conditions under which [learning] evolved are helpful in understanding its nature.

# Chapter 3: Evolution of Patch Harvesting, an Insight into

# **Early Bilaterian Cognition**

Authors: Anselmo C. Pontes, Ian Whalen, Charles Ofria, and Fred C. Dyer

### 3.1 Introduction

#### 3.1.1 Background

The earliest evidence of animals with a nervous system are trace fossils from the Ediacaran Period (635–541 Ma) [152]. Trace fossils, also known as ichnofosils, are fossilized marks made by an organism, not including the fossilized remains of the organism itself [153]. Burrows of bilaterian animals feeding on microbial mats that once covered shallow marine floors have been preserved in sandstones all over the world, going back as far as approximately 560 Ma [152], [154]–[156]. These fossilized trails are horizontal and curvilinear and typically between 1 and 3mm wide [157], indicating the minute scale of the organisms that made them (fig. 3.1). Later traces, however, can be as wide as 5mm<sup>17</sup>.

<sup>&</sup>lt;sup>17</sup> Some of the simplest horizontal curvilinear trace fossils of the same era were left by deposit feeders, animals similar to those that fed on microbial mats, but that lived in deep waters below the photic zone and fed on detritus covering the seafloor [157]. However, the more elaborate trail behaviors are believed to be the product of shallow water microbial mat grazers, and it is on those animals that we focus our research.



**Figure 3.1:** Ediacaran trace fossil. The millimeter wide traces were produced by bilaterians with a nervous system, the first animals with these characteristics on record. Photo by Verisimilus at English Wikipedia, CC BY 2.5, Image downloaded from: https://commons.wikimedia.org/w/index.php?curid=2502886.

The worm-like animals that left these traces are referred to as mat miners or undermat miners [157], [158], after where they fed. Their horizontal burrows were confined to the thin layer of microbial mat between the water, which was possibly low in oxygen, and the toxic and anoxic substrate [152], [157], [159]. These animals were presumably composed only of soft tissues and as such did not leave body fossils (though what are believed to be body impressions of such an animal have recently been found [160]). They were likely closely related to the last common ancestor of today's bilaterians [152], [154], [160], which includes all animals with a central nervous system.

From when they first appear at the end of the Ediacaran and into the early Cambrian<sup>18</sup> [154], these trace fossils show increasingly complex patterns of activity and feeding strategies [152], [157], [161]. They document a critical period when early benthic animals transitioned from sessile to active lifestyles, which included burrowing, grazing, and hunting [155], [162]. They also show a change in feeding strategies, from osmotrophy and filter feeding to heterotrophy [154], [155]. This progressive increase in the complexity of the animals' feeding behavior could indicate an increase in the capacity and sophistication of their sensory and cognitive systems [157]. By studying the changes in these animals' feeding behavior, we can gain insight into the early evolution of animal cognition and nervous systems. We can also better understand the conditions that led to the Cambrian explosion [154], [163].

#### 3.1.2 Previous Research

Researchers have identified some regularities and trends in the fossilized burrows. The earliest traces are simple, showing little directionality, but over time they become increasingly elaborate and compact, showing evidence of more complex responses, such as avoiding crossing other burrows [157], [161]. They also form recognizable motifs, such as spirals and tight parallel meandering, like plowing a field [157], [161], [164]–[166].

According to the earliest interpretations of the fossils, the behavior of these animals was governed by a set of genetically encoded preferences (taxes): "phobotaxis", avoiding crossing burrows; "thigmotaxis", preferring to move alongside an existing burrow; and "strophotaxis", a

<sup>&</sup>lt;sup>18</sup> The Cambrian Period is from 541 to 485.4 Ma.

propensity to make 180° turns after traveling a certain distance [161], [167]. Another, more recent, interpretation is that the "taxes" we observe are not hardwired behaviors; instead, they reflect the animal's response to the concentration and distribution of food in the environment [153]. The animal therefore only has a single taxis, which is to move along the food gradient.

The differences in trail morphology over geological time are frequently considered evidence that the microbial mats these animals fed on were in decline and becoming patchier [157], [166]. Some researchers who favor the hypothesis of multiple hardwired taxes suggest that the animals responded by evolving more efficient grazing behaviors. While those who favor the hypothesis of a single taxis suggest that the animals' behavior did not necessarily change; the more compact trails are artifacts of animals following nutrient gradients in shrinking microbial mats.

Researchers from both sides of the taxes argument have built computer simulations based on their hypotheses [153], [157], [168], and both sides have been successful in demonstrating that, under their set of conditions, pre-programmed agents can reproduce harvesting patterns similar to those found in the fossil record. The reasons for patch harvesting trails becoming more complex and compact over evolutionary time have, therefore, remained disputed.

#### 3.1.3 Recent Evidence

The Ediacaran Period presented great environmental challenges for life, with great swings in global temperature and oxygen concentrations [152], [155], [159]. Furthermore, recent studies indicate that the ecology of the shallow marine environments from which most of the trace

fossils come was more dynamic and heterogeneous than previously thought [154]–[156], [169]. The widespread matgrounds covering the sea floor were composed of a variety of species, including sessile animals, algae, and microbial communities, as well as decaying matter [154], [155]. These microbial mats were prone to disturbance from waves, storms, coastal erosion, etc., as well as the activity of competing mat miners [155], [156]. It is therefore almost certain that the microbial mats were already broken into patches when the first mat miners appeared.

The activity of mat miners was likely further restricted by the concentration of oxygen, as mat miners would only graze where there was both food and oxygen [154], [159], [170]. Oxygen concentrations during the Ediacaran Period were much lower than today and varied greatly over time [155], [159]. Among the species that made up the microbial mats, however, were photosynthetic organisms; these created well-oxygenated oases where animals could thrive [159], [171], [172].

Regarding the suggestion that microbial mats were in decline, recent evidence indicates that shallow water matgrounds were actually increasing in abundance and diversity at the same time that the mat miners' burrows were becoming more sophisticated [155]. There is no consistent evidence that microbial mats were in decline before the Cambrian Period [154], [155], [173].

#### 3.1.4 Our Experiments

Previous computational experiments that attempted to recreate nutrient harvesting behavior were primarily based on pre-programmed (human-designed) agents whose behavior would

vary depending on a few calibration parameters [168]. These agents were also capable of peripheral, and sometimes, long range sensing [153]. Few experiments attempted to use evolutionary algorithms to evolve the agent's behavior, and those that did either failed to reproduce the more complex behavior patterns that we observe in the fossil record [174], or else they did not evolve the agent's control system itself, instead evolving calibration parameters for human-designed agents [175], [176]. As such, these studies could provide only limited insight into the selective pressures, the evolutionary relationships among the different burrowing behaviors, and the landscape of potential burrowing strategies. Two good reviews of these modeling efforts can be found in Hayes, 2003 [164] and Plotnick, 2003 [177].

For our computational experiments, we chose to once again use the Avida platform because of its successful track record in similar studies involving both the evolution of navigation and nutrient-gathering behaviors, such as gradient ascent and trail following [106], [178]. Moreover, Avida allowed us to evolve the control system of the agent's behavior in an openended manner. Afterwards, we were able to analyze the behavioral algorithms that evolved and assess how they conformed to the different taxes that have been proposed. Avida also permits reconstructing the evolutionary trajectory of any given lineage, which allowed us to identify potential relationships between behaviors over evolutionary time.

We performed an exploratory study taking into consideration the latest evidence and interpretation of the Ediacaran fossil record. Recent evidence points towards abundant but patchy microbial mats over the entire period in which we find burrowing traces. We therefore tested an environment with a constant abundance of nutrients over evolutionary time, and
where nutrients were distributed in homogeneous patches with no gradients<sup>19</sup>. The aim was to determine whether such an environment would lead to the evolution of the behavioral patterns observed in the fossil record. We further simplified the environment by assuming that there was no competition with neighboring organisms: each organism born was placed alone in its own testing environment.

We chose to limit the type of sensory system that could evolve and, by so doing, test whether even a very crude sensory apparatus is sufficient to evolve the behaviors observed in the fossil record. Specifically, we assumed that each Avida organism could evolve to sense at most the quality of the nutrient at its current location. This ability would allow it to distinguish four conditions: nutrient present, nutrient absent, edge of nutrient patch, and nutrient already consumed.

We also investigated the effect that patch shape, coverage, and distribution have on the types of harvesting strategies that evolved, and whether different patch structures favored the evolution of particular burrowing strategies.

#### 3.1.5 Our Findings

In our experiments, harvesting behavior became more complex over the course of evolution, tending to maximize patch coverage. We were able to replicate the stereotypical motifs from

<sup>&</sup>lt;sup>19</sup> We agree with Carbone and Narbonne, 2014 [157], that it is unlikely that food gradients were solely responsible for the complexification of trail designs occurring simultaneously all over the world. Therefore, we decided to test the condition where patches did not have gradients and were not in decline.

the fossil record and observe how organisms transition from one type of strategy to another over generations. We also found that when patches were fragmented and dispersed, organisms evolved to balance exploration and exploitation. Remarkably, this ability involved the evolution of memory utilization to decide when to leave a patch and seek another one.

Innate responses, such as phobotaxis, evolved in some of our organisms. However, what appears as strophotaxis in our system was simply an organism's response to reaching the edge of a patch, and what appears as thigmotaxis was an artifact from other behaviors that produced compact parallel trails.

Overall, in a regime where fitness depends on the amount of nutrients consumed during a lifetime, cognitive improvements in response to patch geometry and distribution were sufficient to produce the various complex strategies seen in the fossil record, without any need for complex or long-range sensory capacity, changes in nutrient availability, or intra-patch competition.

## 3.2 Methods

For these experiments, we used the Avida platform described in Chapter 2. This time, the behavioral task consisted of harvesting nutrient patches within an arena. We tested eight different environments, each containing four virtual arenas with unique patch configurations (table 3.1). The arena size was 50 x 50, larger than in Chapter 2, and could contain one or more nutrient patches. Patches could be polygonal or have an irregular ('organic') shape (e.g., fig. 3.2).

**Table 3.1:** Environments in the order they were used in the experiments. Each environment consisted of four arenas with different patch configurations. Each organism experienced only one arena in its lifetime.

Environment name	Number of patches	Patch shape	Edge cues	Average coverage
Rectangular with edges	gular with edges 1 large Rectangular		Yes	41%
Rectangular without edges	1 large	Rectangular	No	41%
Rectangular with hole and edges	1 large	Rectangular with hole	Yes	27%
Irregular with edges	1 large	Irregular	Yes	55%
Irregular with holes and edges	1 large	Irregular with holes	Yes	49%
Connected patches with edges	6 small	Rectangular patches with linking corridors	Yes	57%
Disconnected patches with edges	6 small	Separated rectangular patches	Yes	49%
Disconnected patches without edges	6 small	Separated rectangular patches	No	49%



**Figure 3.2:** Sample Arena with an irregular shaped nutrient patch and edge nutrients. Each arena contained a patch of nutrients with a unique shape. At the start of its life, each organism was placed alone in a randomly selected arena, within a nutrient patch, at a consistent location and orientation.

To eliminate co-evolution dynamics, organisms were neither allowed to interact directly nor indirectly through environmental modification. At the start of its life, each organism was placed alone in a randomly selected arena, within a nutrient patch, at a consistent location and orientation. The behavioral task required the organism to harvest as many nutrients as possible within its lifetime, while minimizing visits to empty locations (off-patch). We measured an organism's performance in the arena using a task quality score (also described in Chapter 2), which was calculated by counting the number of nutrients consumed, subtracting the number of visits to empty locations and dividing by the total number of nutrients in the arena, not allowing the total to be negative. Revisiting previously consumed nutrient locations did not

affect the task quality score, however, revisiting empty locations did.

We also used the same instruction set as in Chapter 2, excluding the move back instruction.

Therefore, the instructions the organisms could acquire through evolution to interact with the

environment were sense current, move ahead, turn right, and turn left (table 3.2).

 Table 3.2: Environmental Interaction instructions used in the experiment

Avida instruction	Description
Move ahead	Caused the organism to move one step in the direction it was facing.
Turn right	Caused the organism to turn 45 degrees to the right.
Turn left	Caused the organism to turn 45 degrees to the left.
Sense current	Provided the organism with an integer corresponding to the content of its current location (nutrient, empty location, previously consumed nutrient, and edge nutrient).

If organisms evolved the ability to sense the content of their current location, they would sense

nutrients, empty locations, and previously consumed nutrient locations as distinct integers (3, -

1, and 1, respectively). In some environments, the edges of a patch could also be sensed as the integer 0.

We evolved 200 populations for each of the eight environments (table 3.1), for a total of 1,600 replicates. We seeded each population with a single sessile and naive organism capable only of reproducing. After a period of evolution that lasted 200,000 updates (a unit of time in Avida),

we selected the three populations with the highest overall task quality<sup>20</sup> from each environment and analyzed the navigation strategies of their predominant organisms (most abundant genotype).

In addition, we selected the most successful predominant organism overall for a full lineage study, which consisted of testing every one of its ancestors on the behavioral task to reconstruct the evolutionary history leading to the final strategy.

All other configurations not mentioned above, such as maximum population size, mutation rates, and maximum organism age were the same as in Chapter 2 (Appendix A).

## 3.3 Results

## 3.3.1 Evolved Behaviors Resemble the Fossil Record

Starting from a sessile and naive ancestor, our evolutionary experiments produced organisms capable of moving and sensing, which responded to the presence of nutrients in the environment by systematically harvesting them in patches while avoiding empty locations. Over the course of evolution, more complex and efficient harvesting strategies emerged in a sequence similar to the one observed in the Ediacaran trace fossil record [157]. In addition, these different strategies reproduced the stereotypical fossil trail patterns, such as "scribbling", "meandering" and spiraling [161], [166] (table 3.3).

<sup>&</sup>lt;sup>20</sup> The overall task quality of a population equates to the Average Maximum Task Quality (AMTQ) explained in Chapter 2.

These harvesting behaviors evolved despite the organisms' limited potential for sensing environmental cues, and in an environment with constant nutrient availability over time, where nutrients were distributed in patches of homogeneous quality.

We also found that innate responses, such as phobotaxis and what resembles thigmotaxis and strophotaxis, evolved anew in some of our organisms. However, in our system, what seemed like strophotaxis, an organism's propensity for occasional 180 degree turns, was in fact a response to reaching the edge of a patch. In addition, what appeared to be phobotaxis was an artifact of other behaviors that produced compact parallel trails, and not an active sensing and hugging of an existing trail.

#### **3.3.2 Evolved Behaviors Fall into Four Stereotypical Strategies**

After analyzing the three top performing predominant organisms from each of the eight environments, we observed that, despite variations, their behaviors could be easily classified in four main strategies: "reactive meandering", "pattern cycling", "plowing", and "spiraling" (table 3.3). Pattern cycling (fig. 3.3) could be further divided into "edge-hugging" and "edgereflecting", depending on how the organism reacted to edge cues (fig. 3.4).

**Table 3.3: Behavioral strategies found across all treatments.** For each environment, we chose the three populations with highest overall task quality across the 200 replicates and analyzed the navigation strategy of their predominant organisms. Although there was a great deal of variation, all behaviors could be classified into four major strategies.

Behavioral Strategy	Description
Reactive meandering	Organism moved straight forward until it encountered an edge cue or an empty location, then it changed direction and continued moving straight. The resulting path crisscrossed the patch, often revisiting the same locations several times, while many locations were never visited (fig. 3.5).
Pattern cycling	Organism moved in a closed cycle, such as the shape of an octagon, or a square, offsetting the center of the cycle along a linear trajectory. The organism reacted to edge cues or empty locations by changing trajectory. Depending on how they reacted to edge cues, this strategy could be further divided into two types: edge-hugging and edge-reflecting. This strategy revisited locations often but tended to cover most of a patch (figs. 3.3 and 3.4).
Plowing	Organism moved in straight lines until reaching the edge of the patch, then turned around and continued in a parallel line offset by one cell. As the organism moved across a patch, it produced a pattern reminiscent of plowing a field, which tended to cover most of a patch while minimizing the number of locations revisited. It reacted to edge cues and empty locations by turning and starting another arm of the trail (fig. 3.6).
Spiraling	Organism spiraled inwards within a patch, moving straight and reacting to empty and previously consumed nutrients by making turns. When it neared the center of the spiral it would leave the patch in a straight line until it found another patch, where it would start spiraling again. This strategy covered most of the patches the organism visited, while minimizing the number of revisited locations. The decision of when to leave the current patch for another was based on memory of previous turning conditions (fig. 3.7).



**Figure 3.3:** Example of the pattern cycling strategy (with edge-hugging) in the environment Rectangular with Edge. Green and red circles indicate the start and end of the trail.



**Figure 3.4:** Illustration of the edge-reflecting (left) and edge-hugging (right) variations of the pattern cycling strategy





## **3.3.3 Evolved Strategies Depended on Patch Structure**

In our experiment, organisms evolved to respond to high-level features of the environment such as the edges of the patch or the trail left by the organism. We did not see a harvesting

strategy based on local search such as proposed by Plotnick and Koy [153].

Plowing and spiraling were the most efficient strategies in the environments where they

evolved, reaching higher task quality than competing strategies. Reactive meandering and

pattern cycling were the least efficient strategies and reached comparable levels of task quality.

However, reactive meandering proved the most robust strategy, and the only one that evolved

in all environments (table 3.4).

Table 3.4: Strategies that evolved among the three best performing populations from ea	ach
environment.	

	Strategies Evolved				
Environment Name	Reactive meandering	Pattern cycling	Plowing	Spiraling	
Rectangular with edges	Yes	Yes	Yes		
Rectangular without edges	Yes				
Rectangular with hole and edges	Yes		Yes		
Irregular with edges	Yes	Yes	Yes		
Irregular with holes and edges	Yes	Yes			
Connected patches with edges	Yes	Yes			
Disconnected patches with edges	Yes				
Disconnected patches without edges	Yes			Yes	

When analyzing the effect of patch shape, coverage, and distribution on the evolution of harvesting strategies, we found that environments with single, solid patches produced a wider variety of strategies than those with multiple patches or single patches with holes. However, patch coverage, the ratio between the area covered with nutrients and the total area of the arena, had no discernible effect on the diversity of the strategies that evolved. Finally, only in an environment with multiple, disconnected patches did we see the spiraling strategy evolve, and it involved a balance between exploration and exploitation. The presence of edge cues did not have a consistent effect. Environments with single rectangular patches produced a wider variety of strategies when edge cues were available, while environments with multiple disconnected patches produced a wider variety of strategies (including spiraling) when edge cues were not available.

#### 3.3.4 In-Depth Results from a Single Environment, Rectangular with Edges

The environment rectangular with edges was one of two that evolved the highest diversity of strategies: reactive meandering, pattern cycling, and plowing.

Among the three, reactive meandering had the lowest task quality performance. The organism that employed it reacted to edge cues but not to previously consumed nutrients (fig. 3.5), leading to wasted movements in already-explored areas of the patch.

Pattern cycling displayed intermediate performance. The organism that employed it reacted to edge and empty location cues but not to previously consumed nutrients (fig. 3.3). It also revisited positions often, however, due to its systematic and tightly spaced cycling, it was able to fully exploit those areas that it visited.

Plowing had the best performance. As with pattern cycling, the organism reacted to edge and empty location cues, but not to previously consumed nutrients (fig. 3.6). However, due to its systematic navigation in parallel lines, it was able to exploit most of the patch while not revisiting positions as frequently as the other two strategies. It was noteworthy that, before the organism started plowing, it moved in a straight line until it encountered one of the edges of the patch, then it would start the plowing pattern in the opposite direction. This initial step allowed it to harvest the entire patch in a single sweep.



Figure 3.6: Example of the plowing strategy in the environment rectangular with hole and edges

#### 3.3.5 The Spiraling Strategy Evolved Memory Usage

Spiraling evolved in a single environment, disconnected patches without edges. This was one of two environments made of multiple, small, disconnected patches. Spiraling had some interesting features. It evolved in an environment without edge cues, which we expected to make the task more difficult. However, the organism that used this strategy was sensitive to previously consumed nutrients and used these cues to guide its navigation. It was also noteworthy that this organism spiraled inwards. When placed on a patch, the organism moved straight until it found one of the edges, then it moved along the patch's perimeter, and finally spiraled towards the center. However, before reaching the center of the spiral, it would leave the current patch and move in a straight line until it found another patch, where it would start another inwards spiral.

The decision of when to leave the current patch was based on a complex calculation involving the organism's memory of the conditions under which it made previous turns, such as by reaching the edge of the patch or previously consumed nutrients. Therefore, this strategy also implemented a balance between exploration and exploitation.



**Figure 3.7:** Example of the spiraling strategy in the environment disconnected patches without edges. The arena is toroidal and the organism begins its navigation at the green dot and ends at the red. It spirals inward but leaves the patch before reaching the center. It starts a new spiral upon encountering a new patch.

#### 3.3.6 Lineage Analysis Shows Different Strategies Evolving from One Another

We selected the overall top performing organism, which produced the plowing strategy in the environment rectangular with hole and edges (fig. 3.6) and analyzed its ancestral lineage to reconstruct the evolution of its behavior. We found that starting from a sessile ancestor, the lineage first evolved pattern cycling, then transitioned into reactive meandering, and finally plowing (fig 3.8). This result shows that the three strategies could evolve into one another over the generations within the same lineage. Below we describe the strategies found at intermediary points during evolution.



**Figure 3.8:** Evolutionary history of the plowing strategy in the environment rectangular with hole and edges. On the left, is the evolution of task quality over time. On the right, the different navigation strategies from selected ancestors along the lineage.

**Organism 1:** This organism used a Pattern Cyling strategy that hugged one of the edges of the patch. It moved slowly and did not travel very far in the patch before reproducing. However, this organism could already react to edge and empty location cues, which shows that the ability to sense and discriminate environmental cues evolved early.

**Organism 2:** This organism performed reactive meandering, and moved faster than Organism 1. It always responded to edge cues, avoiding visiting empty locations. Although it was able to visit most of the patch, much of it was left unharvested.

**Organism 3**: This organism also performed reactive meandering. Interestingly, it often alternated responding to edge cues and empty location cues. This made its navigation less systematic than its ancestor's, but it harvested more of the patch.

**Organism 4**: This organism used the plowing strategy. It responded to both edge and empty location cues. It systematically exploited the areas it visited, however, it did not visit the entire patch.

**Organism 5**: This organism also used the plowing strategy, but it moved more quickly and further than Organism 4, therby harvesting most of the patch.

**Organism 6**: This was the final organism in this lineage. It used the plowing strategy, but with an important addition to that of Organism 5. Before starting the plowing pattern, it moved in a straight line until reaching the edge of the patch. Then, it performed a series of turns that allowed it to position itself inside the patch, close to the edge and parallel to it. Then it plowed the patch in one sweep.

## **3.4 Discussion and Conclusions**

#### 3.4.1 Complex Trails Require Healthy Patches

It has frequently been assumed that Ediacaran microbial mats were declining over time, from homogeneous coverage when mat mining first arose to increasingly patchy and rarefied coverage towards the Cambrian [157]. This assumption was incorporated in previous computational models of mat mining behavior [153], [157], [168].

A more recent interpretation of fossils and other data from the period instead favors the thesis that Ediacaran matgrounds were naturally patchy and were not declining before the Cambrian [154]–[156], [169]. In support of this modern view, our results show that, in a patchy but stable environment, mat mining behavior evolves in a similar progression and produces remarkably similar trail designs to those seen in the fossil record, due simply to competition for faster reproduction.

Although our work is exploratory, our results also suggest that the most complex harvesting behaviors only evolve in environments where patches are large and homogeneous, indicative of healthy and abundant microbial mats. In environments where patches are excessively irregular and fragmented, indicative of stressed and shrinking microbial mats, the more systematic harvesting strategies would not be effective and, indeed, do not evolve. Such strategies include plowing and spiraling, which require large expanses to be efficient and which, in excessively irregular and fragmented environments, have no competitive advantage over simpler behaviors, such as pattern cycling and reactive meandering. If our early results hold, the more complex and efficient trail patterns observed at the end of the Ediacaran and early Cambrian indicate healthy and abundant microbial mats, not a decline as previously assumed.

#### 3.4.2 Patch Boundaries May Have Guided Mat Mining Behavior

Previous computational models that successfully reproduced the trail patterns found in the fossil record assumed that the animals' behavior was either based on three taxes (phobotaxis, thigmotaxis, and strophotaxis) [168], [175] or on a single gradient-following taxis [153]. Our results, however, show that alternative scenarios are possible.

In our experiments, organisms evolved behaviors that were always sensitive to patch boundaries, regardless of whether they were sensitive to existing trails or whether they followed innate patterns. Even the more systematic behaviors, plowing and spiraling, were shaped by the boundaries of the environment and not by thigmotaxis or strophotaxis.

Assuming that Ediacaran microbial mats were naturally patchy, it would have been evolutionarily advantageous for mat miners to use patch boundaries and other persistent features of the environment to guide their behavior. In Chapter 2, we showed how cognition evolves to rely on persistent environmental cues to make decisions, whether these cues last for generations or only for periods within the organism's lifetime.

Moreover, as our results show, a simple sensory capacity, such as contact chemoreception at the anterior end of the animal, would likely be sufficient to detect features such as existing burrows and the boundaries of the harvestable area. Boundaries could have been demarcated by different features, such as empty ground, sessile organisms, low oxygen, and decaying organisms. In any case, this sensory data would have provided sufficient input for the observed behaviors.

#### 3.4.3 Peculiar Spiraling Behavior Balances Exploration and Exploitation

The behaviors that evolved in our experiments produced trails remarkably similar to those from Ediacaran and early Cambrian fossils. Reactive meandering shows close resemblance to "twodimensional avoidance traces" in Carbone and Narbonne, 2014 [157], while pattern cycling and plowing resemble "scribbling" and "meandering" in Seilacher 1967 and 2007, who also documents different types of spiraling behavior [161], [166].

Our organism's spiraling behavior was particularly interesting because it evolved in a multipatch environment and it spiraled inwards. Inwards spiraling is thought to be more efficient than outwards spiraling, as it allows the organism to fully exploit an area by first establishing its limits. It also allows the organism to estimate how much nutrient is left in the patch based on how tight the coiling is becoming [164].

In fact, our spiraling organism did not fully exploit its patches. Before it reached the center of a spiral, it would leave in search of another patch. The organism balanced exploration and exploitation in a manner reminiscent of Charnov's Marginal Value Theorem (MVT) [179], where the optimal time an organism remains on a food patch before leaving for another one is a function of the time already spent and the time required to get to another patch.

Moreover, the organism relied on memory to decide when to leave a patch. Behaviors based on memory use are considered more advanced than hardwired ones [106]. What is striking is that

this more complex cognitive function evolved in the context of balancing exploration and exploitation instead of maximizing grazing coverage.

The efficiency of these early animals' grazing patterns has been the focus of previous discussions about their cognitive capacity and whether their nervous systems became more complex over time [157]. Although there is no straightforward correlation between the complexity of the grazing behavior and the complexity of the nervous system [152], there is a qualitative difference between a cognitive system that performs hardwired behaviors and one that uses memory and performs complex calculations to decide when to switch patches.

Assuming that Ediacaran microbial mats were naturally distributed in patches, it may not have been so much the need to optimize grazing patterns that drove the evolution of cognition. Instead, it may have been another challenge that mat miners faced: deciding when to leave a patch and how to find the next one.

## Chapter 4: Beyond Associative Learning, the Early

# **Evolution of Configural Learning**

Authors: Anselmo C. Pontes, Andrew Mitchell, Charles Ofria, and Fred C. Dyer

## **4.1 Introduction**

#### 4.1.1 Background and Motivation

Learning is classified in order of presumed complexity, which may also be the order in which it evolves[180], [181]. The simplest learning forms, habituation and sensitization, are known as nonassociative learning [180]. Unambiguous associations between pairs of elements of a set of stimuli and responses, such as in Pavlovian and instrumental conditioning, are known as elemental learning [181], [182]. Associations involving configurations of stimuli that interact in space and time, creating ambiguity that can only be solved in context, such as in discrimination and rule learning, are known as configural learning (or non-elemental learning) [180]–[183].

In Chapter 2, we confirmed the hypothesis that cognitive abilities, such as learning, are modular and evolve in a stereotypical sequence from simple to more complex by building upon previous competences [57], [94], [96], [184]–[186]. We decided to extend this experiment beyond instrumental conditioning and investigate whether we could evolve more complex forms of learning as well as other cognitive abilities.

#### 4.1.2 Our Experiments

We began by adding a new type of cue to our Avida environment from Chapter 2. This modification allowed us to create cue combinations with different meanings. We expected that this more complex environment would foster the evolution of a form of configural learning in which the organism was required not only to learn new cues but to respond differently to them depending on previous experience with a separate cue. We also expected that if configural learning evolved, it would recapitulate many of the same steps in the evolution of instrumental conditioning from Chapter 2.

In addition, we were interested in determining how the complexity of the behavioral control algorithms that evolved in this study would compare with those from Chapter 2. Although learning abilities do not always correlate with the complexity of an animal's nervous system [180], [181], there is a trend of increasing ability when going from nerve nets to centralized nervous systems, and according to brain size among closely related species. For example, jellyfish (which have on the order of 10<sup>4</sup> neurons, organized in nerve nets) are capable of habituation and sensitization but not associative learning [180], [187]–[189], and fruit flies (which have on the order of 10<sup>5</sup> neurons in their brain) are capable of learning associations between cues presented simultaneously but not associations that depend on previous experiences such as abstract concepts [180], [190]. Honeybees (which have on the order of 10<sup>6</sup> neurons in their brain), however, are capable of all these, in addition to other forms of learning previously thought exclusive of vertebrates with much larger brains, such as learning by observation [180], [181], [191]–[193].

#### 4.1.3 Our Findings

We succeeded in evolving organisms capable of solving the configural learning task; however, they were exceedingly rare. Out of 2,400 replicate populations, only two evolved organisms with this capability.

We singled out one of these organisms and analyzed its evolutionary history. As in the previous study, evolution followed a discrete sequence of behavioral stages, starting with moving and sensing, followed by reflexive trail navigation using error recovery, then associative learning, and finally, configural learning.

When comparing the organisms that evolved configural learning in this study with those that evolved instrumental conditioning in Chapter 2, we found that although the lengths of their behavioral control programs were similar, the complexity of the algorithm responsible for the configural learning behavior was considerably higher. The rare evolution of these more complex behavioral control algorithms may indicate a technical barrier that limits the cognitive capacity of the organisms that can be evolved with the current version of Avida.

## 4.2 Methods

We continued using the Avida platform and the behavioral task consisting of navigating a trail of nutrients in a virtual arena described in Chapter 2, where the nutrients could cue the organism on the direction to follow if they evolved the ability to sense and interpret the cues correctly. For this experiment, we used a new type of cue, called "sharp turn", in addition to the four used in Chapter 2 (forward, left turn, right turn and empty location). A sharp turn cue

immediately preceding a left or right turn cue indicated that the correct action was to turn 90 degrees instead of 45 degrees.

The sharp turn, empty location, and forward cues, were represented by integer values that remained constant throughout evolution (1, -1, and 0, respectively). While the right and left turn cues were also represented by integers, they were randomly drawn from the interval between 2 and 101 (inclusive) every time an organism was placed in the arena and remained constant only during the organism's lifetime.

For an organism to perform optimally in this environment, it had to associate at least one of the two turn cues with the correct meaning, identifying the remaining cue by exclusion. Additionally, the organism had to discriminate when to turn 90 or 45 degrees based on the presence or absence of the sharp turn cue immediately before a turn cue.

We performed two experiments, each with a different environment consisting of four arenas with unique trail configurations. Both environments were based on the nutrient cued environment from Chapter 2. In the first experiment, the direction of each of the first two turns, although random, could be predicted by counting the number of forward cues that preceded the first turn. In the second experiment, the direction of the first two turns was also random but only the direction of the second turn could be predicted by counting the number of forward cues preceding the first turn.

In addition, the trails in both environments started with 45-degree turns and the first appearance of a 90-degree turn was at approximately the 25% mark of each trail (fig. 4.1). This

exclusion of 90-degree turns from the first portion of the trail was meant to simplify the early stages of evolution and reduce the number of cues an organism had to respond to before it could navigate far into the trail. The 25% mark was chosen based on the results of preliminary experiments, and it afforded a balance between facilitating the evolution of simpler behavioral building blocks with maintaining a strong selective pressure for the new cues to be identified and acted upon.



**Figure 4.1:** Environment with four unique arenas used in the second experiment. At the beginning of each organism's life cycle, we placed it on a nutrient at the start of the trail in a randomly selected arena, facing the next nutrient. The direction of the first two turns in any trail was random. However, the direction of the second turn could be predicted from the number of nutrients preceding the first turn. The first 90-degree turn only appeared at approximately the 25% mark of each trail.

In both experiments, we used the same Avida instruction set as in Chapter 2 which included the

move back instruction (Appendix B). All other configurations were the same as in Chapter 2

(Appendix B), such as maximum population size (3600), arena size (25 x 25), and total duration

of each experiment (250,000 updates). We evolved 600 population replicates in the first

experiment and 1,800 in the second. Each population was seeded with the same naive organism of Chapter 2, which lacked any instruction for behavioral control other than the one necessary for reproduction.

We also performed a full lineage study of the final predominant organism (most abundant genotype) of the population with the highest AMTQ (Chapter 2) from the second experiment. This study consisted of analyzing the performance of each of the organism's ancestors and characterizing the main behavioral transitions along the lineage.

Finally, we measured the length and complexity of the algorithm of the predominant organism from Experiment 2. We used the Cyclomatic complexity measure [194] to determine the complexity of each algorithm. This measure is based on the number of branching points and alternative paths of execution in an algorithm. It requires tracing the execution of the organism's program, flowcharting its algorithm, converting it into a graph, and evaluating the number of edges, nodes and connected components.

## 4.3 Results

In our first experiment, we evolved 600 population replicates and found only one where the predominant organism was able to perform the task. This organism learned by imprinting; that is, it used the number of nutrients at the start of the trail to disambiguate the direction of the first turn and associate one of the turn cues with the correct direction. The organism was also able to perform the configural learning task and turn 45 or 90 degrees depending on the presence of the sharp turn cue. However, its behavior did not generalize to all trail configurations. When placed in a trail where the initial segment did not conform with the

pattern it had evolved in, the organism was not able to make the association, instead leaving the trail and ceasing movement.

For the second experiment, we hypothesized that if we made the start of the trail harder to predict, we would foster the evolution of a generalizable learning behavior, similar to the generalizable imprinting described in Chapter 2, in which the organisms evolved to learn the association by making a mistake, going off the trail, and recovering. Therefore, we used an environment where the first turn could not be predicted from the number of nutrients preceding it, although the second turn could.

We evolved 1800 population replicates until we found one where the final predominant organism was able to perform the configural learning task in this more challenging environment. This organism learned by making a mistake at a turn cue, stepping off the trail, and using error recovery to return to the trail and turn to the correct direction (Fig. 2). In subsequent tests, this behavior proved to generalize across all alternate trail configurations and, therefore represents evidence in support of our hypothesis. Remarkably, when we tested this organism on trails where the turn cues were changed or swapped midway, the organism was able to relearn and reversal learn, even though its ancestors had not experienced these conditions during evolution. When tested in an environment where the first turn was a 90degree one, the organism was still able to learn the association even though the environment in which it evolved always started with 45-degree turns.





**Figure 4.2:** Path of an organism on a nutrient trail demonstrating configural learning. In this trail, turn cues have different meanings depending on the context (i.e., if they are preceded by a sharp turn cue or not). The organism makes a mistake on the second turn, stepping off the trail, but recovers and subsequently associates the meaning of the cue with the correct direction. Afterwards, it is able to extrapolate the learned cue to different contexts.

We flowcharted the organism's algorithm (Appendix B) and calculated its algorithmic

complexity (cyclomatic complexity), which was more than twice as much as the cyclomatic

complexity we found among the organisms that evolved general relearning in Chapter 2 (13 vs.

6).

We also measured the organism's length, which was 84 instructions, of which only 78 were

executed. This is in the same range as the organisms that evolved general relearning in Chapter

2, which was between 55 and 136 instructions long, of which between 53 and 110 instructions

were executed.

Finally, we performed a lineage study on the organism's ancestral line. We found that the ancestral behaviors evolved in succession, following the order that we predicted: First moving and sensing evolved followed by error recovery behavior that allowed the organism to navigate most of the trail with frequent mis-steps (fig. 4.3). This reflexive behavior was followed by the evolution of associative learning, where the organism was able to associate the turn cues with the correct directions and navigate 45-degree turns, but was not able to extrapolate the learned association to the 90-degree turns, and would navigate these by error recovery (fig. 4.3). Only after this last behavior was in place, configural learning evolved (fig. 4.2).



**Figure 4.3:** At left, path of an ancestral organism that navigated by error recovery. At right, Path of an ancestral organism that was capable of associative learning but not configural learning. The organism learns the cue association in the second turn and uses it to navigate all 45-degree turns. However, it is not capable of extrapolating the association for the 90-degree turns and uses error recovery instead.

### 4.4 Discussion

#### 4.4.1 Extra Cognitive Abilities May Contribute to Adaptation

Confirming our expectation, configural learning evolved as the last step in a sequence of increasingly complex behaviors, recapitulating the evolution of instrumental conditioning from Chapter 2. We were surprised, however, by the evolution of certain additional cognitive abilities that were not directly selected for in our environments, namely relearning, reversal learning, and learning cue associations at 90-degree turns. In Chapter 2, relearning and reversal learning evolved only in environments that contained cue reversals along the trails.

According to the theory of sensitive periods and other constraints on learning [84], [89], [147]– [149], [195], [196], learning is costly, and it is optimized by evolution based on the needs of an organism in its environment. One form of optimization is to restrict the periods in an animal's life when it is most capable of learning (sensitive periods). For example, in Chapter 2, some ancestral organisms that were capable of relearning and which evolved in an environment without cue reversals eventually gave rise to organisms that could only imprint once (Appendix B, Section B.6).

It is possible that, given enough time, evolution would have continued to shape our organism's learning ability, optimizing it for this particular environment. Nevertheless, the possibility remains that as more complex cognitive abilities evolve, they may carry extra features that, at least initially, increase the generality of the organism's intelligence, which can be beneficial in adapting to different environments.

#### 4.4.2 We Are Approaching a Complexity Barrier

Complex cognitive abilities evolve by building upon simpler ones. Therefore, the more complex the ability, the longer it should take to evolve anew and the rarer it should be among replicate populations. In our system, however, it seems that rarity increases exponentially with both behavioral and algorithmic complexity.

In Chapter 2, where the algorithmic complexity of both error recovery and relearning was 6, error recovery evolved in 35% of the replicate populations, imprinting evolved in 9%, while relearning evolved in only 2%. In the configural learning experiments, where the algorithmic complexity of the generalizable version of configural learning was 13, the imprinting version evolved in 0.17% of the replicate populations while the generalizable version evolved in 0.06% and required up to 1800 replicates for a single population to succeed at the task.

From a technical point of view, it becomes costly to perform digital evolution experiments using such a large number of replicates. It is possible that if we had designed our environment differently, or if we had seeded the populations with organisms already capable of associative learning, configural learning would have evolved more readily. However, seeding an experiment with pre-evolved organisms likely restricts the open-endedness of behaviors that will evolve and may bias the order in which novel traits appear. In addition, if rarity continues to increase exponentially with behavioral complexity, even seeding populations with pre-evolved organisms will no longer be sufficient to make an experiment practical.

It seems that we reached a complexity barrier with Avida, a problem that other researchers have encountered with different systems [197], [198]. Although Avida has proven to be a versatile experimental tool in digital evolution, making invaluable contributions to our understanding of Darwinian evolution and training a generation of digital evolution scientists, Avida's current genetic representation is inadequate for evolving organisms beyond a certain level of complexity. This limitation mainly stems from Avida's difficulty in evolving multiple coexisting loops, reusable modules such as subroutines and functions, and memory usage beyond a few registers. Although some of these issues may be mitigated by modifying the instruction set and the virtual hardware, this would likely only raise the complexity barrier and not solve the problem. What is needed is a new genetic representation that facilitates the evolution of complexity.

# **Chapter 5: Evolution of Allosteric Regulation in**

## Cyanobacteria

## **5.1 Introduction**

In the preceding chapters, we investigated some of the earliest and most important steps in the evolution of natural cognition: navigation, nutrient harvesting, and associative learning. As we attempted to increase the complexity of the cognitive tasks, the evolution of the organisms capable of performing those tasks took exponentially longer. Such complexity barriers frequently limit evolution potential in artificial systems and are a recurrent problem with present digital evolutionary methods [197], [198].

However, this does not appear to be a problem for biology, which has evolved cognitive systems of incredible complexity and intelligence. Cognitive systems in living organisms also have unique properties, which may be the reason for their evolvability and scalability. For example, regulatory networks in cells have extraordinarily flexible topologies with substantial modularity and redundancy [59], [199], [200]. In addition, cells employ multiple mechanisms of regulation [201], and seem to exploit noise rather than attempting to minimize it [58], [202], [203]. It appears that, in the process of creating life, nature created a language to implement evolvable control systems of arbitrary complexity [204]–[206].

Biology also has the advantage of scale over artificial systems, which we strive to mitigate by deploying our understanding of evolution, both natural and digital, to optimize the process and reduce the necessary time scales and population sizes. Such techniques include using artificially high mutation rates, multi-island environments, and carefully crafted selective pressures. However, biology's advantage in scale does not diminish the importance of natural genetic representations. In fact, the vast time scales of natural evolution provide the ultimate long-term test of evolvability by implicitly selecting for organisms with underlying genetic representations that are capable of producing beneficial complex traits. Therefore, in order for artificial evolutionary systems to achieve similar levels of adaptive complexity, we must also take inspiration from the aspects of natural systems that allowed them to express complex phenotypes so successfully.

#### 5.1.1 A New Digital Evolution Platform

Inspired by biology, I carefully translated the cell's chemical regulatory language into a digital format, and implemented it in a new evolution platform named Elfa (Evolutionary Lab for Flexible Agents). I attempted to preserve the essential features of cell regulation that make it a powerful evolvable control systems language, including much of the stochasticity, while taking advantage of the digital format to optimize other mechanisms, thus avoiding the steep costs of chemical simulations.

In Elfa, the evolving agents are equivalent to cells, and the genetic representation is based on a binary string chromosome, with genes capped by basal promoters and enhancers that bind transcription factors, similar to Banzhaf [207], [208]. However, Elfa's representation

incorporates additional biological mechanisms that allow more complex networks to evolve, such as ligand binding, agonism, proteases, and assembly-program gene products.

In order to demonstrate the platform and test its capabilities, I performed an experiment where a digital equivalent of cyanobacteria can evolve allosteric regulation anew and take advantage of the daily light cycle.

#### 5.1.2 Evolution of Allosteric Regulation Experiment

As far as we know, oxygenic photosynthesis evolved only once in the ancestors of modern cyanobacteria [209]. All green algae and plants inherited this capacity from an ancestor that endosymbiotically acquired a cyanobacterium. In this experiment, I investigated the evolution of gene regulation on the equivalent of an early prokaryote oxygenic photosynthetic cell.

This experiment consisted of exposing populations of cells that expressed their genes at a low, constant rate and had no regulatory ability, to an environment where the light irradiance followed a daily cycle. In addition, each cell contained chemical signals whose concentration correlated with the light intensity in the environment, as well as the cell's level of stress, and energy reserves. My hypothesis was that over the generations these cells would evolve higher photosynthetic output and reproduce faster, maximizing the use of the available light.

Cells could evolve higher fitness via mutations that adjusted the rate at which their genes were expressed, without adding any dynamic regulation mechanism, or they could evolve regulatory mechanisms that allowed them to time the production of proteins, as well as growth and cell division, according to the environmental light cycle. Moreover, the experimental setup allowed different regulatory mechanisms to evolve to solve the same problem. Some would involve direct regulation of one or more of the original proteins via allostery, while others would involve the evolution of new ligand-regulated transcription factors.

#### 5.1.3 Findings

All populations evolved higher fitness over time by adjusting the expression level of their genes. Some populations also evolved allosteric regulation of the original proteins, but this was rare. Regulated cells took advantage of the light cycle by concentrating their growth and reproduction during the light period and reducing protein expression at night. This behavior technique allowed them to harvest much more energy from the environment, and reach much higher fitness than the fittest cells that did not use regulation.

### 5.2 Methods

#### 5.2.1 Cell and Population Organization in Elfa

In Elfa, virtual cells have a prototypical shape, with a given volume of cytosol bounded by a semi-permeable membrane. The cell's volume and surface area change as the cell grows and divides. The cell's cytosol contains a single circular chromosome, proteins, ligands, and other molecules, which all can interact. The cell's chromosome is made of binary digits, organized into 32-bit segments that correspond to integer numbers. The combination of 32-bit segments, within a certain range of values, can encode for operons, each containing one or more genes. Operons can have one or more basal promoters, as well as one or more upstream enhancers.
Segments with integer values outside of the coding range are interpreted as non-coding sequences (if they are not transcribed) or linker domains (if they are transcribed).

During replication, the cell's chromosome can suffer mutations. These could be simple point substitutions (bit flips), as well as multiple segment duplications, transpositions, and deletions. Mutations in an operon's basal promoter, for example, could alter the promoter's affinity for the RNA polymerase and change the expression level of that operon's genes. Mutations could also change the properties of proteins encoded in genes, for example making them sensitive to a ligand. In addition, mutations could cause genes to duplicate or become pseudogenes. In sum, much like in biology, mutations can shape the architecture of chromosomes in Elfa, which encodes not only the genes but also the regulatory regions that determine how those genes are expressed.

Cells are grouped in a population that is well mixed and capped in size. When a cell reaches a threshold amount of cell growth factor and possesses sufficient energy reserves (details below) it replicates, causing it to divide into two daughter cells. Once a population reaches the maximum size, the birth of any additional cell causes the removal of another cell at random. This constraint creates a strong selective pressure for fast replication.

### 5.2.2 Ancestral Cells

The ancestral genotype that seeded all experimental populations possessed only three genes, transcribed at different, but constant, rates that were proportional to their basal promoter

affinity. These genes encoded a photosynthetic complex, an RNA polymerase, and a growth factor, described below.

The photosynthetic complex was responsible for converting CO<sub>2</sub>, H<sub>2</sub>O, and light into sugar that the cell used for energy. The sugar conversion rate in a cell was proportional to the abundance of the photosynthetic complex and the light irradiance at a given time and limited by the amount of CO<sub>2</sub> that diffused into the cell. If the cell ran out of CO<sub>2</sub> in the presence of light, this complex would produce a harmful compound, Reactive Oxygen Species (ROS), instead of sugar. At high concentration, ROS would slow protein translation and cause damage to the cell's components including somatic mutations.

The RNA polymerase (RNAp) was responsible for transcribing the cell's genes. The rate at which a gene was transcribed varied with the abundance of RNAp.

The cell growth factor was responsible for causing the cell to grow, and also for signaling the initiation of cell replication. When the abundance of this factor reached a certain threshold, the cell would attempt to replicate its genome and divide into two daughter cells.

In addition to the three gene products described above, this representation allowed the encoding of proteins with the function of transcription factors. Transcription factors bound to enhancers and either promoted or repressed protein transcription of the operon downstream. Proteins could evolve a ligand-sensitive domain that temporarily bound ligands according to their affinity sequence and affected the function of the protein, either increasing or decreasing its activity.

#### 5.2.3 Ligands in Elfa

Elfa provides each cell with signals in the form of ligands, which convey information about the conditions in the environment and within the cell. The strength of a signal is proportional to the corresponding ligand concentration. However, the ancestral genotype was insensitive to these signals. In order to sense them, the cells had to first evolve proteins with ligand-sensitive domains with affinity for the specific ligands.

These were the five types of ligands used in this experiment:

The reactive oxygen species (ROS) ligand signaled the amount of this harmful compound. The ligand was produced as the cell ran out of CO<sub>2</sub> during photosynthesis. Higher concentrations of ROS increased the probability of somatic mutations, protein damage, translation slowdown, cell aging and loss of fat reserves.

The cyclic AMP (cAMP) ligand signaled starvation stress. This ligand was produced as the cell ran out of sugar reserves. At low concentrations, it indicated that fat reserves were being converted into sugar. However, in high concentrations, it indicated that the fat reserves were running out. Higher concentrations of cAMP increased the chance of somatic mutations, protein damage, and translation slowdown.

The irradiance level ligand indicated the amount of light reaching the cell. Its concentration followed a 24-hour cycle of night and day, where underwater daylight irradiance is modeled according to the literature [210].

The sugar reserves ligand indicated the amount of sugar immediately available as energy for the cell. It was produced as the cell performed photosynthesis and accumulated sugar.

The fat reserves ligand indicated the amount of energy in long-term storage. It was produced when the sugar reserves reached a maximum threshold, and any excess sugar was converted into fat.

All proteins had a half-life, which determined the average time that they remained in the cell before being recycled. Cells, on the other hand, had no age limit but accumulated harmful protein aggregates (plaque) over time [211].

Plaque was formed when proteins were damaged by excessive ROS, could not be recycled, and accumulated in aggregates, which in turn could cause further damage to healthy proteins (cell 'aging'). When a cell replicated, all of its plaque was segregated into one of the daughter cells, while the other was born 'rejuvenated' [212]. However, at cell division, all proteins and energy reserves of the mother cell were divided between the daughter cells.

### 5.2.4 Cellular Costs and Constraints

There were a number of costs and constraints to the cell's actions. Cell replication, once initiated, could not be stopped, and lasted an amount of time proportional to the length of its chromosome. During this period, the energetic cost of replication was subtracted from the cell's energy reserves (sugar and fat). A cell that initiated replication and ran out of energy reserves, died. Transcribing genes, and translating proteins, also entailed energy costs as well as time delays, proportional to the gene's length, and to the protein copy number. This dynamic created a selective pressure for efficient genetic encoding and regulation. By imposing the cell with costs and constraints while providing sensory feedback signals, we avoided adding a number of artificial limitations (such as cell age or genome size) or rates (e.g., somatic mutation rate or transcription rate). In addition, we provided more of an opportunity for the evolution of complex regulatory networks.

#### 5.2.5 Experimental parameters

For this experiment, I evolved 1,012 population replicates, each capped at 200 cells, for a period of 730 virtual days. Cells experienced the light environment corresponding to 4.0 meters of ocean water, at 0.0 degrees latitude. All simulations started with light levels equivalent to those that they would experience at 6:00 am on January 1<sup>st</sup>. Mutation rates were set at 1/50 for point mutations, 1/1,200 segment duplication, 1/1,200 for segment deletion, and 1/1,500 for segment transposition.

### 5.2.6 Analyses

At the end of the experiment, I sorted the populations according to whether they had fixed allosteric regulation, transcription factors, or neither. Elfa provides population reports that make this classification straightforward. In addition, I collected samples from each of the final populations that evolved allosteric regulation and mapped their regulatory network. This mapping, however, is not automated. It requires analyzing individual cells' genes to identify

molecular affinities and analyzing the cells' internal contents and dynamics under different conditions.

In order to measure the fitness of a given population, I collected sample cells and incubated them for a period of 30 simulated days without mutations to test their viability. Due to the probability of somatic mutations under stress conditions, populations typically contained a number of damaged and unviable cells. This issue was exacerbated in the populations that did not evolve regulation since these cells were insensitive to the stress signals. Therefore, measuring the fitness of a given population was a computationally intensive process.

Once I had identified the fitness of a set of populations, and I analyzed the regulatory network of their representative genotypes, I selected a few cells that exemplified the different traits that evolved and plotted their daily cycle of growth and division, noting any salient dynamics. These plots were based on reports that Elfa can produce for individual cells or lineages.

## 5.3 Results

Out of the 1,012 replicate populations we evolved<sup>21</sup>, only 18 fixed allosteric regulation. However, 14 of these regulated populations reached much higher fitness levels than the fittest non-regulated population. No population fixed allosteric regulation of transcription factors.

<sup>&</sup>lt;sup>21</sup> The unusual number of replicates was due to the number of CPU cores, and threads that could efficiently run in the computer I used for this experiment.

However, two populations fixed transcription factors without regulation as a mechanism to increase their gene expression levels.

The cells that evolved regulation could be divided into three groups according to which of their proteins evolved allostery (fig. 5.1). The first group evolved growth factor allostery and consisted of four populations. The second group evolved RNA polymerase (RNAp) allostery and consisted of ten populations. The final group evolved both growth factor and RNAp allostery and consisted of four populations. Note that allostery of the photosynthetic complex was not a viable mechanism of regulation, due to its large protein count that would easily quench most types of ligands.

On average, the fittest group was the one where both the growth factor and the RNAp were regulated (Average = 5.5 generations per day), the second fittest was the one with regulated RNAp (Average = 4.5 generations per day). The regulated growth factor group had an average fitness of 3.1 generations per day.



**Figure 5.1:** The 18 populations that evolved allosteric regulation could be categorized into three groups according to which of their proteins evolved allosteric regulation. The first group (four populations) evolved growth factor regulation. The second group (10 populations) evolved RNA polymerase (RNAp) regulation. The last group (four populations) evolved both growth factor and RNAp regulation. All final populations had higher fitness than the ancestral one (green dotted line at 0.9 generations per day). In addition, most populations with regulation (14/18) had higher fitness than the best non-regulated final population (orange dashed line at 3.4 generations per day).

The ligand that was most commonly used for regulation was ROS. ROS is produced when cells

run out of CO<sub>2</sub> during photosynthesis, due to the slow diffusion of CO<sub>2</sub> from the environment

into the cell. In large quantities, ROS can damage proteins and energy reserves, slow down

translation, and cause mutations. Therefore, cells that use ROS for regulation are quick to

respond to even small amounts of this ligand by increasing their growth rate in order to replicate faster. Cell division serves to alleviate the shortage of CO<sub>2</sub>, since the newborn daughter cells have a higher surface area to volume ratio than the mother cell, thus benefitting from increased CO<sub>2</sub> diffusion from the environment.

In figures 5.2 to 5.5, I contrast the ancestor cell that seeded the experiment (fig. 5.2) with cells that evolved allostery from each of the three groups described earlier. Note that when RNAp is regulated, the expression of all genes is affected. In addition, it is worth noting that not all ligand affinities contribute to the fitness of the cell. Some are simply instances of cross-talk, or interference.



**Figure 5.2:** Gene regulatory network of the ancestral cell with no regulation. On the top are the five ligands available to the cell: ROS, fat reserves (Res), sugar reserves (Sug), irradiance (Light), and cAMP. The ancestral cell does not sense the concentrations of any of these ligands.



**Figure 5.3:** Gene regulatory network of one of the cells with regulated growth factors. Here, the growth factor could bind three different ligands with different affinities, and with different effects. The fat reserves ligand (Res) would bind weakly and had an agonistic effect. The irradiance ligand (Light) would also bind weakly but would have a reverse agonistic effect. Finally, the cAMP ligand would bind strongly and also have a strong reverse agonistic effect. In practice, high fat reserves would promote a slight increase in growth factor production, while both the presence of light (useful for building more reserves) and cAMP (indicating starvation) reduced the growth factor production and thus delayed replication.



**Figure 5.4:** Gene regulatory network of one of the cells with regulated RNAp. Here, the RNAp could bind two different ligands with different affinities, and with different effects. The ROS ligand would bind strongly and also cause a strong agonistic effect, while the sugar reserves ligand (Sug) would bind weakly and have a weak reverse agonistic effect. In practice, the presence of ROS, even in small amounts, would cause the cell to increase the expression rate of all its genes, including the growth factor, causing it to replicate faster. Replication increased the ratio of surface area to volume in the daughter cells, temporarily alleviating the source of stress. High sugar reserves (indicating active photosynthesis), on the other hand, would slightly reduce the expression rate of all genes, and contribute to slow down replication and build energy reserves.



**Figure 5.5:** Gene regulatory network of one of the cells with both growth factor and RNAp regulation. Here, the RNAp could bind three different ligands while the growth factor could bind two. The ROS ligand would bind strongly to both RNAp and growth factor causing strong agonistic effects. The fat reserves ligand (Res) could also bind to both RNAp and growth factor, but it caused a weak reverse agonistic effect. Finally, the sugar reserves ligand (Sug) would bind weakly to the RNAp and have a weak reverse agonistic effect. Similar to the case in fig. 4, ROS would have a strong effect on gene expression by promoting the activity of RNAp, but with even more emphasis on cell growth and replication, since it also directly promoted the activity of the growth factor. However, high fat and sugar reserves (indicating active photosynthesis) would slightly reduce the rate of protein expression and the activity of the growth factor, which favored the accumulation of energy reserves.

In figures 5.6 and 5.7, I contrast the daily cycle of two cell lineages, one capable of regulation

and another that expresses its genes at a constant rate. Regulated cells are typically more

resilient than non-regulated ones since they respond to stress signals, cAMP or ROS, thus

minimizing harm. Non-regulated cells on the other hand, are often mismatched with the

environment. Due to competition for faster replication in the fixed size population, non-

regulated cells are often swinging from starvation stress at night to ROS stress during the day

and suffering from high levels of mutation.



**Figure 5.6:** A twenty-four hour period of one of the cell lineages with RNAp allosteric regulation. In the presence of ROS, which causes an agonistic effect on their RNAp, these cells grow faster and divide, thus increasing their surface area and the absorption of CO<sub>2</sub>. Sharp drops in cell volume indicate cell division. Note that cells have steeper growth curves and also replicate more often during the day than at night.



**Figure 5.7:** A twenty-four hour period of one of the fittest cell lineages without regulation. These cells grow and replicate at a constant rate despite the environmental light cycle and their own internal stress. As a result, they often suffer damage and mutations due to ROS accumulation when they reach their photosynthetic limit.

## **5.4 Discussion and Conclusions**

Overall, the experiment's results supported my hypothesis that cells would evolve not only to maximize their photosynthetic output but also to take advantage of signals that conveyed information about the state of the environment. In addition, the results revealed some unexpected dynamics, such as the use of ROS triggered growth and cell division to relieve ROS stress instead of simply reducing the rate of photosynthesis by limiting the expression of the photosynthetic complex.

Although the cells in the few populations that fixed allostery gained a great fitness advantage, most populations repeatedly evolved allostery without it ever fixing. In fact, most populations appeared to select against allostery. One possible explanation is that the evolution of allostery is disruptive with a net detrimental effect, more often than it is purely adaptive or neutral. If this is the case, allostery will only be able to fix in the rare chance that it does not create side effects or if compensatory mutations occur rapidly to overcome the harm.

More surprising is the lack of fixation of regulation based on transcription factors. It is possible that this form of regulation was slower and less efficient than the allosteric regulation of the target proteins themselves. However, even if this is the case, transcription-factor-based regulation should evolve if there are a sufficiently high number of parameters that the cells could control, since there are only so many signals to which a protein could be sensitive.

#### 5.4.1 Elfa as an open-ended digital evolution platform

As mentioned earlier, a cell's regulatory mechanisms, and its underlying genetic representation, amount to a language of evolvable, open-ended control systems. In developing Elfa, I reinterpreted this language in a digital format. My initial results have demonstrated a variety of biologically realistic evolutionary outcomes, especially given the small populations sizes of only 200 cells and the relatively short time frames. While any fixation of de novo regulation is impressive under these conditions, my work going forward is not only to experiment with the platform but also to reevaluate my model in order to optimize its performance and make sure that it captures the essential features of biology.

As it stands, Elfa's model already demonstrates open-ended properties akin to biology. For example, sensors can not only evolve anew, but their gain and specificity can also be modified by evolution. In addition, evolution can adjust many cell parameters, such as size, replication rate, aging rate, and somatic mutation rate. However, the model could be made even more open-ended by having cells with additional evolvable traits such as evolvable ligands, evolvable RNA polymerase's affinity sequence, evolvable chromosome repair machinery, and evolvable horizontal gene transfer mechanisms

Elfa also has properties that make it a good candidate for evolving control systems for realworld applications, such as agents that acquire their own control laws via evolution, are robust to noise, function in simulated real-time, respond to analog signals, and are capable of dynamic analog responses.

In nature, cells spent three billion years evolving increasingly effective cognitive systems, like that of the amazing choanoflagellate [213]–[215]. It then took only 600 million years for cells to evolve the human brain. If our time scale is number of generations instead of years, the discrepancy in length between these two periods becomes even starker. Thus, I believe that the path to evolving general intelligence starts with evolving the cognition of a cell.

APPENDICES

# APPENDIX A

Supplementary Material for Chapter 2

# A.1 Supplementary Methods

Authors: Anselmo C. Pontes, Robert B. Mobley, Charles Ofria, Christoph Adami, and Fred C. Dyer

This text is adapted from the supplementary material of Am. Nat. 2020. Vol. 195, pp. E1–E19 [75]. © 2019 by The University of Chicago. CC BY-NC 4.0. DOI: 10.1086/706252

We used the Avida default setting for population size (3600 organisms), mutation rates (5% probability for insertion and deletion, and 0.75% probability of instruction change) and behavioral grid size (25 x 25, toroidal grid) [114]. We set the number of instructions that an organism had to execute before it could reproduce to 1500, and the maximum age to 50 times the genome length. In addition, we used an instruction set based on the optimum "genetic hardware" [216], also known as "hardware 3".

Our ancestral organism used the "reproduce" instruction, which performs the full selfreplication process, instead of the Avida default "copy loop", which implements self-replication as sequence of several instructions executed in a loop. This simplification allowed us to focus our attention on the evolution of behavior, since the evolution of self-replication in Avida has already been studied extensively [217]. Therefore, the ancestral organism we used to seed the population was made of 100 instructions: 99 "nop" instructions followed by one reproduce instruction. The Avida instructions mentioned throughout the text are also known by their Avida mnemonic, listed in table A.1. We ran each replicate for 250 thousand updates, where update is an Avida unit of time. The five behavioral strategies we used to categorize the organisms' phenotypes (table 2.2) were discerned during exploratory experiments in which we surveyed a great number of behaviors in a wide range of conditions. We defined the different strategies by observing commonalities in the paths that each organism followed when exposed to different environments and analyzing the organism's algorithm. Subsequently, when we performed the experiments described in this paper, we used these predefined categories and classified each organism by observing its path in different environments and also tracing the execution of its behavioral control program, which included monitoring the organism's internal memory.

Figures A.1 to A.4 show the four environments used in experiment 1, and Figure A.5 shows the environment used in experiment 2.

Instruction name	Avida mnemonic		
Reproduce	repro		
Sense Current	sg-sense		
Rotate Right	sg-rotate-r		
Rotate Left	sg-rotate-l		
Move Ahead	sg-move		
Move Back	sg-move-b		

Table A.1: Avida instructions mnemonic references

The organisms we selected for lineage study in experiment 2 (fig. 2.5) were the 10 most sophisticated final predominant organisms capable of relearning - those capable of reversal

learning any two symbol combinations while retaining the learned associations long-term and generalizing their behavior to any trail configuration.

The custom version of Avida used in this study can be found here:

https://github.com/mercere99/Avida-AssociativeMemory



**Figure A.1:** One fixed turn environment. This environment consists of four different trails. In all of them, the first turn is to the right. The green circle indicates where the organism is placed at the start of the trail, facing the next nutrient.



**Figure A.2:** Two fixed turns environment. This environment consists of four different trails. In all of them, the first turn is to the left and the second turn is to the right. The green circle indicates where the organism is placed at the start of the trail, facing the next nutrient.



**Figure A.3:** Nutrient cued environment. This environment consists of four different trails. In each of them, the number of nutrients before the first turn is even if the first turn is to the right and odd if the first turn is to the left. The green circle indicates where the organism is placed at the start of the trail, facing the next nutrient.



**Figure A.4:** Random start environment. This environment consists of four different trails. In all of them, the number of nutrients before the first turn is the same (3), and each of the four possible combinations of first and second turns are contemplated, therefore, making the start pattern truly random from the point of view of the organism. The green circle indicates where the organism is placed at the start of the trail, facing the next nutrient.



**Figure A.5:** Cue reversal environment. This environment consists of four different trails with the same start pattern as the nutrient cued environment. In each of them, the number of nutrients before the first turn is even if the first turn is to the right and odd if the first turn is to the left. The pink circle indicates the point where the turn cues are reversed. The green circle indicates where the organism is placed at the start of the trail, facing the next nutrient.

## A.2 Preliminary Experiment

This experiment was identical in design to experiment 1, with one exception. The instruction set we used did not contain the move back instruction. The only sensor and effector instructions were sense current, rotate right, rotate left and move ahead.

We performed 50 evolutionary replicates for each environment, except for the nutrient cued one where we extended the number of replicates to 150 because we had observed a high degree of variability and sought to determine whether a pattern would emerge.

### A.2.1 Preliminary Experiment – Results

As in experiment 1, the evolution of learning (i.e., imprinting) was rare, but it occurred in all predictable-start environments (one fixed turn, two fixed turns, and nutrient cued), and never in the control environment (random start). The predictable-start environments also produced a wider range of navigational strategies and much higher task quality scores than the random start environment. No replicate in the random start environment produced an organism that could complete the trail. The average maximum task quality we observed in this environment was 5% (table A.2) for an organism that used a rigid strategy of moving straight on forward cues and always turning left on a turn cue.

We analyzed over 100 out of the 300 replicates across all experimental conditions and found that, just as in experiments 1 and 2, organisms from different evolutionary replicates generated a consistent set of phenotypes. Their behavioral strategies always fell into one of the easily

recognizable categories described in table 2.2, with some hybrids of two or more strategies also observed.

In this experiment, error recovery evolved significantly less often in the predictable-start environments than in experiment 1, where the move back instruction was available (Fisher's exact test: one fixed turn, p = 0.0095; two fixed turns, p = 0.0125; nutrient cued, p < 0.0001). In addition, we observed a significantly lower average maximum task quality in all but the one fixed turn environment (Kruskal–Wallis test: one fixed turn, H = 3.0309, df = 1, p-value = 0.08169; Two Fixed Turns, H = 6.7623, df = 1, p-value = 0.00931; nutrient Cued, H = 19.522, df = 1, p < 0.0001; random start, H = 25.301, df = 1, p-value = 4.904e-07), with the largest effect in the nutrient cued environment (fig. S6).

The error recovery strategy that evolved in this experiment used quite an elaborate routine. When an organism encountered a turn cue, it would attempt to turn to one side. If its first turning attempt led it into an empty location, it would turn around on the spot (execute four 45 degree turns) and move one step forward in order to get back on the trail, then it would turn to the opposite direction that it had tried the first time and move forward. The organisms that used this strategy did not learn from their errors and performed the same behavior at every turn of the trail. However, this result demonstrated that even a basic Avida building block such as the move back behavior could evolve from an assembly of simpler actions.



**Figure A.6:** Distribution of Average Maximum Task Quality (AMTQ) per environment, comparing the preliminary experiment and experiment 1. Each violin plot represents the distribution of task qualities across replicates for a given environment. The preliminary experiment did not use the move back instruction (orange), experiment 1 did (blue). The difference between experiments was significant in all but the one fixed turn environment (Kruskal–Wallis test: one fixed turn, H = 3.0309, df = 1, p-value = 0.08169; two fixed turns, H = 6.7623, df = 1, p-value = 0.00931; nutrient cued, H = 19.522, df = 1, p < 0.0001; random start, H = 25.301, df = 1, p-value = 4.904e-07).

**Table A.2:** Preliminary experiment summary of results. Performance and strategies of the organisms with AMTQ equal or higher than 25%, organized by environment\*.

		Predictable-start environments			Control environment
		One fixed turn	Two fixed turns	Nutrient cued	Random start
Replicates in which organisms finished the trail	Proportion of replicates	13/50	7/50	4/150	0/50
	Strategies evolved (no. of replicates)	Imprinting (9) Error recovery (4)	Imprinting	Imprinting (3) Error recovery (1)	N/A
	Highest AMTQ observed (strategy)	99% (Imprinting)	98% (Imprinting)	98% (Imprinting)	N/A
Replicates in which organisms did not finish the trail (AMTQ ≥ 25%)	Proportion of replicates	9/50	7/50	17/150	0/50
	Strategies evolved	Imprinting Error recovery	Imprinting Error recovery	Error recovery Path predicting Searching Imprinting Hybrids of two or more strategies	N/A

\*Note: We examined only a sample of organisms that had less than 25% AMTQ. Those that were examined displayed previously described strategies and did not travel far on the trail.

In the course of our investigations, we came across intriguing possibilities that we plan to

explore in future studies. For example, in this preliminary experiment, we observed the

greatest variety of corrective actions that led the organisms back to the trail, since we were not

using the move back instruction. Most corrective actions led organisms to skip nutrients or take extra steps through empty locations, resulting in a lower task quality. However, we often observed these less efficient forms of error recovery in hybrid strategies with searching. This finding led us to suspect that error recovery could have evolved as an optimization of a more stereotypical searching strategy. Since error recovery is one of the most successful strategies we observed, it would be interesting to analyze the lineages where it evolved in this study, in order to reconstruct its evolutionary path in greater detail.

### A.2.2 Preliminary Experiment – Bestiary

In the one fixed turn environment, we observed other strategies and variations that were not among the top performers but are still worth mentioning. Some organisms evolved different path predicting strategies. Others evolved different methods of error recovery. In one, the organism did not possess a single rotate left instruction in its genome and performed all turns by rotating right a number of times. In another, if the organism made a mistake at a turn and stepped off the trail, it would bypass the turn cue and try to find the first nutrient after it. This behavior was reminiscent of a search procedure.

In the nutrient cued environment, we also observed some noteworthy strategies among the organisms that could not complete the trail. Most were hybrid strategies combining two or more behaviors (except relearning).

We found two organisms that used a flexible path predicting strategy based on a mnemonic rule and limited memory that could reach up to 29% of the maximum task quality. This strategy

combined some genetic encoding of trail patterns with temporary memory storage of cue values, which allowed the organism to distinguish a turn cue from the last one it had sensed, and repeat the turn in the same direction or not.

Another two organisms used a hybrid strategy combining path predicting and error recovery. They navigated the first portion of the trail based on genetically-encoded trail patterns, and as soon as the pattern no longer applied and they stepped off the trail, they would use error recovery in order to get back on the trail and continue to navigate, reaching up to 31% task quality.

A different hybrid strategy we observed in three organisms also started with path predicting. However, when their encoded pattern no longer applied and they stepped off the trail, they would continue to search for another portion of the trail and enter it, even if in the opposite direction. It appeared as if they hard-coded a portion of a trail within each trail.

We found one organism that used a hybrid strategy combining error recovery and searching. It primarily navigated by error recovery, however, if it encountered two left turns in a row it would not be able to find its way back to the trail. Instead, it continued searching for another stretch of the trail, and reentered it, even if in the opposite direction. It would then resume navigation by error recovery.

Finally, the best performing hybrid strategy combined imprinting, error recovery, and searching. This organism did not complete the path but could reach a task quality of 84%. It performed imprinting at the beginning of the path and would continue to navigate using the

associated cue in memory. However, similar to the previous organism, it had trouble navigating two left turns in a row and would step off the trail. Once off the trail, it would perform an error recovery behavior (moving and turning 90 degrees, four times) to get back on the trail. Then it would continue navigating based on the imprinted cue. However, when it reached the end of some trails, it would start a search behavior until it found another part of the trail and reentered it, even if in the opposite direction, in which case it navigated the trail by employing the error recovery strategy.

We examined only a sample of organisms that had less than 25% task quality. Of these, the ones examined displayed previously described strategies and did not travel far on the trail.

### A.3 Experiment 1 – Additional Results and Bestiary

In both the nutrient cued and one fixed turn environments, we found a variation of the error recovery strategy in which the organisms averted stepping off the trail in non-default turns if they were in close succession. They tended to repeat the non-default turn when encountering another one in a sequence. This behavior allowed them to make fewer "mistakes" and achieve a higher task quality (up to 87% of the maximum) than the organisms that used the typical form of error recovery.

Analyzing the lineages of the organisms that used the more rigid form of imprinting, which did not generalize to other trail configurations, we found that most evolved directly from ancestors that navigated by path predicting. These lineages never evolved error recovery, which explains the inability of these organisms to cope with mistakes that led them off the trail. Interestingly, in the one fixed turn environment, one of the organisms that used the more rigid form of imprinting evolved from ancestors that used error recovery, and that ability was eventually lost during evolution. When learning initially evolved in this lineage, there were ancestral organisms able to both imprint the cue association by using the pattern at the start of the trail, as well as error recover and relearn if the cues changed and they stepped off the trail. These overlapping abilities persisted over many generations, until the behavior was further streamlined, and only the more rigid form of imprinting remained, which provided a small advantage in task quality, in their particular environment. This case supports hypothesis 4 that there must be a need for reversal learning (or relearning) for it to fix, otherwise imprinting alone will suffice and could fix.

Similarly, one of the organisms that evolved generalizable imprinting had an ancestor capable of coping with cue changes along the trail. However, this ability arose from a "short memory" that required it to reassociate the cue at regular intervals, rather than an ability to relearn.



**Figure A.7:** Searching and imprinting hybrid strategy. This organism, which evolved in the nutrient cued environment, is an example of a searching and imprinting hybrid strategy that reached a task quality of 57% of the maximum in this trail. When the organism started navigating the trail, it reacted to turn cues by turning to a default direction. If this direction led it to step off the trail, it would initiate a search procedure, alternating forward moves and turns, until it found another portion of the trail and reentered it. This stint off the trail also primed it for imprinting the cue association the next time it encountered the non-default turn cue. After which it would navigate the remainder of the trail using the learned association.

# A.4 Experiment 2 – Additional Results

We also tested whether the position of the cue reversal on the trail affected the evolution of learning (imprinting and relearning). We used the same cue reversal environment, Turingcomplete instruction set, and sensing and moving instructions as in experiment 2, but varied the position at which the cues were reversed in the trail. We hypothesized that if the cue reversal took place early in the trail, organisms that navigated by imprinting would not have a sufficient selective advantage over those that navigated by error recovery. The results disproved our hypothesis. The position of the cue reversal on the trail had no significant effect on the number of organisms that evolved learning. Initially, we varied the cue reversal position from the 10% mark to the 90% mark of every trail, in 2.5% increments, evolving 200 populations for each increment. Although not significantly different (Fisher's exact test, p = 0.3803), we chose two points that had a large difference in the number of organisms that evolved learning (65% mark and 85% mark) and evolved 1000 populations at each (fig. A.8). Again, the difference was not significant (Fisher's exact test, p = 0.1497).



**Figure A.8:** Evolution of learning according to cue reversal position. Different cue reversal positions along the trail of nutrients, from 10% to 90% of the total length, in 2.5% intervals. In blue, the number of replicates from the first set that evolved learning out of 200 per position. In yellow, the second set with 1000 replicates per position.

#### A.4.1 Experiment 2 – Single Lineage Analysis

Below are phenotypic descriptions and figures from one of the lineages that evolved relearning in experiment 2. This lineage is typical of the evolutionary sequences we observed, where organisms first evolved the capacity for moving, then sensing, followed by a purely reflexive navigation strategy, such as path predicting, searching, or error recovery, and ultimately learning. In each figure, the background graph shows the task quality (TQ) over time, highlighting ancestors at major evolutionary transitions. The graphic on the right shows the path of a particular ancestor in one of the arenas of the cue reversal environment.

### Ancestral organism (TQ = 0.0; Update born: 1):

This was the ancestral organism common to all lineages. Its genome was blank except for a reproduce instruction. Therefore, this organism possessed no behavior and was sessile.
# Oscillator phenotype (TQ = 0.016; Update born: 697):

This organism oscillated in place, moving one step forward and one step back (fig. A.9). It was the first organism of the lineage with a real gain of task quality over the ancestral organism. It did not possess a sense current instruction.



**Figure A.9:** "Oscillator" organism. This organism moved back and forth between its start position (green circle) and its final position (red circle).

### Straight mover phenotype (TQ = 0.032; Update born: 1424):

This organism was the first to be capable of sensing and reacting to environmental cues (fig. A.10). Its behavior was divided in two phases. In the first phase, it moved straight forward, and continued moving while sensing the forward cue. When it no longer sensed the forward cue, it stopped moving forward and entered a second phase where it oscillated in place for the remainder of its lifetime (until it could reproduce). It did not possess rotate instructions.



Figure A.10: "Straight mover" organism. This organism moved forward only until it encountered a turn cue.

## Right turner phenotype (TQ = 0.056; Update born: 32019):

This organism moved forward when sensing forward cues, and always turned right upon sensing any turn cue (fig. A.11). If it encountered a left turn, it would turn to the right and move into an empty location, which caused the organism to stop moving and attempt to



Figure A.11: "Right turner" organism. This organism reacted to turn cues by always rotating right.

#### Path predictor phenotype (TQ = 0.073; Update born: 54071):

This organism moved forwards when sensing forward cues and reacted to the turn cues by turning either to the right or left, depending on how many steps it had moved and the direction of its last turn (fig. A.12). It had encoded a pattern in its behavioral algorithm that reflected the different starts of the four trails for this particular environment (nutrient cued). In other words, it had a limited flexible response that would differentially match the pattern of the first few segments of any of the four trails it could encounter. When the pattern no longer matched, and the organism made a "wrong" turn, it stopped moving, and attempted to reproduce. (In three of the paths it was able to pass the first two turns and in one of the paths it reached past the fifth turn.)



**Figure A.12:** "Path predictor" organism. This organism had an encoded pattern in its behavioral algorithm that reflected the different starts of the four trails for this particular environment (nutrient cued). When the pattern no longer matched the trail, and the organism made a "wrong" turn, it stopped moving.

#### Error recoverer 1 phenotype (TQ = 0.089; Update born: 64201):

This was the first organism in the ancestral lineage to use the error recovery strategy (fig. A.13). However, it moved slowly and reproduced before reaching the end of the trail. The organism reacted to the turn cues by first trying to turn to the right and moving straight. If that led it into an empty location, it took one step back, and tried rotating to the left in 45 degrees increments and moving straight until it found the trail of nutrients again.



**Figure A.13:** "Error recoverer 1" organism. This organism was the first in this lineage to use the error recovery strategy, however it often wasted movements, which made progress somewhat slow.

#### Error recoverer 2 phenotype (TQ = 0.81; Update born: 66346):

This was the last organism in the ancestral lineage to rely exclusively on the error recovery strategy (figs. A.14 and A.17). Once it evolved, this strategy became predominant in the population and improved in efficiency. This organism was able to navigate much further into the trail in a given time than its early ancestors. For example, when this organism made a wrong turn, instead of trying to find the trail again by turning in 45-degree increments like Error recoverer 1, it took one step back, turned 90 degrees at once, and moved forward.



**Figure A.14:** "Error recoverer 2" organism. This organism was the last one in the lineage to rely exclusively on the error recovery strategy. Its behavior was a streamlined version of its early ancestor (Error recoverer 1), which allowed it to move faster, waste fewer movements, and reach much further into the trail.

### First learner (TQ = 0.98; Update born: 66448):

This was the first organism in this lineage to use the relearning strategy (fig. A.15). Notably, it differed from its immediate ancestor (Error recoverer 2) by a single mutation, which connected the error-recovery module to a previously underutilized memory-storing module (figs. A.16 and A.17). Therefore, this organism evolved relearning directly from error recovery, without passing through imprinting.



**Figure A.15:** "First learner" organism. This organism was the first one in the lineage to use the relearning strategy. It differed from its immediate ancestor (Error recoverer 2) by a single mutation. Its path in one of the arenas (right) shows that that it went off the trail three times: once in the initial learning, a second time when the cues were reversed, and a third time when it detected the trail ended, at which point the organism stopped moving.

#### Final organism (TQ = 0.98; Update born: 249826):

The final predominant organism of the lineage used the same relearning strategy as the First learner and achieved the exact same task quality score. However, its behavioral control algorithm was slightly more compact, which allowed it to perform the same behavior as its ancestor more quickly and reproduce sooner, consequently gaining fitness. Figure A.18 shows a comparison between the First learner and the Final organism genomes. Although the Final organism's genome is longer, the executed portion of its program is slightly more efficient, partly due to shorter loops.



**Figure A.16:** One single mutation separates the error recovery and relearning behaviors. The transition from Error recoverer 2, on the left, to First learner, on the right, occurred due to a single mutation. For the full source code of the First Learner, see fig. A.18.



**Figure A.17:** Change in the behavioral algorithm from Error recoverer 2 to First learner due to a single mutation. The mutation connected the error-recovery module to the memory-storing module. Previously, memory-storing was executed only once, right after the organism was initialized on the trail (left sequence). After the mutation, every time it made a wrong turn and recovered, the organism stored the cue that led it off the trail in memory (right sequence).

Note: Each block in the diagram represents a sequence of one or more instructions. For the full source code of the First learner, see fig. A.18.



**Figure A.18:** Comparison between First learner and Final organism. The picture shows the genomes of the First learner and Final organism, side by side. Lines connecting both genomes indicate corresponding portions of the algorithm. The longer Final organism's genome indicates a substantial accumulation of neutral mutations during evolution. Interestingly, the active parts of the genome were highly conserved, and tended to remain together.

## A.4.2 Experiment 2 – Bestiary

In this experiment, several organisms evolved variations of imprinting in which the cue associations did not last long ("short memory"). The most common of these variations was a hybrid of imprinting and error recovery, where an imprinted association would last only as long as the turns were in the non-default direction. A turn to the default direction would erase the association. Afterwards, a turn to the non-default direction would cause the organism to step off the trail, recover and imprint the association again. These organisms coped with cue reversals by using error recovery until their memory lapsed and they could form a new association. They reached higher task quality scores (up to 93% of the maximum) than organisms using typical imprinting where the cue association did not lapse, but if the cues changed, the organism was not able to reform the association. In another variation of this

strategy, the cue association lapsed after the organism had performed a certain number of movements, and not due to encountering a specific type of cue.

# **A.5 Follow-Up Experiment**

As a follow up to experiment 2, which selected for reversal learning, we designed an experiment to test whether limiting the amount of computational memory available to an organism would make the evolution of learning behavior more difficult.

In all previous experiments, we used a Turing-complete instruction set, which provided an organism with six addressed memory positions, each capable of storing a single 32-bit integer, and two push-pop memory stacks, each capable of storing ten 32-bit integers [216]. This time, we used a non-Turing-complete, minimal memory instruction set. This new instruction set provided an organism with only two addressed memory positions, and no stacks. We used the same cue reversal environment (fig. A.5) and sensing and moving instructions as in experiment 2, and performed 200 evolutionary replicates under this condition.

Our hypothesis was that by providing the organisms with less memory to store information and perform computation, the evolution of learning would be more difficult. We believed that two registers were the minimum amount of memory necessary to solve the learning task, since one register is always used by the sense current instruction to store the currently sensed cue. A second register would be necessary to store a previously learned cue in order to compare them.

### A.5.1 Follow-Up Experiment – Results

Our hypothesis was not upheld. Relearning evolved under the minimal memory instruction set at a rate not significantly different from the Turing-complete instruction set used in experiment 2 (Fisher's exact test, p = 0.4191). In addition, when comparing the final predominant organisms of each condition, there was no significant difference in the distribution of AMTQ or average AMTQ (Kruskal–Wallis test, H = 0.11008, df = 1, p = 0.7401) (fig. A.19; table A.3).



**Figure A.19:** Distribution of Average Maximum Task Quality (AMTQ) per condition for experiment 2 and follow-up experiment. Each violin plot represents the distribution of AMTQ across replicates for a given condition. The difference between the two conditions was not significant (Kruskal–Wallis test, H = 0.11008, df = 1, p = 0.7401).

Therefore, the amount of available memory appears not be a constraint on the evolution of

learning in our system, as long as the minimum amount necessary to solve the problem is

provided.

**Table A.3:** Experiment 2 and follow-up experiment summary of results. Comparison between the two conditions, standard Turing-complete instruction set, and non-Turing-complete, minimal memory instruction set.

	Experiment 2	Follow-up experiment
Instruction set	Turing-complete, standard	Non-Turing-complete, minimal memory
Number of replicates	900	200
Average AMTQ across all replicates	0.4853988	0.4405631
Best performance	97%	97%
Organisms that evolved relearning:		
Behavior generalized to any trail configuration	10	5
Behavior did not generalize	8	1
Total	18	6

# A.6 How Evolution Continues to Shape Learning

In a final analysis of our data, we looked into how learning continued to evolve after it first appeared. Although our experiments were not designed to investigate this question, our data shows some intriguing patterns consistent with the literature on how evolution optimizes learning abilities for the particular characteristics of each environment.

### A.6.1 Evolutionary Constraints on Learning

Among the organisms that evolved learning (imprinting and relearning) in experiments 1 and 2, some were capable of learning any cue combination (in the range from 1 to 100) while others could learn some combinations but not others. In addition, some organisms learned from the experience of stepping off the path while others learned by using the pattern at the start of the trail to decide when and what to imprint. We looked at the evolution of these and other differences in learning ability in the context of the literature on preparedness and other so-called constraints on learning [84], [88], [89], [147]. A key theme in this literature is that the evolutionary process has produced learning mechanisms that are particularly efficient to form the kinds of associations that an animal may need to use in the environment where it evolved. For example, an animal that relies upon odors for foraging may more quickly learn to associate odors with good or bad foods than to associate visual cues with the same foods [151].

Our studies were not specifically designed to test this idea, but in a more general sense it is clear that digital organisms in some lineages evolved to learn those cue associations that were relevant, and to do so with high efficiency. Furthermore, we could actually examine the evolutionary history of those organisms that evolved the ability to learn, unlike most empirical papers in the "preparedness" literature, where the focus is on the outcome of evolution and not the evolutionary history itself. This enabled us to document specific examples of where the performance of learning mechanisms improved the ability to deal with the learning task presented by the environment, supporting the general assumption of the preparedness literature. For example, some of the organisms that could learn any cue combination had ancestors that could only learn some combinations (the combinations where one of the cues had the value of 1 was particularly troublesome due to the specific assembly instructions available in the Avida instruction set). This was due to some inefficiencies in their learning algorithm as it first evolved, that were later mitigated by evolution. Therefore, in these lineages, organisms improved their learning ability during evolution by increasing the range of cue values they could associate.

We also found some lineages where ancestral organisms required multiple exposures to learn the cue-response association (regardless if it was the initial learning or a reversal learning). However, their descendants gradually evolved to require a lesser number of exposures, and the final organisms were able to learn the cue-response association in a single exposure.

In many of the lineages we analyzed, there were ancestral organism with short memories. They could form the cue-response association, but it would not last long. The organism would have to either reform the association often or navigate the remainder of the trail using other strategies, such as searching or error recovery. However, in all these lineages, we observed a trend of increasing memory duration, with some of the final organisms being able to retain the association indefinitely. A long-lasting memory was advantageous for organisms in experiment 1, where there were no cue reversals, and for organisms using the relearning strategy in experiment 2, where the environment contained only one cue reversal.

We also observed lineages, in both experiments 1 and 2, where ancestral organisms could make the association under two different conditions. They could first imprint the cue association by

using the pattern at the start of the trail, and later, when their memory lapsed, they could imprint (or relearn) the association when they stepped off the trail. This dual ability was lost in the final organism, which specialized in one of the strategies, presumably because it simplified its behavioral algorithm and allowed it to navigate faster and reproduce sooner.

One of these lineages evolved in the one fixed turn environment in experiment 1. Some of its ancestral organisms were capable of relearning since they evolved from error recovery. Later generations also evolved imprinting based on the pattern at the start of the trail (the first turn was always to the right), and both abilities were simultaneously present in the same organism for many generations, until the ability for error recovery was eventually lost, thus completing the transition to imprinting based on the pattern at the start of the trail, which was the only strategy used by the final organism.

In another lineage that evolved in experiment 2, some of the ancestral organisms were capable of relearning since they evolved from error recovery, but they were also capable of imprinting based on the pattern at the start of the trail. Additionally, these organisms had short memories which required them to relearn often. However, in this case, it was the ability for imprinting that was eventually lost during evolution, and the final organism was capable of relearning exclusively.

Beyond this evidence of learning abilities improving during evolution, we observed variations in the outcome of evolution that were not necessarily predicted by a strict interpretation of the "preparedness" concept. In particular, in spite of the fact that the evolutionary pressures should have been similar across replicates in our experiments, we find a high degree of

variation across replicate lineages in the efficiency of particular learning mechanisms and how they arose. This suggests that a future area of study could be to understand historical contingency in the evolution of learning: i.e., to explore systematically how early evolutionary history influences subsequent evolution of learning abilities.

### A.6.2 Imprinting vs. Relearning – Evolution of the Sensitive Period for Learning

In experiment 1, where environments did not contain cue reversals, the only form of learning present in the final populations was imprinting. However, in experiment 2, where the environment contained cue reversals, we found both imprinting and relearning. Moreover, the highest performing imprinting organisms in experiment 2 had short memories, which would allow them to re-form the associations once their memory had lapsed and was beneficial in coping with cue reversals (Section A.4.2). We looked at these contrasting results from the two experiments from the perspective of the literature on sensitive periods of plasticity [148], [150]. In regards to the evolution of learning, the theory states that when environmental conditions are relatively stable during an individual's lifetime, such that the experiences early in life are reliable predictors of future conditions, natural selection will favor restricting learning to a period early in the individual's life, thus reducing the costs of learning. An example is filial imprinting in birds: a chick needs to learn who its mother is only once, since her identity will not change [149]. Conversely, if environmental conditions vary substantially within an individual's lifetime, natural selection will favor lifelong learning.

In experiment 1, environments were stable within the organism's lifetime, since there were no cue reversals, and the organisms that evolved imprinting learned the cue-response association

early in the trail and retained that association thereafter. They were incapable of relearning when tested in environments with cue reversals. Thus, their sensitive period for learning only lasted until the first learning event. However, in experiment 2, where the environment was less stable, since cues were reversed later in life, organisms evolved relearning and short-term imprinting and were able to re-form the association multiple times along the trail. Their sensitive period for learning lasted their entire lives.

For the organisms that evolved in the stable environments of experiment 1, remaining plastic after the initial learning would not have provided any additional benefit. However, in the cue reversal environment of experiment 2, retaining the ability to re-form the association after the initial learning was adaptive. Moreover, the reduction of the sensitive period may be adaptive in the stable environments of experiment 1, since evolution could presumably simplify an organism's algorithm, limiting the execution of the learning module to the early part of its life, thus allowing it to navigate more quickly the remainder of the trail and reproduce sooner.

When analyzing the evolutionary history of the organisms that evolved imprinting in experiment 1, we found some lineages where ancestral organisms had sensitive periods for learning that could last their entire life, that is, when tested in an environment with multiple cue reversals, these organisms could either relearn or re-form the cue-association after their memory had lapsed. Therefore, in some of the lineages of experiment 1, there was a trend towards reducing the sensitive period for learning to the early part of the organism's life. In contrast, in experiment 2, where the environment contained cue reversals, we never observed a reduction of sensitive periods in the lineages that evolved relearning or short-term imprinting.

As mentioned in the previous section (Section A.6.1), we also found lineages, both in experiment 1 and experiment 2, where ancestral organisms were capable of learning using two different strategies, imprinting and relearning. This dual ability was eventually lost during evolution and the final organism used only one of the strategies. In agreement with the theory, the strategy that prevailed in the environment without cue reversals was imprinting, while relearning prevailed in the environment with cue reversals.

In the future, we would like to perform more experiments targeted specifically at investigating the evolution of sensitive periods and shed light on these observations.

# **APPENDIX B**

Supplementary Material for Chapter 4



**Figure B.1.** Flowchart of the predominant organism capable of the generalizable version of configural learning from Experiment 2. Its cyclomatic complexity is 13.

REFERENCES

### REFERENCES

- [1] M. Mitchell, *Artificial Intelligence: A Guide for Thinking Humans*. Penguin UK, 2019.
- [2] R. Lachman, J. L. Lachman, and E. Butterfield, *Cognitive psychology and information processing : an introduction*. Hillsdale, N.J. : Lawrence Erlbaum Associates ; New York : distributed by Halsted Press, 1979. Accessed: Feb. 27, 2020. [Online]. Available: http://archive.org/details/cognitivepsychol00lach
- [3] P. McCorduck, *Machines who think: a personal inquiry into the history and prospects of artificial intelligence*, 25th anniversary update. Natick, Mass: A.K. Peters, 2004.
- [4] A. Newell and H. A. Simon, *Human problem solving*, vol. 104. Prentice-Hall Englewood Cliffs, NJ, 1972.
- [5] R. A. Brooks, *Flesh and Machines: How Robots Will Change Us*. Vintage Books, 2003.
- [6] R. Menzel and J. Fischer, Eds., *Animal Thinking: Contemporary Issues in Comparative Cognition*. The MIT Press, 2011. doi: 10.7551/mitpress/9780262016636.001.0001.
- [7] C. Gallistel, "Learning and Representation," *Learn. Theory Behav. Learn. Mem. Compr. Ref.*, vol. 1, Dec. 2008, doi: 10.1016/B978-012370509-9.00082-6.
- [8] A. Kamil, "A Synthetic Approach to the Study of Animal Intelligence," *Nebr. Symp. Motiv.*, p. 53, 1987.
- [9] R. Dukas, "Evolutionary Biology of Animal Cognition," Annu. Rev. Ecol. Evol. Syst., vol. 35, no. 1, pp. 347–374, 2004, doi: 10.1146/annurev.ecolsys.35.112202.130152.
- [10] S. J. Shettleworth, *Cognition, Evolution, and Behavior*. New York: Oxford University Press, 1998.
- [11] C. R. Gallistel, "Animal Cognition: The Representation of Space, Time and Number," Annu. Rev. *Psychol.*, vol. 40, no. 1, pp. 155–189, Jan. 1989, doi: 10.1146/annurev.ps.40.020189.001103.
- [12] D. H. Ballard, M. M. Hayhoe, and J. B. Pelz, "Memory Representations in Natural Tasks," *J. Cogn. Neurosci.*, vol. 7, no. 1, pp. 66–80, Jan. 1995, doi: 10.1162/jocn.1995.7.1.66.
- [13] D. Kahneman, *Thinking, fast and slow*, 1st pbk. ed. New York: Farrar, Straus and Giroux, 2013.
- [14] J. J. Gibson, *The ecological approach to visual perception*. New York, London: Psychology Press, 2015.
- [15] S. L. Hurley, *Consciousness in action*, 1. Harvard University Press paperback ed. Cambridge, Mass.: Harvard Univ. Press, 2002.

- [16] R. S. Marken, "Perceptual organization of behavior: A hierarchical control model of coordinated action.," J. Exp. Psychol. Hum. Percept. Perform., vol. 12, no. 3, p. 267, 19861101, doi: 10.1037/0096-1523.12.3.267.
- [17] D. C. Dennett, *From bacteria to Bach and back: the evolution of minds*, First edition. New York: W.W. Norton & Company, 2017.
- [18] T. Gomila and P. Calvo, "1 Directions for an Embodied Cognitive Science: Toward an Integrated Approach," in *Handbook of Cognitive Science*, P. Calvo and A. Gomila, Eds. San Diego: Elsevier, 2008, pp. 1–25. doi: 10.1016/B978-0-08-046616-3.00001-3.
- [19] W. T. Powers, *Behavior: The control of perception*. Oxford, England: Aldine, 1973, pp. xi, 296.
- [20] T. Van Gelder, "What Might Cognition Be, If Not Computation?," J. Philos., vol. 92, no. 7, pp. 345–381, 1995, doi: 10.2307/2941061.
- [21] L. de Bruin, A. Newen, and S. Gallagher, Eds., *The Oxford Handbook of 4E Cognition*. Oxford: Oxford University Press, 2018.
- [22] A. Newell, *Unified theories of cognition*. Cambridge, Mass: Harvard University Press, 1990.
- [23] P. Godfrey-Smith, "Environmental complexity and the evolution of cognition," in *The Evolution of Intelligence*, R. Sternberg and J. Kaufman, Eds. London: Lawrence Erlbaum Associates, 2002, pp. 233–249.
- [24] D. A. Levitis, W. Z. Lidicker, and G. Freund, "Behavioural biologists don't agree on what constitutes behaviour," *Anim. Behav.*, vol. 78, no. 1, pp. 103–110, Jul. 2009, doi: 10.1016/j.anbehav.2009.03.018.
- [25] A. Trewavas, "The foundations of plant intelligence," *Interface Focus*, vol. 7, no. 3, p. 20160098, Jun. 2017, doi: 10.1098/rsfs.2016.0098.
- [26] K. Aleklett and L. Boddy, "Fungal behaviour: a new frontier in behavioural ecology," *Trends Ecol. Evol.*, vol. 36, no. 9, pp. 787–796, Sep. 2021, doi: 10.1016/j.tree.2021.05.006.
- [27] I. M. De la Fuente *et al.*, "Evidence of conditioned behavior in amoebae," *Nat. Commun.*, vol. 10, no. 1, Art. no. 1, Aug. 2019, doi: 10.1038/s41467-019-11677-w.
- [28] P. Lyon, "The cognitive cell: bacterial behavior reconsidered," *Front. Microbiol.*, vol. 6, Apr. 2015, doi: 10.3389/fmicb.2015.00264.
- [29] D.-E. Nilsson and J. Marshall, "Lens eyes in protists," *Curr. Biol.*, vol. 30, no. 10, pp. R458–R459, May 2020, doi: 10.1016/j.cub.2020.01.077.

- [30] M. K. Trinh, M. T. Wayland, and S. Prabakaran, "Behavioural analysis of single-cell aneural ciliate, Stentor roeseli, using machine learning approaches," J. R. Soc. Interface, vol. 16, no. 161, p. 20190410, Dec. 2019, doi: 10.1098/rsif.2019.0410.
- [31] F. Cvrčková, V. Žárský, and A. Markoš, "Plant Studies May Lead Us to Rethink the Concept of Behavior," *Front. Psychol.*, vol. 7, 2016, doi: 10.3389/fpsyg.2016.00622.
- X. E. Barandiaran, E. Di Paolo, and M. Rohde, "Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-temporality in Action," *Adapt. Behav.*, vol. 17, no. 5, pp. 367–386, Oct. 2009, doi: 10.1177/1059712309343819.
- [33] C. Misselhorn, "Collective Agency and Cooperation in Natural and Artificial Systems," in Collective Agency and Cooperation in Natural and Artificial Systems: Explanation, Implementation and Simulation, C. Misselhorn, Ed. Cham: Springer International Publishing, 2015, pp. 3–24. doi: 10.1007/978-3-319-15515-9\_1.
- [34] D. A. Sanders, "Artificial intelligence for control engineering," *Control Eng.*, vol. 62, no. 2, p. 38+, Feb. 2015.
- [35] S. A. Umpleby, "A HISTORY OF THE CYBERNETICS MOVEMENT IN THE UNITED STATES," J. Wash. Acad. Sci., vol. 91, no. 2, pp. 54–66, 2005.
- [36] E. Pacherie, "The phenomenology of action: A conceptual framework," *Cognition*, vol. 107, no. 1, pp. 179–217, Apr. 2008, doi: 10.1016/j.cognition.2007.09.003.
- [37] W. R. Ashby, *An introduction to cybernetics*. New York: J. Wiley, 1956.
- [38] K. Sterelny, *The evolution of agency and other essays*. Cambridge, UK ; New York: Cambridge University Press, 2001.
- [39] D. C. Dennett, "Intentional Systems," J. Philos., vol. 68, no. 4, pp. 87–106, 1971, doi: 10.2307/2025382.
- [40] D. Dennett, *Intentional Systems Theory*. Oxford University Press, 2009. doi: 10.1093/oxfordhb/9780199262618.003.0020.
- [41] D. C. Dennett, *The intentional stance*, 7. printing. Cambridge, Mass.: MIT Press, 1998.
- [42] K. E. Boulding, "General Systems Theory-The Skeleton of Science," *Manag. Sci.*, vol. 2, no. 3, pp. 197–208, 1956.
- [43] A. J. Dzieciol and S. Mann, "Designs for life: protocell models in the laboratory," *Chem Soc Rev*, vol. 41, no. 1, pp. 79–85, 2012, doi: 10.1039/C1CS15211D.

- [44] A. Hanson, "Spontaneous electrical low-frequency oscillations: a possible role in Hydra and all living systems," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 376, no. 1820, p. 20190763, Mar. 2021, doi: 10.1098/rstb.2019.0763.
- [45] G. Buzsáki, *Rhythms of the Brain*. New York: Oxford University Press, 2006. doi: 10.1093/acprof:oso/9780195301069.001.0001.
- [46] R. A. Baines and M. Landgraf, "Neural development: The role of spontaneous activity," *Curr. Biol.*, vol. 31, no. 23, pp. R1513–R1515, Dec. 2021, doi: 10.1016/j.cub.2021.10.026.
- [47] J. Hernández-Orallo, B. S. Loe, L. Cheke, F. Martínez-Plumed, and S. Ó hÉigeartaigh, "General intelligence disentangled via a generality metric for natural and artificial intelligence," *Sci. Rep.*, vol. 11, no. 1, p. 22822, Nov. 2021, doi: 10.1038/s41598-021-01997-7.
- [48] A. Visioli, "Maximizing the Impact of Control at All Levels," *Front. Control Eng.*, vol. 1, 2020, doi: 10.3389/fcteg.2020.602469.
- [49] S. Bennett, "CONTROL AND THE DIGITAL COMPUTER: THE EARLY YEARS," p. 6.
- [50] T.-M. Yi, Y. Huang, M. I. Simon, and J. Doyle, "Robust perfect adaptation in bacterial chemotaxis through integral feedback control," *Proc. Natl. Acad. Sci.*, vol. 97, no. 9, pp. 4649–4653, Apr. 2000, doi: 10.1073/pnas.97.9.4649.
- [51] S. A. Frank, "Homeostasis, environmental tracking and phenotypic plasticity. I. A robust control theory approach to evolutionary design tradeoffs," *bioRxiv*, p. 332999, May 2018, doi: 10.1101/332999.
- [52] S. A. Frank, *Control Theory Tutorial: Basic Concepts Illustrated by Software Examples*. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-319-91707-8.
- [53] A. Mitra, A.-M. Raicu, S. L. Hickey, L. A. Pile, and D. N. Arnosti, "Soft repression: Subtle transcriptional regulation with global impact," *BioEssays*, vol. 43, no. 2, p. 2000231, 2021, doi: https://doi.org/10.1002/bies.202000231.
- [54] A. Kurz, "Physiology of Thermoregulation," *Best Pract. Res. Clin. Anaesthesiol.*, vol. 22, no. 4, pp. 627–644, Dec. 2008, doi: 10.1016/j.bpa.2008.06.004.
- [55] J. A. Russell, G. Leng, and A. J. Douglas, "The magnocellular oxytocin system, the fount of maternity: adaptations in pregnancy," *Front. Neuroendocrinol.*, vol. 24, no. 1, pp. 27–61, Jan. 2003, doi: 10.1016/S0091-3022(02)00104-8.
- [56] H. El-samad, J. P. Goff, and M. Khammash, "Calcium Homeostasis and Parturient Hypocalcemia: An Integral Feedback Perspective," J. Theor. Biol., vol. 214, no. 1, pp. 17–29, Jan. 2002, doi: 10.1006/jtbi.2001.2422.
- [57] B. F. Skinner, "The evolution of behavior," J. Exp. Anal. Behav., vol. 41, no. 2, pp. 217–221, 1984.

- [58] A. Kurakin, "Stochastic Cell," IUBMB Life, vol. 57, no. 2, pp. 59–63, 2005, doi: 10.1080/15216540400024314.
- [59] D. J. Nicholson, "Is the cell really a machine?," J. Theor. Biol., vol. 477, pp. 108–126, Sep. 2019, doi: 10.1016/j.jtbi.2019.06.002.
- [60] F. J. Gomez and R. Miikkulainen, "Active Guidance for a Finless Rocket Using."
- [61] J. Yosinski, J. Clune, D. Hidalgo, S. Nguyen, J. C. Zagal, and H. Lipson, "Evolving Robot Gaits in Hardware: the HyperNEAT Generative Encoding Vs. Parameter Optimization," p. 8.
- [62] S. P. Franklin, *Artificial Minds*. Cambridge, MA, USA: A Bradford Book, 1995.
- [63] N. Kitadai and S. Maruyama, "Origins of building blocks of life: A review," *Geosci. Front.*, vol. 9, no. 4, pp. 1117–1153, Jul. 2018, doi: 10.1016/j.gsf.2017.07.007.
- [64] L. Bich, "Autonomous Systems and the Place of Biology Among Sciences. Perspectives for an Epistemology of Complex Systems," in *Multiplicity and Interdisciplinarity: Essays in Honor of Eliano Pessa*, G. Minati, Ed. Cham: Springer International Publishing, 2021, pp. 41–57. doi: 10.1007/978-3-030-71877-0\_4.
- [65] R. I. Vane-Wright, "What is life? And what might be said of the role of behaviour in its evolution?: What is Life?," *Biol. J. Linn. Soc.*, vol. 112, no. 2, pp. 219–241, Jun. 2014, doi: 10.1111/bij.12300.
- [66] F. G. Varela, H. R. Maturana, and R. Uribe, "Autopoiesis: The organization of living systems, its characterization and a model," *Biosystems*, vol. 5, no. 4, pp. 187–196, May 1974, doi: 10.1016/0303-2647(74)90031-8.
- [67] H. R. Maturana and F. J. Varela, *Autopoiesis and Cognition: The Realization of the Living*, vol. 42. Dordrecht: Springer Netherlands, 1980. doi: 10.1007/978-94-009-8947-4.
- [68] M. Vitas and A. Dobovišek, "Towards a General Definition of Life," *Orig. Life Evol. Biospheres*, vol. 49, no. 1, pp. 77–88, Jun. 2019, doi: 10.1007/s11084-019-09578-5.
- [69] F. Wong and J. Gunawardena, "Gene Regulation in and out of Equilibrium," *Annu. Rev. Biophys.*, vol. 49, no. 1, pp. 199–226, 2020, doi: 10.1146/annurev-biophys-121219-081542.
- [70] L. von Bertalanffy, "The Theory of Open Systems in Physics and Biology," *Science*, vol. 111, no. 2872, pp. 23–29, Jan. 1950, doi: 10.1126/science.111.2872.23.
- [71] P. C. W. Davies, *The demon in the machine: how hidden webs of information are solving the mystery of life*. London: Allen Lane, 2019.
- [72] A. W. Fenton, "Allostery: an illustrated definition for the 'second secret of life," *Trends Biochem. Sci.*, vol. 33, no. 9, pp. 420–425, Sep. 2008, doi: 10.1016/j.tibs.2008.05.009.

- [73] A. Jones, Ed., "Grand Research Challenges in Information Systems," *Rep. First Conf. Inf. Syst.*, p. 35, 2003.
- [74] R. Dale, "GPT-3: What's it good for?," Nat. Lang. Eng., vol. 27, no. 1, pp. 113–118, Jan. 2021, doi: 10.1017/S1351324920000601.
- [75] A. C. Pontes, R. B. Mobley, C. Ofria, C. Adami, and F. C. Dyer, "The Evolutionary Origin of Associative Learning," Am. Nat., vol. 195, no. 1, pp. E1–E19, Jan. 2020, doi: 10.1086/706252.
- [76] D. Hume, A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning Into Moral Subjects. London: Oxford University Press, 1738.
- [77] P. Godfrey-Smith, *Complexity and the Function of Mind in Nature*. New York: Cambridge University Press, 1996. doi: 10.1017/CBO9781139172714.
- [78] B. H. Weber and D. J. Depew, *Evolution and learning: The Baldwin effect reconsidered*. Cambridge, Mass: MIT Press, 2003.
- [79] R. A. Duckworth, "The role of behavior in evolution: a search for mechanism," Evol. Ecol., vol. 23, no. 4, pp. 513–531, Jul. 2009, doi: 10.1007/s10682-008-9252-6.
- [80] S. Ginsburg and E. Jablonka, "The evolution of associative learning: A factor in the Cambrian explosion," *J. Theor. Biol.*, vol. 266, no. 1, pp. 11–20, Sep. 2010, doi: 10.1016/j.jtbi.2010.06.017.
- [81] R. L. Brown, "Learning, evolvability and exploratory behaviour: extending the evolutionary reach of learning," *Biol. Philos.*, vol. 28, no. 6, pp. 933–955, Nov. 2013, doi: 10.1007/s10539-013-9396-9.
- [82] D. C. Dennett, *Darwin's dangerous idea: evolution and the meanings of life*. New York: Touchstone, 1996.
- [83] R. Dukas, "Effects of learning on evolution: robustness, innovation and speciation," *Spec. Sect. Behav. Plast. Evol.*, vol. 85, no. 5, pp. 1023–1030, May 2013, doi: 10.1016/j.anbehav.2012.12.030.
- [84] M. E. Seligman, "On the generality of the laws of learning," *Psychol. Rev.*, vol. 77, no. 5, pp. 406–418, 1970, doi: 10.1037/h0029790.
- [85] D. W. Stephens, "Change, regularity, and value in the evolution of animal learning," *Behav. Ecol.*, vol. 2, no. 1, pp. 77–89, Mar. 1991, doi: 10.1093/beheco/2.1.77.
- [86] F. Mery and T. J. Kawecki, "Experimental evolution of learning ability in fruit flies," Proc. Natl. Acad. Sci., vol. 99, no. 22, p. 14274, Oct. 2002, doi: 10.1073/pnas.222371199.
- [87] R. Dukas and J. M. Ratcliffe, Eds., Cognitive Ecology II. Chicago; London: University of Chicago Press, 2009. doi: 10.7208/chicago/9780226169378.001.0001.

- [88] S. J. Shettleworth, *Cognition, Evolution and Behavior*. Oxford; New York: Oxford University Press, 2010.
- [89] M. Domjan, "Biological or Evolutionary Constraints on Learning," in Encyclopedia of the Sciences of Learning, N. M. Seel, Ed. Boston, MA: Springer US, 2012, pp. 461–463. doi: 10.1007/978-1-4419-1428-6\_89.
- [90] B. R. Moore, "The evolution of learning," *Biol. Rev.*, vol. 79, no. 2, pp. 301–335, May 2004, doi: 10.1017/S1464793103006225.
- [91] A. S. Dunlap, M. W. Austin, and A. Figueiredo, "Components of change and the evolution of learning in theory and experiment," *Anim. Behav.*, vol. 147, pp. 157–166, Jan. 2019, doi: 10.1016/j.anbehav.2018.05.024.
- [92] A. S. Dunlap and D. W. Stephens, "Components of change in the evolution of learning and unlearned preference," *Proc. R. Soc. B Biol. Sci.*, vol. 276, no. 1670, pp. 3201–3208, Sep. 2009, doi: 10.1098/rspb.2009.0602.
- [93] G. F. Miller and P. M. Todd, "Exploring Adaptive Agency I: Theory and Methods for Simulating the Evolution of Learning," in *Connectionist Models*, D. S. Touretzky, J. L. Elman, T. J. Sejnowski, and G. E. Hinton, Eds. San Mateo, CA: Morgan Kaufmann, 1991, pp. 65–80. doi: 10.1016/B978-1-4832-1448-1.50013-5.
- [94] M. J. Wells, "Sensitization and the Evolution of Associative Learning," in Neurobiology of Invertebrates: Proceedings of the Symposium Held at the Biological Research Institute of the Hungarian Academy of Sciences (Tihany) September 4–7, 1967, J. Salánki, Ed. Boston, MA: Springer US, 1968, pp. 391–411. doi: 10.1007/978-1-4615-8618-0\_28.
- [95] G. Razran, *Mind in Evolution: An East-West Synthesis of Learned Behavior and Cognition*. Boston: Houghton Mifflin, 1971.
- [96] R. D. Hawkins and E. R. Kandel, "Is there a cell-biological alphabet for simple forms of learning?," *Psychol. Rev.*, vol. 91, no. 3, pp. 375–391, 1984, doi: 10.1037/0033-295X.91.3.375.
- [97] R. D. Hawkins and E. R. Kandel, "Steps toward a cell-biological alphabet for elementary forms of learning," in *Neurobiology of Learning and Memory*, G. Lynch, McGaugh James L, and N. M. Weinberger, Eds. New York: Guilford Press, 1984, pp. 385–404.
- [98] M. van Duijn, "Phylogenetic origins of biological cognition: convergent patterns in the early evolution of learning," *Interface Focus*, vol. 7, no. 3, p. 20160158, Jun. 2017, doi: 10.1098/rsfs.2016.0158.
- [99] J. S. Duerr and W. G. Quinn, "Three Drosophila mutations that block associative learning also affect habituation and sensitization," Proc. Natl. Acad. Sci., vol. 79, no. 11, pp. 3646–3650, Jun. 1982, doi: 10.1073/pnas.79.11.3646.

- [100] A. C. Roberts and D. L. Glanzman, "Learning in Aplysia: looking at synaptic plasticity from both sides," Trends Neurosci., vol. 26, no. 12, pp. 662–670, 2003, doi: https://doi.org/10.1016/j.tins.2003.09.014.
- [101] M. Gagliano, V. V. Vyazovskiy, A. A. Borbély, M. Grimonprez, and M. Depczynski, "Learning by Association in Plants," *Sci. Rep.*, vol. 6, p. 38427, Dec. 2016.
- [102] H. L. Armus, A. R. Montgomery, and J. L. Jellison, "Discrimination Learning in Paramecia (*P. caudatum*)," *Psychol. Rec.*, vol. 56, no. 4, pp. 489–498, Oct. 2006, doi: 10.1007/BF03396029.
- [103] Fernando Chrisantha T et al., "Molecular circuits for associative learning in single-celled organisms," J. R. Soc. Interface, vol. 6, no. 34, pp. 463–469, May 2009, doi: 10.1098/rsif.2008.0344.
- [104] T. J. Ord and E. P. Martins, "Evolution of behaviour: phylogeny and the origin of present day diversity," in *Evolutionary Behavioral Ecology*, Westneat, David and Fox, Charles W, Eds. New York: Oxford University Press, 2010, pp. 108–128.
- [105] J. B. Losos, "Seeing the Forest for the Trees: The Limitations of Phylogenies in Comparative Biology: (American Society of Naturalists Address)," Am. Nat., vol. 177, no. 6, pp. 709–727, 2011, doi: 10.1086/660020.
- [106] L. M. Grabowski, D. M. Bryson, F. C. Dyer, C. Ofria, and R. T. Pennock, "Early Evolution of Memory Usage in Digital Organisms.," in Artificial Life XII: Proceedings of the Twelfth International Conference on the Synthesis and Simulation of Living Systems, Cambridge MA, 2010, pp. 224–231.
- [107] R. T. Pennock, "Models, simulations, instantiations, and evidence: the case of digital evolution," J. *Exp. Theor. Artif. Intell.*, vol. 19, no. 1, pp. 29–42, Mar. 2007, doi: 10.1080/09528130601116113.
- [108] C. O. Wilke, J. L. Wang, C. Ofria, R. E. Lenski, and C. Adami, "Evolution of digital organisms at high mutation rates leads to survival of the flattest," *Nature*, vol. 412, no. 6844, pp. 331–333, Jul. 2001, doi: 10.1038/35085569.
- [109] R. E. Lenski, C. Ofria, R. T. Pennock, and C. Adami, "The evolutionary origin of complex features," *Nature*, vol. 423, no. 6936, pp. 139–144, May 2003, doi: 10.1038/nature01568.
- [110] S. S. Chow, C. O. Wilke, C. Ofria, R. E. Lenski, and C. Adami, "Adaptive Radiation from Resource Competition in Digital Organisms," *Science*, vol. 305, no. 5680, p. 84, Jul. 2004, doi: 10.1126/science.1096307.
- [111] F. M. Codoñer, J.-A. Darós, R. V. Solé, and S. F. Elena, "The Fittest versus the Flattest: Experimental Confirmation of the Quasispecies Effect with Subviral Pathogens," *PLOS Pathog.*, vol. 2, no. 12, p. e136, Dec. 2006, doi: 10.1371/journal.ppat.0020136.
- [112] L. M. Grabowski, W. R. Elsberry, C. Ofria, and R. T. Pennock, "On the Evolution of Motility and Intelligent Tactic Response," in *Proceedings of the 10th Annual Conference on Genetic and*

*Evolutionary Computation*, New York, NY, USA, 2008, pp. 209–216. doi: 10.1145/1389095.1389129.

- [113] C. Ofria, D. M. Bryson, and C. O. Wilke, "Avida," in *Artificial life models in software*, M. Komosinski and A. Adamatzky, Eds. London: Springer London, 2009, pp. 3–35.
- [114] C. Ofria, C. T. Brown, and C. Adami, *Avida*. 2015. [Online]. Available: https://github.com/mercere99/Avida-AssociativeMemory
- [115] F. C. Dyer, "Bees acquire route-based memories but not cognitive maps in a familiar landscape," Anim. Behav., vol. 41, no. 2, pp. 239–246, Feb. 1991, doi: 10.1016/S0003-3472(05)80475-0.
- [116] F. C. Dyer, "Cognitive ecology of navigation," in Cognitive ecology: The evolutionary ecology of information processing and decision making, Chicago, IL, US: University of Chicago Press, 1998, pp. 201–260.
- [117] T. S. Collett and M. Collett, "Memory use in insect visual navigation," Nat. Rev. Neurosci., vol. 3, no. 7, pp. 542–552, Jul. 2002, doi: 10.1038/nrn872.
- [118] L. M. Grabowski, "The evolutionary origins of memory use in navigation," PhD Thesis, Michigan State University, East Lansing, MI, 2009. [Online]. Available: https://search.proquest.com/docview/304931513
- [119] S. W. Zhang, K. Bartsch, and M. V. Srinivasan, "Maze Learning by Honeybees," Neurobiol. Learn. Mem., vol. 66, no. 3, pp. 267–282, Nov. 1996, doi: 10.1006/nlme.1996.0069.
- [120] S. W. Zhang, M. Lehrer, and M. V. Srinivasan, "Honeybee Memory: Navigation by Associative Grouping and Recall of Visual Stimuli," *Neurobiol. Learn. Mem.*, vol. 72, no. 3, pp. 180–201, Nov. 1999, doi: 10.1006/nlme.1998.3901.
- [121] E. L. Charnov, "Optimal Foraging: Attack Strategy of a Mantid," Am. Nat., vol. 110, no. 971, pp. 141–151, 1976.
- [122] A. C. Pontes, R. B. Mobley, C. Ofria, C. Adami, and F. C. Dyer, "Data from: The Evolutionary Origin of Associative Learning." American Naturalist, Dryad Digital Repository, Aug. 11, 2019. [Online]. Available: http://doi:10.5061/dryad.f45gh6s
- [123] R. Hadar and R. Menzel, "Memory Formation in Reversal Learning of the Honeybee," *Front. Behav. Neurosci.*, vol. 4, p. 186, 2010, doi: 10.3389/fnbeh.2010.00186.
- [124] G. B. Bissonette and E. M. Powell, "Reversal learning and attentional set-shifting in mice," *Neuropharmacology*, vol. 62, no. 3, pp. 1168–1174, 2012.
- [125] G. Xue, F. Xue, V. Droutman, Z.-L. Lu, A. Bechara, and S. Read, "Common Neural Mechanisms Underlying Reversal Learning by Reward and Punishment," *PLOS ONE*, vol. 8, no. 12, p. e82169, Dec. 2013, doi: 10.1371/journal.pone.0082169.

- [126] A. S. Dunlap and D. W. Stephens, "Reliability, uncertainty, and costs in the evolution of animal learning," *Curr. Opin. Behav. Sci.*, vol. 12, pp. 73–79, Dec. 2016, doi: 10.1016/j.cobeha.2016.09.010.
- [127] Z. D. Blount, C. Z. Borland, and R. E. Lenski, "Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*," *Proc. Natl. Acad. Sci.*, vol. 105, no. 23, pp. 7899–7906, Jun. 2008, doi: 10.1073/pnas.0803151105.
- [128] R. E. Lenski, "Experimental evolution and the dynamics of adaptation and genome evolution in microbial populations," *ISME J.*, vol. 11, no. 10, pp. 2181–2194, Oct. 2017, doi: 10.1038/ismej.2017.69.
- [129] O. S. Soyer and R. A. Goldstein, "Evolution of response dynamics underlying bacterial chemotaxis," *BMC Evol. Biol.*, vol. 11, no. 1, p. 240, Dec. 2011, doi: 10.1186/1471-2148-11-240.
- [130] C. Ofria, W. Huang, and E. Torng, "On the Gradual Evolution of Complexity and the Sudden Emergence of Complex Features," *Artif. Life*, vol. 14, no. 3, pp. 255–263, May 2008, doi: 10.1162/artl.2008.14.3.14302.
- [131] H. H. McAdams, B. Srinivasan, and A. P. Arkin, "The evolution of genetic regulatory systems in bacteria," *Nat. Rev. Genet.*, vol. 5, no. 3, pp. 169–178, Mar. 2004, doi: 10.1038/nrg1292.
- [132] N. Kashtan and U. Alon, "Spontaneous evolution of modularity and network motifs," *Proc. Natl. Acad. Sci.*, vol. 102, no. 39, pp. 13773–13778, Sep. 2005, doi: 10.1073/pnas.0503610102.
- [133] G. P. Wagner, M. Pavlicev, and J. M. Cheverud, "The road to modularity," Nat. Rev. Genet., vol. 8, no. 12, pp. 921–931, Dec. 2007, doi: 10.1038/nrg2267.
- [134] T. D. Johnston, "Selective Costs and Benefits in the Evolution of Learning," in Advances in the Study of Behavior, vol. 12, J. S. Rosenblatt, R. A. Hinde, C. Beer, and M.-C. Busnel, Eds. New York: Academic Press, 1982, pp. 65–106. doi: 10.1016/S0065-3454(08)60046-7.
- [135] C. Carbone and G. M. Narbonne, "When Life Got Smart: The Evolution of Behavioral Complexity Through the Ediacaran and Early Cambrian of NW Canada," J. Paleontol., vol. 88, no. 2, pp. 309– 330, 2014, doi: 10.1666/13-066.
- [136] P. M. Todd and G. F. Miller, "Exploring adaptive agency II: Simulating the evolution of associative learning," in *Proceedings of the first international conference on simulation of adaptive behavior on From animals to animats*, 1991, pp. 306–315.
- [137] E. Izquierdo and I. Harvey, "The Dynamics of Associative Learning in an Evolved Situated Agent," in *Advances in Artificial Life*, Berlin, Heidelberg, 2007, pp. 365–374.
- [138] E. Izquierdo, I. Harvey, and R. D. Beer, "Associative Learning on a Continuum in Evolved Dynamical Neural Networks," *Adapt. Behav.*, vol. 16, no. 6, pp. 361–384, Dec. 2008, doi: 10.1177/1059712308097316.

- [139] C. Breazeal, *Designing Sociable Robots*. Cambridge, MA: The MIT Press, 2004. doi: 10.7551/mitpress/2376.001.0001.
- [140] J. Panksepp, *Affective neuroscience: The foundations of human and animal emotions*. Oxford: Oxford University Press, 2004.
- [141] A. R. Damasio, *Descartes' error: emotion, reason, and the human brain*. London: Penguin, 2005.
- [142] S. Singh, R. L. Lewis, and A. Barto, "Where Do Rewards Come From?," *Proc. Annu. Meet. Cogn. Sci. Soc.*, vol. 31, no. 31, Jan. 2009.
- [143] S. Singh, R. L. Lewis, A. G. Barto, and J. Sorg, "Intrinsically Motivated Reinforcement Learning: An Evolutionary Perspective," *IEEE Trans. Auton. Ment. Dev.*, vol. 2, no. 2, pp. 70–82, Jun. 2010, doi: 10.1109/TAMD.2010.2051031.
- [144] A. B. Bond, A. C. Kamil, and R. P. Balda, "Serial reversal learning and the evolution of behavioral flexibility in three species of North American corvids (*Gymnorhinus cyanocephalus, Nucifraga columbiana, Aphelocoma californica*)," J. Comp. Psychol., vol. 121, no. 4, pp. 372–379, 2007, doi: 10.1037/0735-7036.121.4.372.
- [145] M. Cauchoix, E. Hermer, A. S. Chaine, and J. Morand-Ferron, "Cognition in the field: comparison of reversal learning performance in captive and wild passerines," *Sci. Rep.*, vol. 7, no. 1, p. 12945, Oct. 2017, doi: 10.1038/s41598-017-13179-5.
- [146] S. D. Buechel, A. Boussard, A. Kotrschal, W. van der Bijl, and N. Kolm, "Brain size affects performance in a reversal-learning test," *Proc. R. Soc. B Biol. Sci.*, vol. 285, no. 1871, p. 20172031, Jan. 2018, doi: 10.1098/rspb.2017.2031.
- [147] A. S. Dunlap, "Biological Preparedness," in *Encyclopedia of Animal Cognition and Behavior*, J. Vonk and T. Shackelford, Eds. Cham: Springer International Publishing, 2017, pp. 1–7. doi: 10.1007/978-3-319-47829-6\_1301-1.
- [148] P. Bateson, "How do sensitive periods arise and what are they for?," *Anim. Behav.*, vol. 27, pp. 470–486, May 1979, doi: 10.1016/0003-3472(79)90184-2.
- [149] E. Cashdan, "A sensitive period for learning about food," Hum. Nat., vol. 5, no. 3, pp. 279–291, Sep. 1994, doi: 10.1007/BF02692155.
- [150] T. W. Fawcett and W. E. Frankenhuis, "Adaptive explanations for sensitive windows in development," *Front. Zool.*, vol. 12, no. 1, p. S3, Aug. 2015, doi: 10.1186/1742-9994-12-S1-S3.
- [151] A. S. Dunlap and D. W. Stephens, "Experimental evolution of prepared learning," *Proc. Natl. Acad. Sci.*, vol. 111, no. 32, pp. 11750–11755, Aug. 2014, doi: 10.1073/pnas.1404176111.
- [152] G. E. Budd, "Early animal evolution and the origins of nervous systems," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 370, no. 1684, p. 20150037, Dec. 2015, doi: 10.1098/rstb.2015.0037.

- [153] R. E. Plotnick and K. Koy, "LET US PREY: SIMULATIONS OF GRAZING TRACES IN THE FOSSIL RECORD," p. 13, 2005.
- [154] G. E. Budd and S. Jensen, "The origin of the animals and a 'Savannah' hypothesis for early bilaterian evolution: Early evolution of the animals," *Biol. Rev.*, vol. 92, no. 1, pp. 446–473, Feb. 2017, doi: 10.1111/brv.12239.
- [155] M. L. Droser, L. G. Tarhan, and J. G. Gehling, "The Rise of Animals in a Changing Environment: Global Ecological Innovation in the Late Ediacaran," *Annu. Rev. Earth Planet. Sci.*, vol. 45, no. 1, pp. 593–617, Aug. 2017, doi: 10.1146/annurev-earth-063016-015645.
- [156] M. L. Droser and J. G. Gehling, "The advent of animals: The view from the Ediacaran," *Proc. Natl. Acad. Sci.*, vol. 112, no. 16, pp. 4865–4870, Apr. 2015, doi: 10.1073/pnas.1403669112.
- [157] C. Carbone and G. M. Narbonne, "When life got smart: the evolution of behavioral complexity through the Ediacaran and early Cambrian of NW Canada," J. Paleontol., vol. 88, no. 2, pp. 309– 330, 2014.
- [158] L. A. Buatois and M. G. Mángano, "Ediacaran Ecosystems and the Dawn of Animals," in *The Trace-Fossil Record of Major Evolutionary Events*, vol. 39, M. G. Mángano and L. A. Buatois, Eds. Dordrecht: Springer Netherlands, 2016, pp. 27–72. doi: 10.1007/978-94-017-9600-2\_2.
- [159] M. Gingras *et al.*, "Possible evolution of mobile animals in association with microbial mats," *Nat. Geosci.*, vol. 4, no. 6, pp. 372–375, Jun. 2011, doi: 10.1038/ngeo1142.
- [160] S. D. Evans, I. V. Hughes, J. G. Gehling, and M. L. Droser, "Discovery of the oldest bilaterian from the Ediacaran of South Australia," *Proc. Natl. Acad. Sci.*, Mar. 2020, doi: 10.1073/pnas.2001045117.
- [161] A. Seilacher, "Fossil Behavior," Sci. Am., vol. 217, no. 2, pp. 72–83, 1967.
- [162] T. Monk and M. G. Paulin, "Predation and the Origin of Neurones," Brain. Behav. Evol., vol. 84, no. 4, pp. 246–261, 2014, doi: 10.1159/000368177.
- [163] S. Ginsburg and E. Jablonka, "The evolution of associative learning: A factor in the Cambrian explosion," *J. Theor. Biol.*, vol. 266, no. 1, pp. 11–20, 2010.
- [164] B. Hayes, "Computing Science: In Search of the Optimal Scumsucking Bottomfeeder," Am. Sci., vol. 91, no. 5, pp. 392–396, 2003.
- [165] R. Gougeon, D. Néraudeau, A. Loi, and M. Poujol, "New insights into the early evolution of horizontal spiral trace fossils and the age of the Brioverian series (Ediacaran–Cambrian) in Brittany, NW France," *Geol. Mag.*, pp. 1–11, Jan. 2021, doi: 10.1017/S0016756820001430.
- [166] A. Seilacher, *Trace fossil analysis*. Berlin: Springer, 2007.

- [167] R. Richter, "Flachseebeobachtungen zur Paläontologie und Geologie. 9. Zur Deutung rezenter und fossiler Mäander-Figuren," *Senckenbergiana*, vol. 6, pp. 141–157, 1924.
- [168] D. M. Raup and A. Seilacher, "Fossil Foraging Behavior: Computer Simulation," *Science*, vol. 166, no. 3908, pp. 994–995, 1969.
- [169] J. G. Gehling and M. L. Droser, "Textured organic surfaces associated with the Ediacara biota in South Australia," *Earth-Sci. Rev.*, vol. 96, no. 3, pp. 196–206, Oct. 2009, doi: 10.1016/j.earscirev.2009.03.002.
- [170] S. Xiao, Z. Chen, C. Zhou, and X. Yuan, "Surfing in and on microbial mats: Oxygen-related behavior of a terminal Ediacaran bilaterian animal," *Geology*, vol. 47, no. 11, pp. 1054–1058, Nov. 2019, doi: 10.1130/G46474.1.
- [171] W. Ding *et al.*, "Early animal evolution and highly oxygenated seafloor niches hosted by microbial mats," *Sci. Rep.*, vol. 9, no. 1, pp. 1–11, Sep. 2019, doi: 10.1038/s41598-019-49993-2.
- [172] H. Wang *et al.*, "A benthic oxygen oasis in the early Neoproterozoic ocean," *Precambrian Res.*, vol. 355, p. 106085, Apr. 2021, doi: 10.1016/j.precamres.2020.106085.
- [173] L. A. Buatois, G. M. Narbonne, M. G. Mángano, N. B. Carmona, and P. Myrow, "Ediacaran matground ecology persisted into the earliest Cambrian," *Nat. Commun.*, vol. 5, no. 1, p. 3544, Mar. 2014, doi: 10.1038/ncomms4544.
- [174] O. Hammer, "Computer simulation of the evolution of foraging strategies: application to the ichnological record," *Palaeontol. Electron.*, 1998, doi: 10.26879/98005.
- [175] F. Papentin, "A Darwinian evolutionary system: III. Experiments on the evolution of feeding patterns," J. Theor. Biol., vol. 39, no. 2, pp. 431–445, May 1973, doi: 10.1016/0022-5193(73)90110-0.
- [176] F. Papentin and H. Röder, "Feeding patterns: the evolution of a problem and a a problem of evolution," *Neues Jahrb. Für Geol. Paläontol.*, no. 3, pp. 184–191, 1975.
- [177] R. E. Plotnick, "Ecological and L-system based simulations of trace fossils," Palaeogeogr. Palaeoclimatol. Palaeoecol., vol. 192, no. 1–4, pp. 45–58, Mar. 2003, doi: 10.1016/S0031-0182(02)00678-8.
- [178] L. M. Grabowski, W. R. Elsberry, C. Ofria, and R. T. Pennock, "On the Evolution of Motility and Intelligent Tactic Response," in *Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation*, New York, NY, USA, 2008, pp. 209–216. doi: 10.1145/1389095.1389129.
- [179] E. L. Charnov, "Optimal foraging, the marginal value theorem," *Theor. Popul. Biol.*, vol. 9, no. 2, pp. 129–136, Apr. 1976, doi: 10.1016/0040-5809(76)90040-X.

- [180] C. J. Perry, A. B. Barron, and K. Cheng, "Invertebrate learning and cognition: relating phenomena to neural substrate: Invertebrate learning and cognition," *Wiley Interdiscip. Rev. Cogn. Sci.*, vol. 4, no. 5, pp. 561–582, Sep. 2013, doi: 10.1002/wcs.1248.
- [181] R. Menzel, B. Brembs, and M. Giurfa, "1.26 Cognition in Invertebrates," in *Evolution of Nervous Systems*, J. H. Kaas, Ed. Oxford: Academic Press, 2007, pp. 403–442. doi: 10.1016/B0-12-370878-8/00183-X.
- [182] J.-M. Devaud, T. Papouin, J. Carcaud, J.-C. Sandoz, B. Grünewald, and M. Giurfa, "Neural substrate for higher-order learning in an insect: Mushroom bodies are necessary for configural discriminations," *Proc. Natl. Acad. Sci.*, vol. 112, no. 43, pp. E5854–E5862, Oct. 2015, doi: 10.1073/pnas.1508422112.
- S. Glautier, "Configural Cues in Associative Learning," in *Encyclopedia of the Sciences of Learning*, N. M. Seel, Ed. Boston, MA: Springer US, 2012, pp. 759–762. doi: 10.1007/978-1-4419-1428-6\_1777.
- [184] O. S. Soyer and R. A. Goldstein, "Evolution of response dynamics underlying bacterial chemotaxis," *BMC Evol. Biol.*, vol. 11, no. 1, p. 240, Dec. 2011, doi: 10.1186/1471-2148-11-240.
- [185] G. F. Miller and P. M. Todd, "Exploring adaptive agency I: Theory and methods for simulating the evolution of learning," 1991.
- [186] R. D. Hawkins and E. R. Kandel, "Steps toward a cell-biological alphabet for elementary forms of learning," *Neurobiol. Learn. Mem. Guilford Press N. Y.*, pp. 385–404, 1984.
- [187] H. Bode, S. Berking, C. N. David, A. Gierer, H. Schaller, and E. Trenkner, "Quantitative analysis of cell types during growth and morphogenesis in Hydra," *Wilhelm Roux Arch. Für Entwicklungsmechanik Org.*, vol. 171, no. 4, pp. 269–285, Dec. 1973, doi: 10.1007/BF00577725.
- [188] C. Skogh, A. Garm, D.-E. Nilsson, and P. Ekström, "Bilaterally symmetrical rhopalial nervous system of the box jellyfish Tripedalia cystophora," J. Morphol., vol. 267, no. 12, pp. 1391–1405, 2006, doi: 10.1002/jmor.10472.
- [189] K. Cheng, "Learning in Cnidaria: A systematic review," *Learn. Behav.*, vol. 49, no. 2, pp. 175–189, Jun. 2021, doi: 10.3758/s13420-020-00452-3.
- [190] J. I. Raji and C. J. Potter, "The number of neurons in Drosophila and mosquito brains," PLOS ONE, vol. 16, no. 5, p. e0250381, May 2021, doi: 10.1371/journal.pone.0250381.
- [191] R. Menzel and M. Giurfa, "Cognitive architecture of a mini-brain: the honeybee," *Trends Cogn. Sci.*, vol. 5, no. 2, pp. 62–71, Feb. 2001, doi: 10.1016/S1364-6613(00)01601-6.
- [192] O. J. Loukola, C. Solvi, L. Coscos, and L. Chittka, "Bumblebees show cognitive flexibility by improving on an observed complex behavior," *Science*, vol. 355, no. 6327, pp. 833–836, Feb. 2017, doi: 10.1126/science.aag2360.
- [193] S. R. Howard, A. Avarguès-Weber, J. E. Garcia, A. D. Greentree, and A. G. Dyer, "Numerical ordering of zero in honey bees," *Science*, vol. 360, no. 6393, pp. 1124–1126, Jun. 2018, doi: 10.1126/science.aar4975.
- [194] T. J. McCabe, "A Complexity Measure," IEEE Trans. Softw. Eng., vol. SE-2, no. 4, pp. 308–320, Dec. 1976, doi: 10.1109/TSE.1976.233837.
- [195] S. J. Shettleworth, *Cognition, evolution, and behavior*. Oxford University Press, 2010.
- [196] T. W. Fawcett and W. E. Frankenhuis, "Adaptive explanations for sensitive windows in development," *Front. Zool.*, vol. 12 Suppl 1, p. S3, 2015, doi: 10.1186/1742-9994-12-S1-S3.
- [197] G. S. Hornby, H. Lipson, and J. B. Pollack, "Generative representations for the automated design of modular physical robots," *IEEE Trans. Robot. Autom.*, vol. 19, no. 4, pp. 703–719, Aug. 2003, doi: 10.1109/TRA.2003.814502.
- [198] M. Schoenauer, P. Savéant, and V. Vidal, "Divide-and-Evolve: a Sequential Hybridization Strategy Using Evolutionary Algorithms," in Advances in Metaheuristics for Hard Optimization, P. Siarry and Z. Michalewicz, Eds. Berlin, Heidelberg: Springer, 2008, pp. 179–198. doi: 10.1007/978-3-540-72960-0\_9.
- [199] R. Albert, "Scale-free networks in cell biology," J. Cell Sci., vol. 118, no. Pt 21, pp. 4947–4957, Nov. 2005, doi: 10.1242/jcs.02714.
- [200] T. J. Gibson, "Cell regulation: determined to signal discrete cooperation," *Trends Biochem. Sci.*, vol. 34, no. 10, pp. 471–482, Oct. 2009, doi: 10.1016/j.tibs.2009.06.007.
- [201] Y. Timsit and S.-P. Grégoire, "Towards the Idea of Molecular Brains," Int. J. Mol. Sci., vol. 22, no. 21, Art. no. 21, Jan. 2021, doi: 10.3390/ijms222111868.
- [202] A. Kurakin, "Self-organization versus Watchmaker: stochastic dynamics of cellular organization," *Biol. Chem.*, vol. 386, no. 3, pp. 247–254, Mar. 2005, doi: 10.1515/BC.2005.030.
- [203] A. Kurakin, "Self-organization versus Watchmaker: ambiguity of molecular recognition and design charts of cellular circuitry," J. Mol. Recognit., vol. 20, no. 4, pp. 205–214, 2007, doi: 10.1002/jmr.839.
- [204] J. Collado-Vides, "A syntactic representation of units of genetic information—A syntax of units of genetic information," J. Theor. Biol., vol. 148, no. 3, pp. 401–429, Feb. 1991, doi: 10.1016/S0022-5193(05)80245-0.
- [205] J. Collado-Vides, "Grammatical model of the regulation of gene expression.," *Proc. Natl. Acad. Sci.*, vol. 89, no. 20, pp. 9405–9409, Oct. 1992, doi: 10.1073/pnas.89.20.9405.
- [206] D. Bray, "Protein molecules as computational elements in living cells," *Nature*, vol. 376, no. 6538, pp. 307–312, Jul. 1995, doi: 10.1038/376307a0.

- [207] W. Banzhaf, "Artificial Regulatory Networks and Genetic Programming," in *Genetic Programming Theory and Practice*, Springer, Boston, MA, 2003, pp. 43–61. doi: 10.1007/978-1-4419-8983-3\_4.
- [208] W. Banzhaf, "On Evolutionary Design, Embodiment, and Artificial Regulatory Networks," in Embodied Artificial Intelligence, Springer, Berlin, Heidelberg, 2004, pp. 284–292. doi: 10.1007/978-3-540-27833-7\_22.
- [209] W. W. Fischer, J. Hemp, and J. E. Johnson, "Evolution of Oxygenic Photosynthesis," *Annu. Rev. Earth Planet. Sci.*, vol. 44, no. 1, pp. 647–683, 2016, doi: 10.1146/annurev-earth-060313-054810.
- [210] J. T. O. Kirk, Light and Photosynthesis in Aquatic Ecosystems. Cambridge University Press, 2010.
- [211] N. G. Bednarska, J. Schymkowitz, F. Rousseau, and J. Van Eldere, "Protein aggregation in bacteria: the thin boundary between functionality and toxicity," *Microbiology*, vol. 159, no. 9, pp. 1795– 1806, 2013, doi: 10.1099/mic.0.069575-0.
- [212] E. J. Stewart, R. Madden, G. Paul, and F. Taddei, "Aging and Death in an Organism That Reproduces by Morphologically Symmetric Division," *PLOS Biol.*, vol. 3, no. 2, p. e45, Feb. 2005, doi: 10.1371/journal.pbio.0030045.
- [213] B. S. C. Leadbeater, *The choanoflagellates: evolution, biology, and ecology*. Cambridge, United Kingdom: Cambridge University Press, 2015.
- [214] T. Brunet, B. T. Larson, T. A. Linden, M. J. A. Vermeij, K. McDonald, and N. King, "Light-regulated collective contractility in a multicellular choanoflagellate," *Science*, vol. 366, no. 6463, pp. 326– 334, Oct. 2019, doi: 10.1126/science.aay2346.
- [215] T. Brunet and N. King, "The Origin of Animal Multicellularity and Cell Differentiation," *Dev. Cell*, vol. 43, no. 2, pp. 124–140, Oct. 2017, doi: 10.1016/j.devcel.2017.09.016.
- [216] D. M. Bryson and C. Ofria, "Understanding Evolutionary Potential in Virtual CPU Instruction Set Architectures," *PLoS ONE*, vol. 8, no. 12, p. e83242, Dec. 2013, doi: 10.1371/journal.pone.0083242.
- [217] C. Ofria and C. Adami, "Evolution of Genetic Organization in Digital Organisms," in *Evolution as Computation*, 2002, pp. 296–313.