

# **BIOBOTS: A NEW TYPE OF MULTICELLULARITY**

DICTYOSTELIUM AS A MODEL FOR COMPUTER DESIGNED ORGANISMS

**Simon Van Parys**

Student ID: 01710902

Supervisors: Prof. Dr. Ir Wim Van Criekinge, Dr. Ir. Michiel Stock,

Tutors: Prof. Dr. Ir. Francis wyffels, Ms. Mattijs Biondina

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of master in Bioscience Engineering.

Academic year: 2021 - 2022

De auteur en promotor geven de toelating deze scriptie voor consultatie beschikbaar te stellen en delen ervan te kopiëren voor persoonlijk gebruik. Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot de verplichting uitdrukkelijk de bron te vermelden bij het aanhalen van resultaten uit deze scriptie.

The author and promoter give the permission to use this thesis for consultation and to copy parts of it for personal use. Every other use is subject to the copyright laws, more specifically the source must be extensively specified when using results from this thesis.

Gent, June 10, 2022

The promotor,



Prof. Dr. Ir Wim Van Criekinge, Dr. Ir. Michiel Stock Simon Van Parys

The author,



## **DANKWOORD**

Deze thesis markeert het einde van mijn carrière als bio-ingenieur student. Ik blik heugelijk terug op een leerijke periode, die wel voorbij lijkt gevlogen. Na vijf jaar wandel ik de Universiteit Gent niet enkel buiten met een sterke basiskennis in verschillende bio-ingenieur domeinen, maar leerde ik ook de waarde kennen van andere vaardigheden. Het ontwikkelen van een openstaande, maar toch kritische blik lijkt belangrijker dan ooit. Ook andere vaardigheden zoals, analistisch denken , werken in een team en het plannen van taken heeft de UGent mij bijgebracht. Bij deze wil ik alle professoren, assistenten, technisch en administratief personeel van de Universiteit Gent bedanken om deze leerrijke, maar vooral ook leuke tijd mogelijk te maken.

Vooreerst zou ik graag enkele personen bedanken, zonder wie deze thesis er niet zou gekomen zijn. Vooreerst mijn promotores, Wim Van Crieckinge en Michiel Stock, alsook mijn tutors Francis wyffels en Matthijs Biondina. Ik wil jullie graag bedanken voor de verrijkende gesprekken en brain-storm sessies. Een thesis verloopt zelden zonder slag of stoot, maar hoe dan ook kon ik steeds bij jullie terecht voor antwoorden en raad.

Verder wil ik graag mijn familie, grootouders, broer, zussen, vrienden en kotgenoten bedanken voor alle steun de afgelopen jaren.

Ten slotte zijn er nog drie mensen die ik graag nog eens extra wil bedanken. Vooreerst mijn ouders, voor hun onvoorwaardelijke steun, zowel gedurende het afgelopen jaar als alle jaren ervoor. Zonder jullie zou ik nu niet staan waar ik nu sta. Als laatste wil ik graag mijn vriendin Asa bedanken om steeds een luisterend oor te bieden wanneer nodig, maar vooral ook voor alle leuke momenten die we samen hebben gedeeld.



# CONTENTS

<b>Dankwoord</b>	i
<b>Contents</b>	iv
<b>Nederlandse samenvatting</b>	v
<b>Summary</b>	vii
<b>1 Introduction</b>	1
<b>2 Living robots</b>	5
2.1 Definition . . . . .	5
2.2 Micro robotics . . . . .	5
2.3 Xenobots . . . . .	6
2.3.1 Xenobots 1.0 . . . . .	8
2.3.2 Xenobots 2.0 . . . . .	8
2.3.3 Xenobots 3.0 . . . . .	9
2.4 Applications of biobots . . . . .	10
2.5 General design pipeline . . . . .	12
2.5.1 Building Blocks . . . . .	13
2.5.2 <i>In silico</i> design of biobots . . . . .	17
2.5.3 Evolutionary Algorithms (EA) . . . . .	20
2.5.4 Genetic Algorithms (GA) . . . . .	20
2.5.5 <i>In vivo</i> realisation . . . . .	25
<b>3 Dictyostelium discoideum</b>	29
3.1 Relevance of Dictyostelium as a model organism . . . . .	29
3.2 Life cycle . . . . .	29
3.2.1 Aggregation . . . . .	30
3.2.2 Pseudoplasmodium formation . . . . .	31
3.2.3 Migratory slug . . . . .	31
3.2.4 Culmination . . . . .	32
3.3 Dictybots? . . . . .	33
3.3.1 Unicellular aggregation . . . . .	33
3.3.2 Salt and pepper sorting . . . . .	34
3.3.3 Migratory slug . . . . .	34
3.4 Experiment . . . . .	35
<b>4 Informational individual</b>	37
4.1 Complexity of biology . . . . .	37
4.2 A digression into information theory . . . . .	40
4.2.1 What is information . . . . .	40
4.2.2 Shannon's entropy . . . . .	40
4.3 Quantifying individuality . . . . .	44
4.3.1 Types of individuality . . . . .	47
4.4 Individuality of aggregating Dictyostelium cells . . . . .	48
4.4.1 Abstraction . . . . .	48
4.4.2 Forms of individuality . . . . .	51

**5 Conclusion** **57**

**Bibliografie** **59**

# SAMENVATTING

Van microrobots wordt verwacht dat ze in de toekomst een groot potentieel hebben voor verschillende biomedische en ecologische toepassingen. Echter, het ontwerpen en produceren van microrobots uit traditionele materialen zoals roestvrij staal, aluminium en plastics blijkt moeilijk haalbaar. Recent ontwikkelingen in microrobotica, soft robotica en biologie gaven de aanzet voor het gebruik van een nieuw soort materiaal, namelijk levende cellen. Kriegman et al. (2020) doopten deze eerste *in vivo* realisatie van dit soort robot om tot *Xenobot*, naar de kikker *Xenopus laevis*, het organisme dat de cellen leverde. Hoofdstuk 2 bespreekt de huidige *state of the art* in het snel evoluerende veld van xenobots. Het bespreekt de voordelen van soft robotica ten opzichte van de traditionele robotica en de uitdagingen die gepaard gaan met het schalen van robots naar de microschaal. Verder worden ook voorziene toepassingen van xenobots besproken. Ten slotte, beschrijft Hoofdstuk 2 ook de *pipeline* voor het ontwerpen en produceren van xenobots, zoals voorgesteld in (Kriegman et al., 2020).

Reaggregatie van embryonale cellen is niet enkel voorbehouden voor *Xenopus laevis*' cellen. Hoofdstuk 3 bespreekt het modelorganisme *Dictyostelium discoideum* als celldonor organisme en model systeem voor het maken van biobots. *Dictyostelium discoideum* is een vrijlevende amoeba in de bodem, die een opmerkelijke levenscyclus doorloopt. *Dictyostelium discoideum*, ook wel *slijmzwam* genoemd is een goed bestudeerd modelorganisme, gebruikt om verschillende biologische processen zoals chemotaxis, celsortering en signaltransductie te bestuderen. Wanneer de amoebes hongerstress ondervinden, zijn ze in staat te aggereren en een multicellulair pseudoplasmodium of *slak* te vormen. Deze slak is in staat om gericht te bewegen onder invloed van chemo-, photo en thermotaxis. Aangezien *Dictyostelium* cellen van nature over een aantal exploiteerbare eigenschappen beschikken, worden ze beschouwd als een interessant model systeem om biobots van te bouwen.

Door levende cellen te gebruiken als bouwmaterial voor microrobots, maken wetenschappers gebruik van het aanpassingsvermogen en plasticiteit in de natuur. Echter, brengt dit ook nadelen met zich mee. Door levende materialen te gebruiken, wordt een extra laag aan complexiteit aangebracht. Hoofdstuk 4 stelt een wiskundig kader voor gebaseerd op informatie theorie om de individualiteit van een systeem te quantificeren. De aanname van individualiteit is wijdverspreid binnen biologie, echter zijn er weinig sluitende methoden om het toe te kennen. De multicellulaire *slak*, die gevormd wordt na aggregatie van *Dictyostelium* cellen, lijkt op een individueel organisme. Echter, afhankelijk van de criteria, wordt het niet altijd als individu geclasseerd. Aangezien biobots een nieuwe soort van organismen zijn, kan er ook verwarring ontstaan omtrent de individualiteit van biobots. Door te quantificeren hoe informatie wordt doorgegeven tussen een systeem en zijn omgeving door de tijd, kan de individualiteit van het systeem bepaald worden.

**Sleutelwoorden:** Multicellulariteit, biobot, xenobot, *Dictyostelium discoideum*, soft robotics, micro robotics, informatie theorie, individualiteit.



## **SUMMARY**

Microrobots are projected to have immense potential for various biomedical and ecological applications. However, designing and producing complex micro robots from traditional materials, such as stainless steel, aluminium and plastics, has shown to be challenging. Recent advances in microrobotics, soft robotics, and biology have unlocked a new type of building material: the living cell. Kriegman et al. (2020) baptised the first *in vivo* realisation of this biological robot *Xenobots*, after the frog *Xenopus laevis*, the organism that provided the cells. Building robots out of living materials has some major advantages compared to conventional materials. Chapter 2 of this thesis aims to provide an overview of the current state of the art in the evolving field of biobots. It discusses the potential advantages of soft robotics compared to traditional rigid robotics and the challenges associated with scaling down robotics to the microscale. Further, it discusses some potential real-world applications of biobots. Lastly, it describes the pipeline for designing and production of xenobots and, in extension, biobots, as proposed by Kriegman et al. (2020).

Re-aggregation of embryonic cells is not solely reserved for *Xenopus laevis*' cells. Chapter 3 of this thesis discusses the model organism *Dictyostelium discoideum* as a potential cell donor organism and model system. *Dictyostelium discoideum* is a free-living soil amoeba that goes through a remarkable life cycle. *D. discoideum*, commonly referred to as a slime mould, is a well known, robust model organism used to study biological processes such as chemotaxis, cell sorting and signal transduction. When the amoebae experience starvation stress, it aggregates and forms a multicellular pseudoplasmodium or *slug*, capable of chemo-, photo- and thermotaxis. Since *Dictyostelium* cells natively possess some exploitable features, they are considered a potential material to build biobots from.

Using cells as building materials for micro robots, scientists exploit the plasticity and adaptability found in nature. However, using living materials also introduces an additional level of complexity. Chapter 4 proposes a mathematical framework based on information theory to quantify the individuality of a system. The assumption of individuality is widespread in biology. However, there is little agreement on what an individual is. The migratory *Dictyostelium* might act as an individual organism, but depending on the criteria used, it might not be classified as one. Since biobots are a new species of organisms, there may also be confusion regarding the individuality of biobots. By quantifying how information is passed from a system to itself and its environment, the individuality of the system can be assessed.

**Keywords:** Multicellularity, Biobots, Xenobots, *Dictyostelium discoideum*, Soft robotics, Micro robotics, Information theory, individuality.



# 1. INTRODUCTION

In all shapes and sizes, robots have become an integral part of today's world. From applications ranging from medicine to heavy industrial processes, the economy has become heavily reliant on robots. Because of their wide range of applications, researchers are constantly developing new types of robots. Some might repeatedly perform the same simple task, while others are highly sophisticated and complex. The field of robotics is a quickly evolving one. Designing a tool that can carry out repetitive or heavy tasks is an idea that can be dated back to ancient times. Despite some debate about the exact date in the literature, it was not until 1950 that the first industrial robot was developed (Gasparetto et al., 2019). The *Unimate* developed by Joseph Engelberger and George Devol in the '50s stands in pale contrast to highly sophisticated robots today. Think of the *Perseverance* rover currently on Mars, or modern bio-prosthetics, which can sense and act on neuronal signals. Despite robotics being a fast-evolving field, some fundamental laws devised by science fiction writer Isaac Asimov in 1942 are still applicable (Asimov, 1942). These include:

- **First law:** A robot may not injure a human being or, through inaction, allow a human being to come to harm.
- **Second law:** A robot must obey the orders given to it by human beings except where such orders would conflict with the First Law.
- **Third law:** A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Over the years, robots have made many appearances in popular culture, with famous examples including Arnold Schwarzenegger, as a killer robot in *The Terminator*, the *Sentinels* in *The Matrix* and the more friendly R2-D2 and C-3PO from the *Star wars* saga. Depictions like these have created a public image of the typical appearance of robots as human-like machines built from shiny rigid materials such as stainless steel, aluminium or plastics. However, in the last decades, a lot of attention has been directed to the development of soft alternative materials (Lee et al., 2017). This branch of robotics is known as *soft robotics*. Conventional, rigid robots thrive in highly predictable, controlled environments where they perform a specific task. This is the case in, for example, robots in industrial assembly lines. Due to the stiffness of their materials, rigid robots require complex protection and stability control algorithms (Lee et al., 2017). As a result, placing them in unstructured conditions will likely damage their structure and eventually lead to loss of function. Due to their *softness*, soft robots are less prone to damage, potentially allowing them to bridge the gap between robot and human or biological environments, making them promising candidates for medical and ecological applications. Additionally, soft materials are less likely to cause damage to their environment in the case of a collision. Today, current soft materials include different kinds of elastic polymers and gels (Bilodeau and Kramer, 2017).

---

Another emerging branch of robotics is the branch of micro-robotics. Over the past decades, scientists have strived to develop motile sub-millimetre micro robots (Soto et al., 2022). These microrobots show several promising applications ranging from targeted drug delivery to *in vivo* biosensing, non-invasive medicine and environmental remediation (Li et al., 2017). However, miniaturising traditional robots has proven not to be a viable approach (Palagi and Fischer, 2018). Conventional rigid robots can be seen as an assembly of discrete components, where each component performs its prescribed role. In order to make these parts operate, traditional robots rely on operator inputs or artificial intelligence software, resulting in a clear distinction between hardware and software (Bongard and Levin, 2021). Due to size limitations, implementing these electronic control circuitry has shown not to be feasible in microrobots (Soto et al., 2022). However, in nature, the distinction between hardware and software is less prominent (Bongard and Levin, 2021). Biological systems have the ability to gather information and act on it, directly encoded into their materials. Unicellular microorganisms can actively swim around while sensing their environment and act correspondingly, yet they do not possess over any type of central processor, managing all different ongoing processes (Palagi and Fischer, 2018). Responses are generated through and controlled by a complex network of biochemical and biophysical loops.

Another hurdle when miniaturising conventional robots to the micro-scale stems from fluid dynamics. Moving in liquid at the micro-scale is fundamentally different from the macro-scale (Palagi and Fischer, 2018). In his work *Life at Low Reynold Numbers*, explaining the principles of aquatic locomotion, Purcell proved that conventional swimming techniques at the macro-scale can not be transposed to the micro-scale due to the low inertial forces compared to the viscosity (Purcell, 1977). Under these conditions, microorganisms have developed ingenious alternative motion mechanisms, such as cilia and helical propellers, to achieve locomotion.

In the scope of developing complex microrobots, Kriegman et al. (2020) came up with the idea to build a microrobot entirely from animal cells. This idea led to the development of biobots. Kriegman et al. (2020) succeeded in making these biobots and called their newly formed aggregates *xenobots*, named after *Xenopus laevis*, the organisms from which the cells were derived. The Kriegman et al. (2020) hypothesised that, since xenobots are entirely biological, they would inherit the plasticity and adaptive potential associated with living cells. Chapter 2 gives an overview of the current state of the art in the field of biobots. Further, it reviews the pipeline for designing reconfigurable organisms, proposed in Kriegman et al. (2020).

The ability to re-aggregate after separation is not solely reserved for *Xenopus laevis* cells. Chapter 3 discusses the social amoeba *Dictyostelium discoideum* and its possibility to derive biobots from it. *Dictyostelium sp.* is a well-studied model organism which goes through a remarkable life cycle. When placed under starvation stress, unicellular *Dictyostelium sp.* amoebae can aggregate to form a multicellular migratory pseudoplasmodium. During the formation of the pseudoplasmodium or *slug* cells self-organise into predefined patterns and form highly conserved layers of different cell types. Since *Dictyostelium sp.* cells natively possess the properties of self-assembly and organisation, they stand out as interesting starting points for the production of biobots.

Introducing biological materials for the manufacturing of microrobots also comes with disadvantages. Whereas conventional robots are highly predictable, life is not (Bongard and Levin, 2021). Biobots are considered a completely new form of biological life. However, biological concepts such as *fitness*, *repro-*

ductive success and *individuality* are hardly applicable. Chapter 4 of this thesis discusses the concepts of individuality. The notion of a *biological individual* is crucial when studying its biology. Despite the universal assumption of individuality, there is a lack of a fundamental method to classify it (Krakauer et al., 2020). As of today, different criteria for identifying the individuality of a system exist (Gilbert et al., 2012; Buss, 1988). However, there is little agreement on what individuals are. Chapter 4 proposes a mathematical framework to quantify different types of individuality by studying the flows of information between a system and its environment over time. Krakauer et al. (2020) proposed three different types of individuality: an organismal, colonial and environmentally driven form. These three metrics of individuality are calculated for an illustrative example for *Dictyostelium* cells during aggregation.

---

## **2. LIVING ROBOTS**

### **2.1 Definition**

The first *in vivo* realisation of xenobots by Kriegman et al. (2020) gave rise to an entirely new form of life. Where xenobots derive their name from the frog *Xenopus laevis*, biobots are considered the umbrella term for synthetic living machines. The term biobot includes all kinds of robots that are partially or wholly manufactured from biological materials. Biobots meet most criteria of being 'a living entity' (Levin et al., 2020). However, there is debate over whether biobots should be considered animals, organisms, robots or a combination of the above. Biobots are actively blurring the lines between biology and robotics.

This thesis adopts the definition given by Bongard and Levin (2021), who define a robot as a machine capable of physical actions which have direct impacts on the world and which can sense the repercussions of those actions, and is partly or completely independent of human action and intent. In the scope of this thesis, biobots are considered as microscale robots that are fully comprised of biological material such as living cells.

### **2.2 Micro robotics**

As of today, challenges around developing and manufacturing robots on the micro-scale have not been resolved (Yu et al., 2018). Implementing complex control circuits into microscale robots still poses a major problem, due to size restrictions of electronic compounds (Servant et al., 2015). In addition, the manufacturing of micro robots has shown to be difficult to scale. Although progress has been made through advances in 3D-printing techniques, the high-throughput production of micro robots remains a major challenge, holding back real-world applications of micro robotics (Soreni-Harari et al., 2020; Alcântara et al., 2019). Recent advances in stem cell biology and tissue engineering have pointed to an alternative, biological approach. Kriegman et al. (2020) developed a method to use living cells as a new type of material for the scalable production of micro robots. Kriegman et al. (2020) successfully built a micro robots out of *Xenopus laevis*' cells, which was called *xenobot*.

Xenobots are the first micro-robots that are completely built from biological tissue, thereby inheriting its adaptive potential inherent to life. Almost (see Chapter 3) all higher multicellular organisms go through different strictly regulated stages of development. At this stage, cells show a high degree of self-organisation and emergent behaviour. Living cells already possess the ability to sense their environment, act on it and communicate it through various signalling and computational circuits, sidestepping the previously mentioned conceptual and manufacturing challenges.

---

In a way, biobots resemble organoids, which are three-dimensional *in vitro* synthetic analogues of complex tissues and organs, such as components of the eye, gut, liver, intestine, mammary glands and pancreas. Work on organoids has shown to be valuable in stem cell research and has allowed researchers to drive progenitor cells to a wide variety of different cell types (Rossi et al., 2018). Recent advantages led to the development of full body-on-chip systems (Zhang et al., 2018). However, biobots and organoids differ in some aspects as well. Firstly, organoids are built in order to study existing *in vivo* organs and their corresponding metabolism. In addition, organoids are often produced by bioprinting or by using scaffolds. As a result of this, the potential of living cells to self-organise is not fully utilised. Biobots aim to be a new type of life, built from existing building blocks rather than replicating known morphologies. Additionally, the ability of cells to physically move through and interact with their environment is not exploited.

Examples of biorobots are already present in literature. However, they often come in the form of bio-hybrids such as the phototactic stingray (Park et al., 2016), skeletal muscle propelled micro-biohybrids (Cvetkovic et al., 2014) or the macro scale sea turtle hybrid (Webster et al., 2017). However, hybrids still require synthetic materials to some extent. Hereby, the traditional limitations of traditional robotics are still present. They cannot replicate, self-heal or show inherent emergent and self-organisational features.

Instead, Kriegman et al. (2020) opted for a fully biological machine, allowing robots to exploit the remarkable plasticity and flexibility found in nature.

## 2.3 Xenobots

The first *in vivo* realisation of the previously described biobots was developed by the group around Kriegman et al. (2020). They baptised their newly-achieved synthetic organisms *xenobots*, after *Xenopus laevis*, the organism from which the cells were isolated. Not to be confused with xenotransplantation, where recently a pig heart was successfully transplanted into a human body (Reardon, 2022), where the prefix "xeno-" is derived from the old Greek word  $\xi\epsilon\nu\sigma$ , meaning stranger. *Xenopus sp.* and in extension *Xenopus laevis* or the African clawed frog is a popular and well-studied model organism in a variety of biological fields. Their popularity stems from the fact that they are tolerant to starvation, disease and other stress factors, allowing them to be easily kept in captivity. Female *Xenopus* can produce a large number of offspring all year round. Moreover, *Xenopus* eggs and embryos are unusually large, simplifying bio-imaging in research (Wallingford et al., 2010).

The original xenobots are less than a millimeter in diameter and consist of only two different *Xenopus* cell types, Figure 2.1 shows an example of a *in silico* design and its *in vivo* realisation. *Xenopus* cells are harvested in the blastula stage and still possess most of their pluripotency: having the potential to differentiate in any of the three germ layers. After isolation, cells were dissociated and pooled together in desired amounts. After spontaneous aggregation, contractile tissue made from cardiac progenitor cells is layered on top of the spherical aggregates. Afterwards, the aggregates are shaped into the desired morphology through microsurgery. The formed aggregates were motile and could survive for up to ten days. They were able to move by actively pushing off, by actively expanding and contracting their active

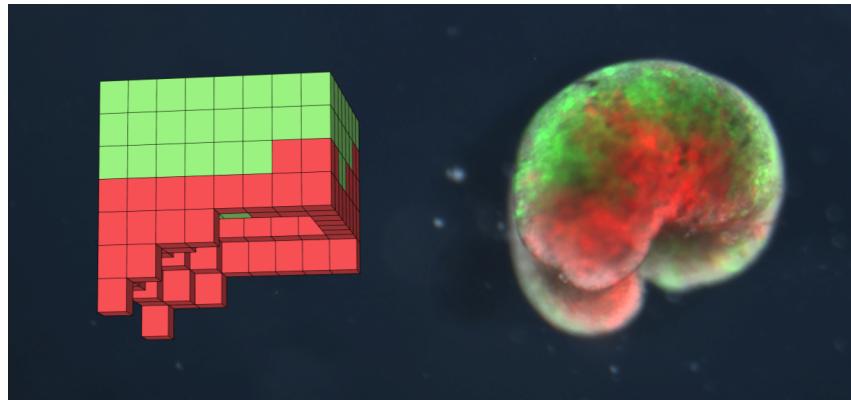


Figure 2.1: The left-hand side shows a computer designed configuration of a xenobot. Passive ectodermic tissue is depicted in green, while active cardiac tissue is shown in red. On the right hand side, the *in vivo* morphology of the xenobot that embodies the *in silico* design is depicted Kriegman et al. (2020)

tissue. The *in vivo* realisation of desired biobots will be discussed in greater detail in Section 2.5.5. Note that the cells of the obtained xenobots have not undergone any type of genetic engineering and are genetically identical to wild-type *Xenopus laevis* cells.

Xenobots are the first *in vivo* realisation of computer-designed organisms. Computer designing of organisms is a new emergent field stemming from synthetic developmental biology and machine learning. The widely accepted Darwinian approach toward evolution states that all organisms evolve, and this evolution happens through mutation, recombination or other sources of genetic variation. The advantageous variations will persist in populations due to the mechanism of survival of the fittest. Evolution is the process that has given rise to biodiversity at all levels of biological organisation. However, xenobots are synthetic constructs, unable to replicate. Therefore they are not prone to traditional evolution or natural selection. However, xenobots are fully evolved in a virtual world through modern machine learning algorithms in a matter of hours or days rather than in the biosphere, where the same processes would take up a period a couple of orders higher. Section 2.5.2 discusses various methods and algorithms applied in order to perform this virtual evolution.

Xenobots and, by extension soft robots show several advantages compared to traditional rigid robots. Recently a lot of work is directed on the development of bioinspired self-healing materials. Self-healing and damage prevention are essential features in nature, moreover all multicellular organisms are intrinsically able to withstand some amount of damage and heal (Cremaldi and Bhushan, 2018). Since xenobots consist of living mammalian cells and can be considered as a living organism (Bongard and Levin, 2021), one could hypothesise that the property of self-healing will be conserved. Indeed, Blackiston et al. (2021) showed the ability to fully self-repair after suffering severe mechanical damage, as shown in Figure 2.2. Damage resistance, or the ability to withstand stressors without getting damaged, and damage resilience, or the ability to continue functioning after being damaged are key features for machines operating under unstructured and unpredictable environments, such as the human body. As of today, three generations of xenobots have been developed.



Figure 2.2: Five-day-old xenobots are mechanically cut by a single incision of about half their diameter. Images show a xenobot, before and after the incision, as well as the repair process, 5 minutes, 10 minutes, 15 minutes and 48 hours after the incision. This individual was able to fully close resolve the wound and regain its spherical shape (Blackiston et al., 2021).

### 2.3.1 Xenobots 1.0

The first generation xenobots, referred to as xenobots 1.0 are the original xenobots proposed by Kriegman et al. (2020). These xenobots consist of two different *Xenopus* cell types: contractile cardiac cells and passive epidermal cells. Out of all possible designs, an evolutionary algorithm was tasked to find the best performing configurations. Through simulations, the performance of a design was assessed by various means, with the main focus on unidirectional locomotion and energy consumption. The *in silico* simulator did not model any locomotion provided by cilia, which are extracellular organelles found on eukaryotic cells. Both motile and non-motile cilia exist. In order to ensure displacement is only due to contractile tissue pushing against the surface of the Petri dish, cilia were suppressed *in vivo* by embryonic microinjection of small interfering RNA (siRNA), coding for the Notch intracellular domain (Deblangre et al., 1999). Despite lacking any form of nervous system, individual cells were shown to be able to cooperate in groups and self-organize into aggregates. Although initially, the activation of cardiomyocytes was modelled as random noise, *in vivo* cells would exhibit spontaneous coherent, phase-matched contractions (Kriegman et al., 2020). They were able to not only self-organise into their artificial morphology, but also self-repair in case of mechanical damage. As previously stated, self-repair properties are nearly impossible to achieve in conventional microrobotics.

### 2.3.2 Xenobots 2.0

Whereas the first generation of xenobots are built from a top-down approach, where first a well performing configuration is simulated and built *in vivo*, Blackiston et al. (2021) developed an alternative bottom-up strategy. These xenobots 2.0 consist entirely out of wild-type *Xenopus laevis* embryonic ectoderm cells. Aggregates become multiciliated, by expressing cilia on the outer cells of the aggregate. The repurposed cilia, which are responsible for distribution of mucus and flow pathogens and other material off the skin, allowed to generate a kind of swimming motion. Loose cells, removed from their normal developmental environment were harvested in the embryonic stage from the animal cap. Dis-

sociated individuals cells were, later brought together in small wells. Under the right conditions, cells began to aggregate and form spherical aggregates of a couple of thousand cells (Blackiston et al., 2021). Although completely consisting of skin tissue, xenobots were self-motile due to cilia present on the surface cells. Aggregates could survive for up to ten days without any external food source. An elaborate self-repair assessment showed that xenobots were able to resolve wounds, and close the injury site after an incisions of about half the diameter. Figure 2.2 depicts the process of self-repair in greater detail.

### 2.3.3 Xenobots 3.0

The last developments in the field of xenobots is the development of the third generation of xenobots (Kriegman et al., 2021). Growth can come in diverse forms: budding, fission, sexual reproduction, viral propagation and so forth. However, up to now, xenobots, although being fully biological in nature, were not able to grow or reproduce. Kriegman et al. (2021) developed a strategy to enable xenobots to reproduce. Similar to the second generation of biobots, cells were harvested from the animal cap of *Xenopus laevis* embryos. Just like in the second generation of xenobots, individuals cells clumped together and formed aggregates, with ciliated cells on their surface. After the formation of the tight aggregates, the xenobots were placed in Petri dishes covered in dissociated *Xenopus laevis* embryonic cells. The combined helical movements pushed together these cells and after some incubation time, the new piles of cells began to aggregate and form a second generation of xenobots, hereby enabling xenobots to produce a type of offspring by means of kinematic self-replication. The process of kinematic self-replication is further illustrated in Figure 2.3.

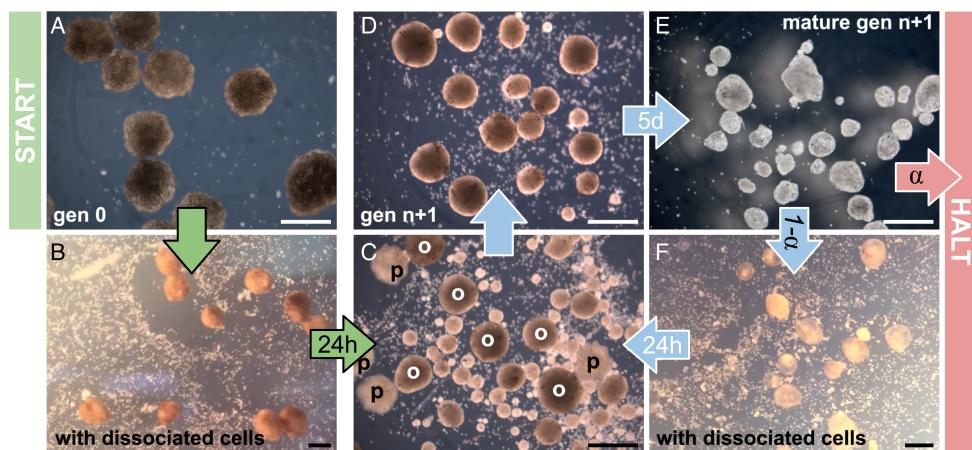


Figure 2.3: Kinematic self-replication in embryonic *Xenopus laevis* cell aggregates. (A) Stem cells, harvested from early stage *Xenopus laevis* embryos are harvested and put in solution. After some incubation time, spherical aggregates of around 3,000 cells are formed. Motility in aquatic environments is acquired by multiciliated cells on the aggregates surface. (B) Fully formed, motile aggregates are placed in a Petri dish dotted with dissociated embryonic stem cells. The collective motion of the aggregates pushes cells into piles. If piles are sufficiently large, new ciliated aggregates are formed (C). A new generation of offspring was formed by the initial generation. This process can be repeated as long as condition for kinematic self-replication are met. Conditions include a suitable temperature, sufficient dissociated cell concentration, amount of mature aggregates, viscosity, geometry of the dish and possible contamination (Kriegman et al., 2021).

---

## 2.4 Applications of biobots

The potential applications of biobots are similar to the ones attributed to conventional microrobots. However, the use of biobots show some remarkable advantages compared to their rigid counterparts. Due to their biological nature, they are fully biodegradable. Additionally they can be made from biocompatible materials and can withstand and repair damages to some extend. These features makes them an attractive tool in a variety of ecological and medical applications.

### Scientific tool

Nowadays, biobots in the form of xenobots are being primarily used as model organisms. Xenobots provide a completely new platform for studying the relationship between an organism's geno- and phenotype. Remarkable is that xenobots are genotypically 100% identical to *Xenopus laevis* organisms, yet phenotypically both look nothing alike. These new organisms allow scientists to study areas of morphospace reached by these new organisms that can not be studied by normal embryogenesis (Ebrahimkhani and Levin, 2021). How an organisms fully develops from a fertilised egg, to an embryo to a highly complex organisms is yet not fully understood. Hence, biobots could serve as a unique model system for studying multicellularity and its evolutionary encoding (Manicka and Levin, 2019). Advances in developmental embryology, synthetic biology and understanding how cells cooperate might revolutionise biomedicine and might unlock the full potential of regenerative medicine (Levin, 2020).

### Microplastics

In the future biobots are expected to be useful for real applications and perform several real-world tasks. It is estimated that something in between 93 and 236 thousand metric tons of microplastics are present in oceans globally (Sebille et al., 2015). These microplastics pose adverse effects on marine life and are transferred along long the food web and are shown to be harmful to human health (Sharma and Chatterjee, 2017). Autonomous motile photocatalytic microrobots are already being proposed as a possible solution for capturing and even degrading microplastics (Beladi-Mousavi et al., 2021). However, traditional robots, made from conventional materials still have some major shortcomings. As previously mentioned, the production of microrobots built from polymers or metals is not a process that is easily scaled up. Another disadvantage of these conventional microrobots is their non-biodegradable character. Residues of metals and plastics from the robots might end up in the food chain and pose potential marine and human health hazards themselves. Since biobots are entirely made from living animal cells, they are per definition biodegradable. Their limited lifespan and inability to reproduce should also limit potential threats towards biodiversity and ecological homeostasis. However, lifespan, reproductive and mutagenic potential, and possible interactions with other organisms should be thoroughly studied when releasing a completely new organisms in ecosystems (Levin et al., 2020). Xenobots have already been shown the exhibit objective manipulation and collection *in silico* as well as *in vivo* Kriegman et al. (2020, 2021). Piling up microplastics, enables the efficient collection of larger aggregates of plastics. In later stages biobots are also projected to enzymatically degrade plastics, hereby adding to the solution the

growing microplastic problem. Adding biobots in ecological systems does not increase the amount of pollution, whereas their rigid counterparts would.

### **Intelligent drug delivery**

Biobots could serve as a new method of targeted drug delivery. As of today, this mainly is done by metallic nanoparticles or even microrobots (Patra et al., 2013). These nanoparticles are often made from scarce and expensive metals and can be potentially toxic to the organisms in which they operate. Motile biobots, responsive to specific biomarkers, could aid in the targeted, fast delivery of drugs such as antibiotics and anti-tumor drugs. In the scope of the problem with antibiotic resistance, targeted drug delivery could be a major tool. Most antibiotics have broad-spectrum working range, hereby a selective pressure is put on non targeted bacteria, inducing antibiotic resistance (Sharma et al., 2012). Targeted delivery increases the specificity of the antibiotic, hereby this selective pressure is not applied and resistance is less likely to emerge. An additional benefit of targeted delivery of antibiotics is the reduction of disturbance of the gut microbiome, limiting the short-term and even long-term adverse effects of antibiotic treatments on human health (Lange et al., 2016). In cancer therapy, targeted drug delivery can be used to increase the efficacy of cancer treatment, resulting in better treatment and less side-symptoms, due to lower dosages and inhibition of drug delivery to non-cancer cells (Bahrami et al., 2017).

### **Biomedical applications**

The immune system is responsible for safeguarding the host from infections and other malignancies. When a foreign subject is introduced to an organism, it elicits a possibly unwanted response from the immune system entirely. However, biobots, completely constructed from the patient's own cells would completely avoid these interactions with the immune system, since they would be recognised as body's own cells. Their biological nature allows them to be biocompatible with possible host organisms, which can be a problem in both nanoparticle technology and conventional microrobotics.

Inside of the human body, they could perform a variety of different tasks. Biobots could remove arterial plaques, preventing atherosclerosis, which can result in heart-attacks. They could be used as a diagnostic tool for early cancer detection, targeted drug delivery or local disease treatment or control.

### **Detoxification**

Advances in synthetic biology allow for precise genetic manipulation and insertions of complex genetic circuitry. Engineered biobots could be used as biosensors for sensing and reporting toxic compounds at environments, which are physically inaccessible to other machines. By means of chemotaxis they could track toxic compounds and if equipped with the right genetic inserts degrade them *in situ*. Due to their biological nature, biobots are allowed to be introduced in both aquatic or terrestrial ecological systems, without causing any additional pollution. Although studies on the impact of introducing new kind of species to an ecosystem should still be conducted (Levin et al., 2020).

## 2.5 General design pipeline

The original paper by Kriegman et al. (2020) proposed a general scalable pipeline for the designing of reconfigurable organisms. Figure 2.4 gives an overview of this process. First of all, one needs to define a clear objective. A popular goal in evolutionary robotics is unidirectional locomotion. Next, all the possible building blocks to achieve this goal must be listed, along with their physical properties. The next step consists of an *in silico* optimisation. An evolutionary algorithm (EA) discovers possible designs combining the possible configurations, using the predefined building blocks. In the next step, the behaviour of a subset of possible designs is simulated and consequently scored based on their performance. The configurations of the best performing designs are retained and allowed to *breed offspring*, which are simulated and scored in their turn. This process repeats itself until termination criteria are met. The best-performing designs are built *in vivo* to validate the predicted behaviour of the designs. The *in vivo* realisation of the original xenobots is performed by microsurgery, making the production of large numbers of biobots infeasible. Section 2.5.2 discusses alternative approaches for the biobot production, potentially allowing for higher throughput production.

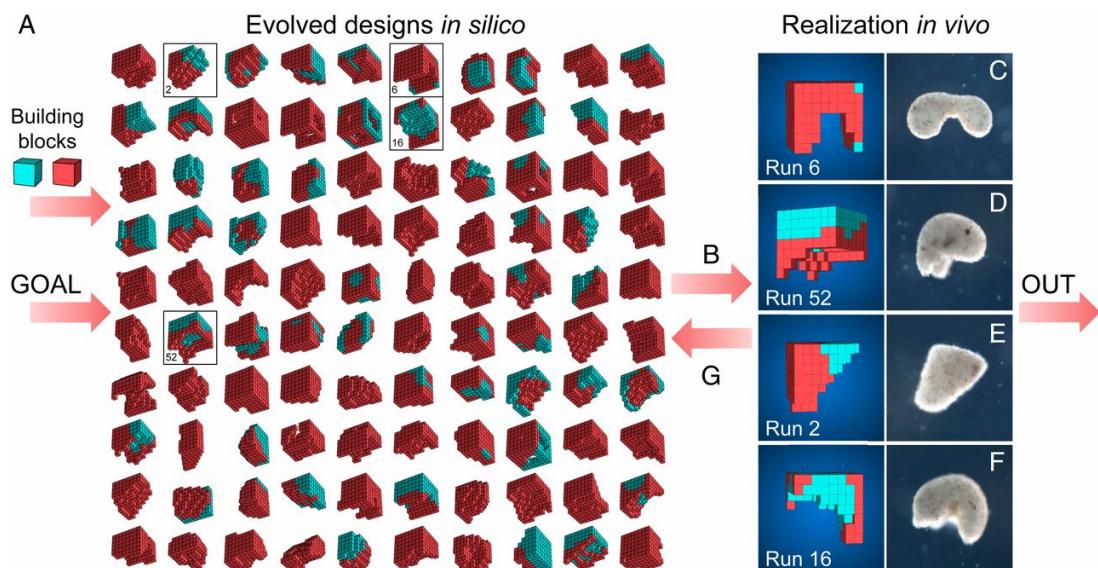


Figure 2.4: The pipeline for designing biobots consists of several distinct steps. It takes two inputs and outputs a set of designs. (A) The first two steps consist of defining a clear behavioural goal and the building blocks at hand. This research defined unidirectional planar locomotion as the desired goal and provided types of building blocks: contractile tissue (red) and passive structural tissue (green). (B) Given the goal and available building blocks, an evolutionary algorithm was tasked to return a set of the best performing configurations. Section 2.5.2 discusses the working mechanisms of this evolutionary algorithm in further detail. Once the algorithm returned a set of solutions, the last step consists of the *in vivo* realisation of the configuration out of living *Xenopus laevis* cardiomyocytes and epidermal progenitor cells to validate the results. (C-F) Four best solutions and their *in vivo* counterparts are depicted (Kriegman et al., 2020).

### 2.5.1 Building Blocks

The proposed pipeline needs two inputs: a goal and the available building blocks. Once a desired behaviour of the biobot is formulated, the next step consists of selecting the right building blocks and defining their physical properties. As previously discussed, nature inherently posses over an incredible amount of diversity. This diversity can be observed on all hierarchical levels of biological organisation. The human body already consists of around 200 different cell types (Bernstein et al., 2010). The toolbox consists of a near-infinite number of different cell types available to scientists when considering all the different species.

Section 2.4 discusses contemporary applications together with some of the anticipated use cases. Kriegerman et al. (2020) optimised organisms to achieve unidirectional locomotion on a flat surface. Unidirectional locomotion is a relatively simple task. Therefore, only a limited amount of different cell types should be required to achieve this goal. However, as tasks get more complex, an increase in the number of distinct building blocks will be required. An obvious next step would be the implementation of chemotaxis, which requires the presence of cells capable of sensing gradients of different kinds of chemical compounds.

#### Cell adhesion

In order to form a functioning multicellular organism, cells must cohere into a structure that can withstand external forces trying to pull them apart (Alberts et al., 2015). Organisms can maintain this functional state of multicellularity by linking cells to one another. Cells can either directly bond with other cells or bind to the *extracellular matrix* (ECM), a complex network of (glyco)proteins, proteoglycans and polysaccharides, excreted by the cells. However, cell-cell and cell-matrix do not solely serve a structural function. It allows cells to communicate and respond to external factors by controlling the orientation and behaviour of the cell's cytoskeleton. Cell junctions and matrix interactions are critical mechanisms for multicellular structures' organisation, function, dynamics, functionality and communication of (Alberts et al., 2015). Due to the diversity in interactions with specific and specialised functions, organisms can create highly complex and diverse classes of tissues. Cell-cell and cell-matrix interactions allow redistributing of mechanical stress. Dysfunctions in adhesive interactions and the communication pathway associated with them can cause or contribute to various health conditions, including a variety of neuromuscular and skeletal disorders and cancer (Lodish et al., 2016).

Cells in tissue can directly adhere to one another. These interactions are called cell-cell interactions. Adhesion is mediated by specialised membrane proteins called cell-adhesion molecules (CAMs) which often cluster in specialised cell junctions (Lodish et al., 2016). Most CAMs belong to one of the four major families: cadherins, immunoglobulin (Ig) superfamily, integrins and selectins, as shown in Fig. 2.5. The adhesion mechanisms in each of the four CAMs can link cells of the same cell type (*homotypic* adhesion) or cells of a different cell type (*heterotypic* adhesion). CAMs that bind to the same type of CAM on an adjacent cell is called *homophilic*, while CAMs that bind to other types of CAMs are called *heterophilic*. CAMs can be broadly distributed along the plasma membrane or come in narrow clusters in a small region on the plasma membrane. Generally, CAMs consist of two important domains.

---

(1) The cytosolic domain allows for binding to the cytoskeleton by recruiting adapter proteins. Adapter proteins directly link CAMs with structural elements of the cytoskeleton, such as actin filaments or intermediate filaments. Further, they play a role in the cell signalling pathway, which controls cellular behaviour, including gene expression of CAMs themselves (Lodish et al., 2016). (2) The other domain is the extracellular domain. The extracellular domain allows binding to CAMs outside the cell's plasma membrane. CAMs can bind to CAMs from an adjacent cell, forming *trans* interactions or CAMs from the same cell, forming *cis* or *lateral* interactions. CAMs can form very tight bonds in a Velcro-like manner, combining many weak interactions. These bonds are often formed in locations where many CAMs are concentrated, like in cell-cell junctions. Some CAMS, like integrins, cadherins and selectins, are calcium dependent, meaning they need  $Ca^{2+}$  to enable cell adhesion.

In a typical epithelial tissue, such as the lining of the small intestine of a vertebrate, three major classes of stable cell-cell junctions can be distinguished, as shown in Fig 2.6 .

Anchoring and tight junctions have a mainly structural role, holding cell together to form functional tissue. Tight junctions are mainly found in epithelial cells, while anchoring junctions can be found in nearly all cell types. Anchoring and tight junctions are made up of three parts: CAMs allow binding to other adhesive molecules in the extracellular space, adapter proteins, which enable the connection of CAMs to the cytoskeletal filaments or signal molecules and the cytoskeletal filaments themselves.

The primary role of gap junctions is the rapid diffusion of small water-soluble molecules between the cytoplasm of adjacent cells, allowing cells to communicate with their neighbours. Gap junctions are structurally different from anchoring and tight junctions and do not play a structural or strengthening role. Tight junctions are only present in epithelial cells. They allow cell sheets to form impermeable barriers that prevent molecules from leaking. In the small intestine, tight junctions act as fences to prevent the diffusion of unwanted molecules from the gut lumen to the intercellular space, as well as the backflow of glucose basal side of the epithelium to the gut lumen. Four different types of anchoring junctions exist in cells, two types of cell-cell adhesion and two types of cell-matrix adhesion. *Adherens* junctions are mediated by cadherins and connect actin filaments of adjacent epithelial cells, usually near the apical surface, below the tight junctions. *Adherens* junctions can form a type of belt across multiple cells and function as a kind of tension cable, distributing mechanical stress and controlling the shape of the tissue. *Desmosomes* allow for strong cell-cell interactions. Therefore, they can be found in different types of tissue that experience high amounts of mechanical stress, such as epithelial, smooth muscle and heart muscle cells. Hemidesmosomes, actin-linked cell-matrix junctions, and focal adhesions are two mechanisms of cell-matrix adhesion that anchor the cell to the ECM. Understanding different cell-cell and cell-matrix interactions is a crucial point in designing biobots. Sections 2.5.5 discusses how synthetic cell signalling pathways could help scale up the *in vivo* production of biobots.

## Multicellularity

Biobots are a new multicellular form of life. Due to their multicellular nature, they are highly complex and possess a near-infinite degree of freedom in their configuration. Multicellularity comes with advantages as well as disadvantages. A popular method for retrieving these selective advantages is the Darwinian approach of evolution. Multicellularity has evolved independently numerous times across

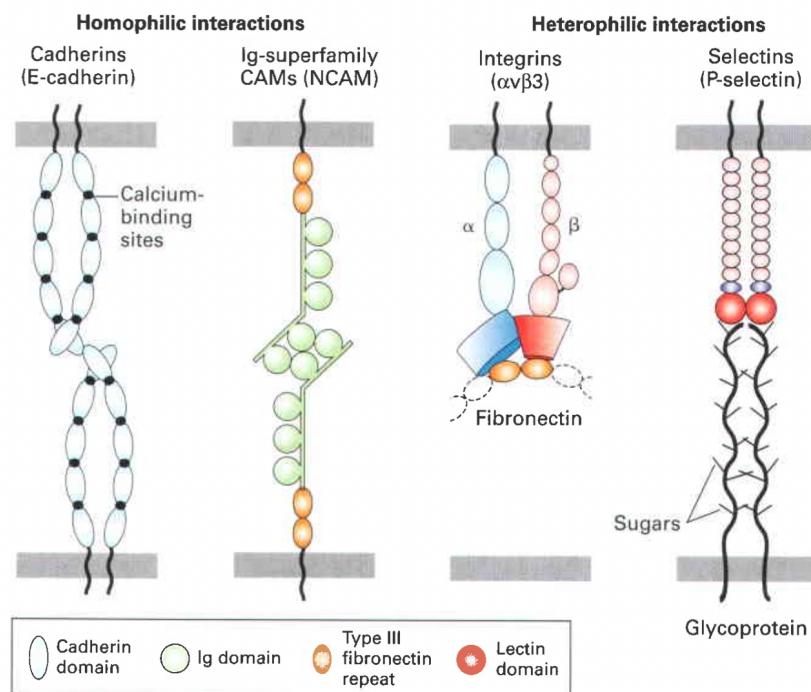


Figure 2.5: Overview of the major cell adhesion molecules (CAMs) and adhesion receptors. Dimeric E-cadherins often form homophilic bonds with other E-cadherins on adjacent cells. CAMs from the Ig-superfamily CAMs can both bind homophilic and heterophilic bonds. Heterodimeric integrins can act as CAMs or as adhesion receptor molecules. As adhesion receptors, they bind to large multi adhesive matrix proteins, such as fibronectin, of which only a small part is shown in this figure. Selectins are single-chain transmembrane glycoproteins that contain a carbohydrate-binding domain that recognizes specific carbohydrate residues and can thus be considered a type of lectin. Most CAMs contain multiple distinct domains, commonly also found in other types of CAMs (Lodish et al., 2016).

multiple lineages (Grosberg and Strathmann, 2007). According to the Darwinian approach, multicellular organisms must have a selective advantage in particular niches compared to unicellular organisms, otherwise, they would not have emerged from them. Most of the macroscopic life on this planet is multicellular. However, most unicellular lineages did not evolve towards multicellular forms, probably because the cost (diffusion limitations) did not outweigh the potential benefits (Tong et al., 2021). In addition, multicellularity is not solely reserved for eukaryotes. Some argue that multicellularity in bacteria and cyanobacteria can arise in the form of biofilms, filaments or aggregates, although this is not considered as real multicellularity by some and debate still exists in the literature (Lyons and Kolter, 2015). There must also disadvantages when operating as a group rather than a single cell. Diffusion limitations play a significant role. However, nature has evolved ingenious ways to solve this problem. Think of the respiratory and circulatory systems in higher vertebrates. Other disadvantages include the energetic cost for the synthesis of CAMs and signalling molecules, reduced freedom of movement, as well as the exploitation of so-called 'cheaters' within a cooperative system (West et al., 2006).

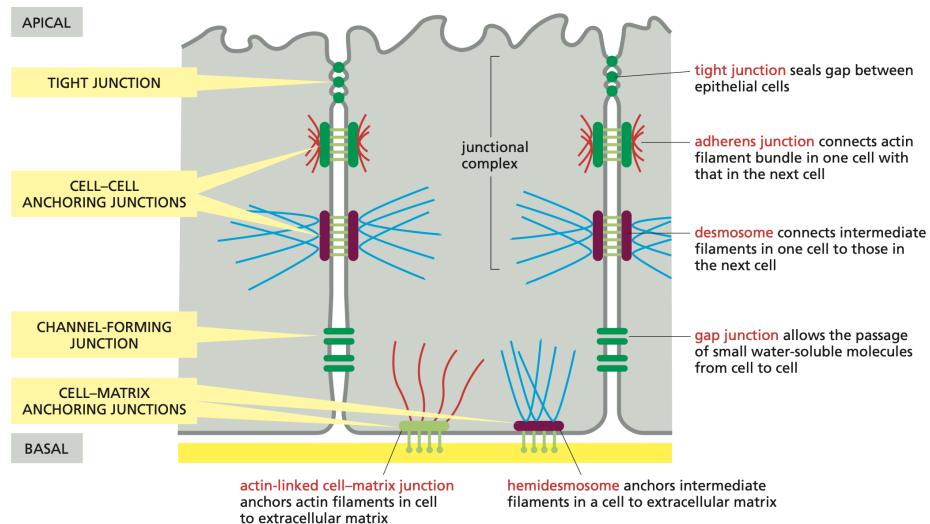


Figure 2.6: Overview of possible junctions in a vertebrate epithelial cell. Tight junctions seal the apical parts of the cells. Followed by a layer of adherens junctions and desmosomes (both anchoring junctions) and gap junctions (Alberts et al., 2015).

### Evolution of multicellularity

Multicellularity has naturally evolved multiple times across distinct lineages (Kawabe et al., 2019). Although evolutionary theories are hard to validate, the literature describes several possible selective drivers for the evolution of multicellularity. Tong et al. (2021) came up with several concepts discussed below.

**Predation avoidance:** Phagocytosis is a frequently applied feeding mechanism at the micro-scale. A simple strategy to prevent predation by phagocytosis is forming groups that are simply too large to be eaten.

**Stress resistance:** The formation of biofilms under stress conditions is a widespread response to different types of stress such as nutrient deficiency, salt stress, low or high pH, oxidative stress and heat shock. Outer layers shelter inner cells, leading to a better tolerance to various types of stress.

**Improved extracellular metabolism:** By forming larger aggregates, the loss of extracellular metabolites by diffusion is limited, increasing their concentration. Enabling aggregates to grow at conditions at which unicellular cells would not survive.

**Faster sedimentation:** Stokes' law dictates that the speed at which particles sediment is proportional to their size and density. Group formation increases the size and thus sedimentation rate. This strategy allows organisms to control their position in aquatic environments.

**Increased motility:** Increasing in size allows organisms to increase their motility. Larger organisms can break through or circumvent physical barriers, allowing them access to more light or nutrients.

**Chimerism:** In clonal multicellular organisms, all cells are genetically nearly identical, as a result, the within-group genetic heterogeneity is minimised. In organisms where multicellularity developed through aggregation, such as *D. discoideum*, chimeric organisms are formed.

**Cross-feeding and division of labour:** Multicellularity allows for complementary metabolisms and differentiation. Cross-feeding allows bacteria to increase their efficiency or resource utilisation and growth rate by specialising their metabolism. This division of labor can be a selective advantage in some highly competitive environments.

**Competitive overgrowth:** Increasing size can allow multicellular organisms to overgrow their competitors. This advantage gives them access to light, nutrients, or water resources.

**Efficient utilisation of patchy resources:** Multicellular organisms can transport nutrients through their cytoplasmic network. By bridging patches where little nutrients are available, multicellular organisms are able to grow in places, where unicellular cells would not be able to.

**Production and dispersal of propagules:** Some organisms can form spores upon stress. Multicellularity can aid in both the production and dispersal of these stress-resistant propagules.

## 2.5.2 *In silico* design of biobots

This section discusses the *in silico* design of biobots. Once a clear goal is defined, and the possible building blocks are defined, the next step consists of designing the optimal biobot to perform this task. Given the near-infinite degrees of freedom of the design space, the usage of computational intelligence (CI) is called upon.

### Encoding

The first step in designing a configuration is defining an efficient and precise encoding. Translating real-world problems to a computer-readable format remains one of the most important steps when working with CI. A suitable encoding can describe all relevant information whilst avoiding unnecessary calculations or returning false information.

Kriegman et al. (2020) use Compositional Pattern Producing Networks (CPPN) as an indirect generative encoding for reconfigurable organisms. Stanley (2007) first proposed CPPNs as an efficient and elegant way of artificially encoding development and the geno- and phenotype relation. The hypothesis behind CPPNs is that patterns found in nature can be artificially described as a composition of functions in the form of a graph or network.

CPPNs are a variation of artificial neural networks (ANN). Unlike ANNs, whose nodes are typically activated by a sigmoid function, each node of a CPPN may use a variety of canonical activation functions, such as sines, Gaussians, re-lu and sigmoids. These activation functions allow CPPNs to produce patterns with specific motifs commonly found in nature and avoid producing patchy or mosaic designs (Stanley, 2007). The resulting network is a *composition* of different functions, able to produce patterns such as

*repetition*(e.g. neurons in a brain), *repetition with variation* (e.g. fingers of a hand), *symmetry* (e.g. bilateral symmetry in a majority of higher vertebrates) and *imperfect symmetry* (e.g. organ distribution in humans). One might expect to produce *symmetric* patterns using a Gaussian function as an activation function. Similar, when using periodic functions, such as sines, one expects *repetitional*, segmenting patterns to emerge. An example of the combination of both functions is shown in Figure 2.7. Combining regular functions (e.g. sine and Gaussian) with irregular functions results in *repetition with variation* or *imperfect symmetry* (Risi, 2012).

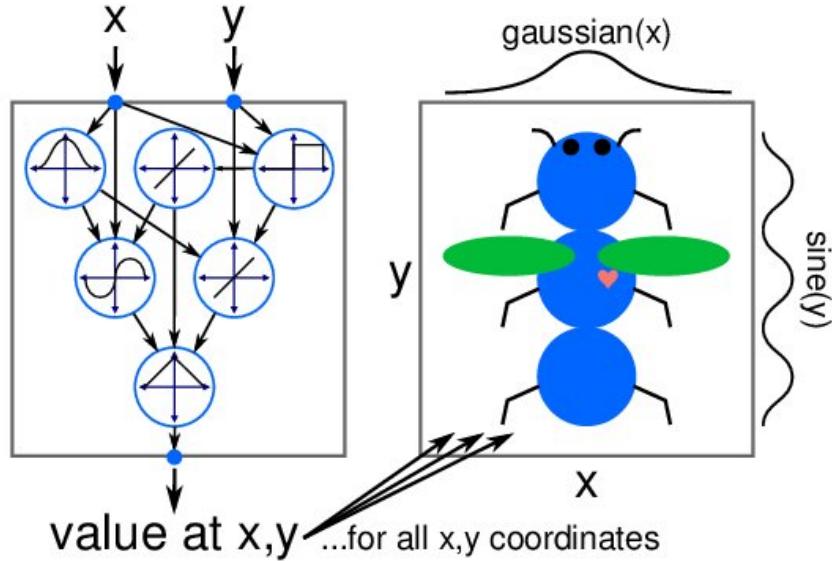


Figure 2.7: A CPPN can generate patterns in a two-dimensional space, with only the  $x$  and  $y$  coordinates as inputs. The bilateral symmetry stems from a Gaussian, and a sine function produces the segmentation and repetition (Coleman et al., 2014).

CPPNs can produce a phenotype that is a function of  $n$  dimensions, where  $n$  is the amount of dimension of the physical solution space (Risi, 2012). A well-known example of the 2D application of CPPNs is the picbreeder.org project (Secretan et al., 2011). Picbreeder.org is a platform where users can make computer designed genetic artworks encoded by CPPNs. Figure 2.8 shows the encoding in further detail.

Note that the encoding is indirect. The phenotype is only achieved by querying the CPPN at every image pixel. Since neighbouring pixels have similar spatial information (coordinates and distance to the centre), they are biased towards forming continuous patterns, preventing the formation of patchy and mosaic patterns. The CPPN can be seen as a kind of blueprint for the design of the biobot. Indirect encoding allows for applying different genetic operators, as later discussed in Section 2.5.4.

Similar to Picbreeder, Cheney and Clune (2013) proposed CPPNs as an indirect encoding for the design of biobots. Instead of pixels, voxels are used. A voxel is the 3D counterpart of pixels in 2D graphics. For each voxel, the CPPN receives four inputs:  $x$ ,  $y$ ,  $z$ , and  $d$ . Where  $(x, y, z)$  represent the cartesian coordinates and  $d$  is the distance to the centre, as shown in Figure 2.9. The CPPN returns two outputs. The first one predicts whether a voxel should be filled, while the second output defines the nature of the material (active or passive) that makes up the voxel. Later this method was expanded by adding an extra CPPN

in series (Corucci et al., 2015). Figure 2.10 shows an overview of the general setup. Both CPPNs take the previously mentioned  $x$ ,  $y$ ,  $z$ , and  $d$  inputs, and a bias  $b$ .

The first morphological CPPN has two outputs. The first dictates whether the voxel is filled or empty. If the voxel is filled, a second output dictates whether it consists of active or passive tissue Corucci et al. (2015); Kriegman et al. (2020).

If the voxel is defined to be filled and active by the first CPPN, the second CPPN dictates the frequency at which the active tissue contracts together with the phase offset. Biobots realise actuation through the periodic contraction and expansion of active tissue. This periodic contraction can be described as a global sinusoidal oscillation. Since there was confusion about how real-world biobots would synchronize the contractions of their tissue, a random phase offset was first implemented. Independently evolving the body (morphology) and brain (control) is a known strategy in robot design (Corucci et al., 2015).

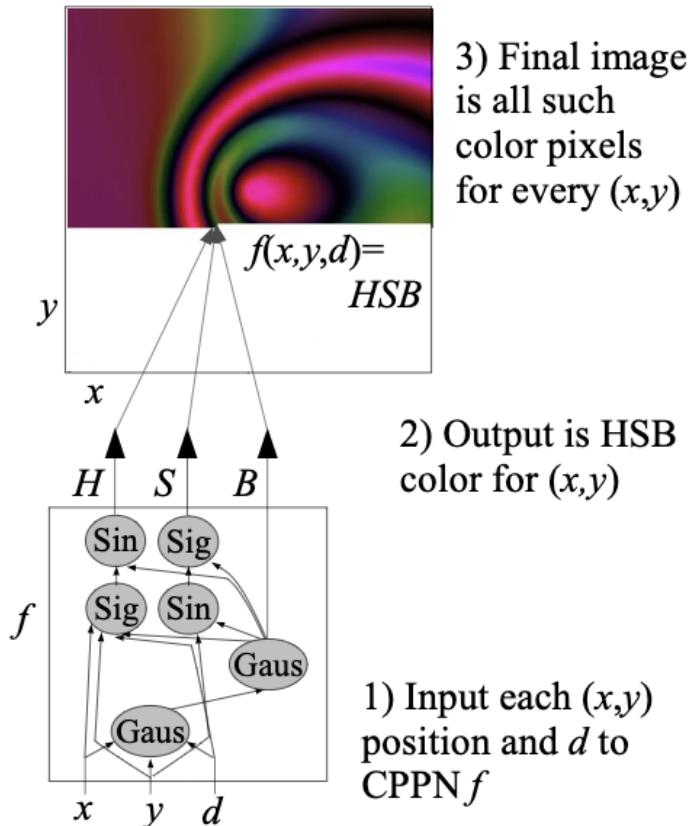


Figure 2.8: A CPPN encoding of a coloured image. The CPPN takes three arguments, the 2D coordinates  $x$  and  $y$  of a pixel and its euclidean distance  $d$  from the centre. For each pixel, three outputs are produced in the HSB format. Finally, a picture is produced by assembling all the pixels in the image (Secretan et al., 2011).

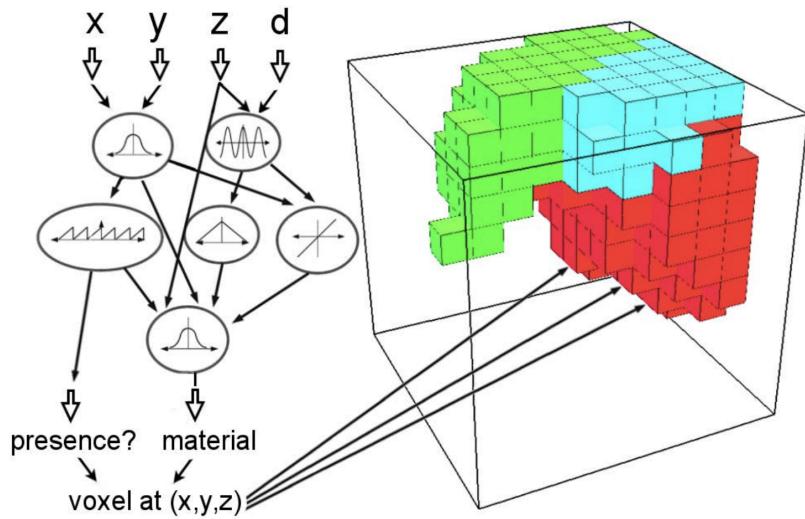


Figure 2.9: For each voxel, a CPPN is iteratively queried. The CPPN takes four inputs:  $x$ ,  $y$ ,  $z$  and  $d$ . Where  $x$ ,  $y$ ,  $z$  denote the cartesian coordinates of the voxel in the 3D grid and  $d$  takes the value of the euclidean distance from the centre of the grid. For each voxel, two outputs are returned. Output one dictates whether the voxel should be filled or not, while output two dictates the material, given output one returns it should be filled. Querying the CPPN for all voxels in the grid returns a the design of the biobot (Cheney and Clune, 2013)

### 2.5.3 Evolutionary Algorithms (EA)

Evolutionary algorithms (EA) are methods to mimic evolutional mechanisms found in nature. Reproduction, mutation, crossover and selection are key drivers in modern evolution theory. Each solution represents an *individual* in the space of possible solutions. For every *individual* or solution, the quality of the solution is calculated by a user-defined function. In biological terms, this would equate to the *individual's fitness* and its associated *genome*. *Individuals* with a high *fitness* score are allowed to *replicate*. *Replication* can be asexual, classical *recombination* or even with multiple *parents* (Eiben et al., 1994). During *replication*, biological mechanisms such as *crossover* and *mutation* are mimicked. Several additional algorithms are proposed to counteract local convergence and explore the entire *fitness landscape*. An important subclass of EAs is the class of Genetic Algorithms (GA).

### 2.5.4 Genetic Algorithms (GA)

Genetic algorithms are a class of EAs. Generally, a GA only requires two inputs: a genetic representation of the solution domain and a fitness function to evaluate the possible configurations. Each solution represents an *organism* with its corresponding *genome*. Further, each *genome* can be broken down into distinct *genes*. Roughly speaking, one can consider a *gene* to encode a particular trait. The search space is defined as the collection of all possible solutions feasible to fulfil the task at hand. The biological synonym would be the *fitness landscape*. In biology, the *fitness landscape* represents all of the possible *genotypes* with their corresponding *fitness*.

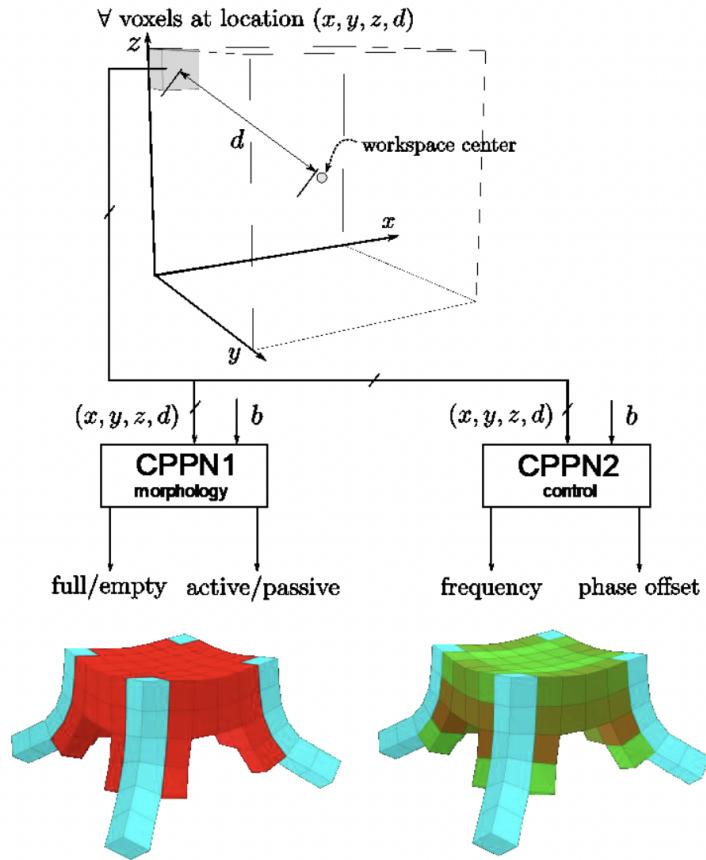


Figure 2.10: Decoupling the evolution of the brain and the morphology is a known strategy in evolutionary robotics. Both CPPNs take four inputs:  $x$ ,  $y$ ,  $z$  and  $d$ . Where  $x$ ,  $y$ ,  $z$  denote the cartesian coordinates of the voxel in the 3D grid and  $d$  takes the value of the euclidean distance from the centre of the grid. The first CPPN returns for each voxel whether it should be filled or empty and in case it is filled, whether the tissue should be active or passive. The second CPPN returns for each active voxel a contraction frequency and a phase offset (Corucci et al., 2015).

### Initialisation

First, the algorithm generates a random set of solution. Every solution is encoded as an *organism*, where its *genome* consists of different genes. The size of the initial population depends on the problem but typically consists of several hundred to thousands of organisms.

### Selection

During each generation of *organisms*, the algorithm assigns a *fitness* score to each individual. Afterwards, all solutions are ranked, and the famous mechanism of *survival of the fittest* is applied. The fittest *organisms* or best performing solutions are allowed to *replicate* and produce *offspring* and form a new generation.

---

A popular objective in evolutionary robotics, and robotics in general, is unidirectional locomotion. However, different *fitness* criteria can be combined by several methods (Coello et al., 2007). What follows is a multi-objective evolutionary algorithm. Corucci et al. (2015) optimised the design according to the following objectives based on the concepts of *Pareto dominance*, as described in Coello et al. (2007):

- maximize distance travelled by the centre of mass (locomotion)
- minimize the percentage of actuated voxels (energy)
- minimize the number of voxels (employed material/ mass)

## Genetic Operators

In GAs, two main genetic operators are employed: mutation and crossover (or recombination).

**Mutation** is the genetic operator analogous to biological mutation. It makes point mutations in the *genome* of the selected *organism* at a user-defined rate  $m_r$ . Mutations introduce genetic diversity into the population. By preventing the best performing individuals from becoming too similar, the convergence into local minima is prohibited.

**Crossover** or recombination is the equivalent of sexual production. The child *genome* is a combination of the parent genomes. Often two parent genomes are combined. Since biological constraints do not limit this process, literature also describes multi-parent recombinations (Eiben et al., 1994). Depending on the application, several different strategies for recombination can be implemented.

Both *mutation* and *crossover* are important mechanisms that ensure the maintenance of genetic diversity within populations. *Mutation* prevents *genomes* from becoming too similar, while *recombination* helps achieve optimal solutions. Both mutation probability and crossover probability are parameters that need to be tuned carefully to find a set of well-performing solutions efficiently.

## Evolution

The most fit individuals of the initial population are allowed to *breed* offspring. The offspring is, in their turn, evaluated, and the best performing individuals are allowed to *breed* and so on. This loop repeats itself until predefined terminating conditions are met. Algorithm 1 describes a typical GA.

## Simulation

In order to evolve the GA, the *fitness* of all designs in each *population* need to be calculated. A popular objective in evolutionary robotics is unidirectional locomotion. Designs that can cover the most distance in a given time period are awarded the highest scores and are allowed to reproduce and create *offspring*.

**Algorithm 1:** Outline of a generic Genetic Algorithm. The algorithm will evolve a number of  $N$  individuals in each generation and return the  $m$  best designs when termination criteria are met (Mitchell, 1996).

---

```
Data: N ≥ 0, m ≥ 1
Result: solution
iteration ← 0;
population ← RandomPopulation(N);
fitness ← TestFitness(population);
while Terminate(fitness, iteration, m) = False do
    parents ← Select(population);
    children ← Crossover(parents);
    population ← Mutate(children);
    fitness ← TestFitness(population);
    iteration ← iteration + 1;
    solution ← MostFit(population, scores, m);
end
```

---

Cheney and Clune (2013) and Kriegman et al. (2020) make use of the open-source VoxCad<sup>1</sup>/Voxelyze<sup>2</sup> modeller and simulator wrote in C++. Most physics engines and simulators are written in C++ for better performance. A Python wrapper called Evosoro<sup>3</sup> has been developed to facilitate experiments Kriegman et al. (2017).

## Neuroevolution

The final part of the *in silico* design consists of performing the design's actual evolutionary optimisation. Since the CPPN encoding strongly resembles the structure of ANNs, genetic algorithms designed for neural network optimisation can be modified and applied. A popular GA for neuroevolution is the NeuroEvolution Augmenting Topologies (NEAT) algorithm, developed by Stanley and Miikkulainen (2002). In traditional reinforcement learning, topology is often fixed, and weights are adapted through training cycles through backpropagation. The NEAT algorithm allows for changes in weights, and the topology or the structure of the network itself. Traditionally the topology is predefined by the researcher. This is often a cumbersome procedure based on prior knowledge and a vast amount of repetitive trial-and-error steps. Neuroevolution was first designed as an alternative to the conventional gradient descent methods in optimising ANNs. Later, the possibility of altering the network was added, omitting the need to predefined the topology.

Evolving networks by GAs allows for applying previously-discussed genetic operators, such as *mutation* and *crossover*, in a sophisticated manner. *Genomes* are a linear representation of the network's connectivity. The *genome* is a list of *connection genes*, where each *gene* refers to two nodes being connected. Thus each *gene* contains the following features of the network: the in-node, the out-node, the weight of the connection, whether or not the connection is active and an *innovation number*, as shown in Figure 2.11. Whenever a new gene appears, the *global innovation number* is incremented by one and it gets

---

<sup>1</sup><https://sites.google.com/site/voxcadproject/>

<sup>2</sup><https://github.com/jonhiller/Voxelyze>

<sup>3</sup><https://github.com/skriegman/evosoro>

assigned to the new gene. Here, the *innovation number* represents the chronology of the first appearance of each gene and allows for *hybridisation* through alignment of genes with the same *innovation number*, a crucial step when recombining genomes.

One of the key mechanisms in evolution is that of *mutation*. In the NEAT algorithm, the *mutation* is implemented so that it can both alter weights and the network structure. Two mechanisms of mutations in the structure of a network are implemented, as shown in Figure 2.11. Mutations always expand the size of the *genome* by adding new genes. The *add connection* mutation adds a new single connection with a random weight between two previously unconnected nodes. The *add node* mutation splits an existing connection and places a new node where the old connection used to be, disabling the old connection and adding two new ones to the *genome*. The new connection leading to the new node gets assigned a weight of 1, while the second connection gets the weight of the connection it replaced, limiting the impact a *mutation* has and allowing for efficient optimisation once a well-performing structure is found.

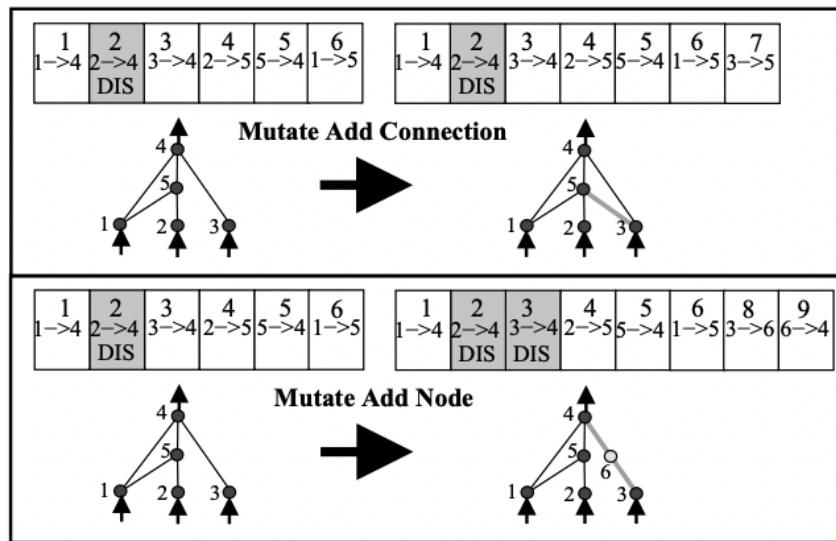


Figure 2.11: Two mechanisms of structural mutation in the NEAT algorithm. Mechanisms either result in the addition of a connection or in the addition of both a node and a connection, always leading to an expansion of the *genome* (Stanley and Miikkulainen, 2002).

For every *gene* that evolves through *mutation*, its origin gets tracked. Whenever a *gene* pops up, an incremental global innovation number is assigned to it. Whenever structures are selected to *mate* or *cross over*, the *offspring* will inherit the same innovation number on each gene, retaining the evolutionary origin of each *gene*. Since the NEAT algorithm knows which genes match up with which, an alignment and *cross over* can be performed. Note the resemblance with chromosome matching in diploid eukaryotes. When two *organisms* are selected to *mate*, analogous *genes* are matched. Genes that do not match can either be disjoint or excess. They represent a structure that is not present in the other genome. The *offspring* contains randomly chosen *genes* from either parent at matching *genes*, whereas all excess or disjoint *genes* come from the more fit parent of the two. This way, the NEAT algorithm performs a linear combination of both *parent genomes* as shown in Figure 2.12.

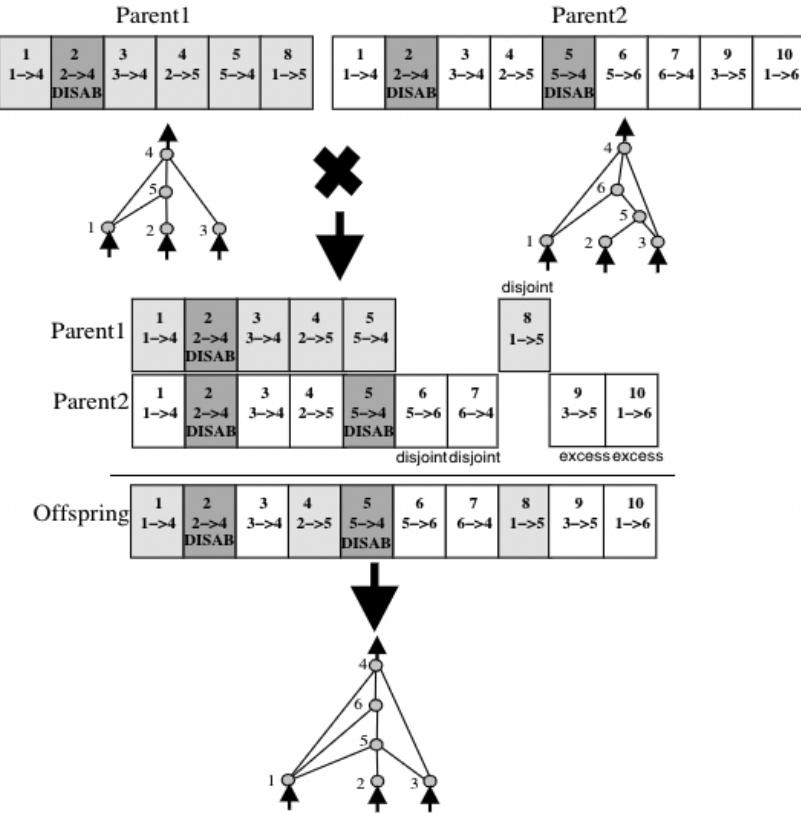


Figure 2.12: Crossing over two distinct topologies for producing offspring. Each gene gets matched up with its corresponding gene on the other genome using its innovation number. Matching genes are inherited randomly, while distinct and excess genes are adopted from the most fit parent. In this example, equal fitness is assumed. As a result disjoint and excess genes are inherited randomly (Stanley and Miikkulainen, 2002).

### 2.5.5 *In vivo* realisation

After the *in silico* design of optimal biobots, the last step consists of the *in vivo* production of the biobot. Currently, this is mainly for validating the predicted behaviour. Once these reconfigurable organisms show real-world applications, as discussed in Section 2.4, other factors, such as scalability, reproducibility and accuracy must also be considered.

#### Microsurgery

Kriegman et al. (2020) used a manual intensive method of producing *in vivo* xenobots. Xenobots are entirely made of embryonic *Xenopus laevis* cells. From the fertilised eggs, the animal cells were manually removed and pooled together in welled dishes. After the outer ectoderm was removed, the inner layers were agitated to dissociate the cells. At this stage, cells are largely pluripotent. After 48 hours, the epidermal cells form a spherical reaggregate. The desired morphology is obtained using microsurgery. A

---

$13\text{-}\mu\text{m}$  wire electrode<sup>4</sup>. Later contractile tissue can be layered on top of the aggregate, as shown in Figure 2.13. This process allowed Kriegman et al. (2020) to validate their pipeline showing remarkably accurate results. Observed *in vivo* behaviour showed to be similar to their *in silico* simulated counterparts.

Microsurgery has shown to be a valid technique for the *in vivo* production of xenobots and allowed researchers to validate their *in silico* predictions. However, this method involves much manual work and would be hard to scale. Real-world application of biobots would require the production of large numbers of biobots and thus, require more scalable methods.

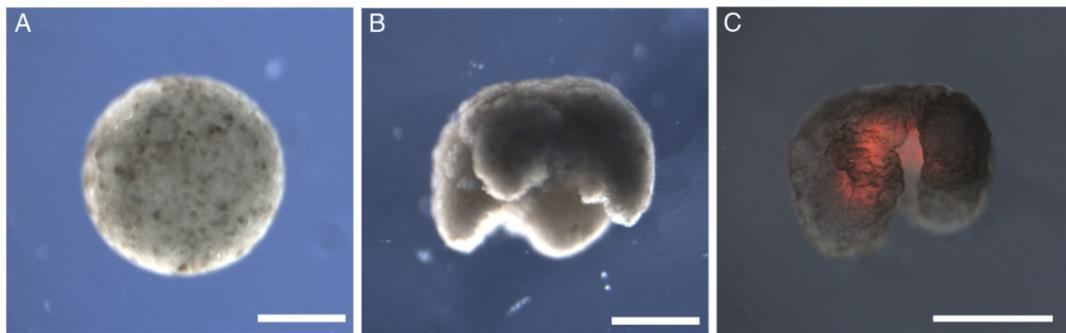


Figure 2.13: The different steps in the *in vivo* production of xenobots using microsurgery. (A) shows the spherical aggregate after pluripotent blastula cells from *Xenopus laevis*. (B) The aggregate after microsurgery by an electronic wire tip. (C) The layering of contractile cardiac progenitor cells at *in silico* predicted places allows for locomotion (Kriegman et al., 2020).

### Artificial cell-cell signalling

Morphological changes can be induced by interfering in the cell-cell signalling of multicellular tissues. As discussed in Section 2.5.1, cell-cell and cell-matrix adhesion mechanisms play a prominent role in the morphology of a specific tissue. Toda et al. (2018) developed a toolkit based on synthetic Notch (synNotch) to alter the self-organisation of multicellular mouse fibroblast aggregates artificially. The Notch signalling pathway is a highly conserved cell interaction mechanism, playing a fundamental role in metazoan development Weiner et al. (1999). The engineered synNotch receptor is composed of the core regulatory domain of the Notch receptor, linked together with a chimeric extracellular recognition domain and a chimeric intracellular transcriptional domain (Morsut et al., 2016). Upon recognising a complementary ligand on an adjacent cell by the synNotch receptor, the receptor undergoes cleavage of the transmembrane region. The released transmembrane region can enter the cell nucleus and alter the expression of user-defined genes. Figure 2.14 shows a schematic representation of the synNotch construct. Various types of genes can be under the control of the transcriptional domain.

To get cells to self-organise in a two-layered structure, Toda et al. (2018) engineered two types of L929 mouse fibroblasts. Cell type A or the *sender* cell, is engineered to express the synNotch ligand CD19 and blue fluorescent protein (BFP). Type B or the *receiver* cell expresses the anti-CD19 synNotch receptor and the corresponding response element. E-cadherin (Ecad) and green fluorescent protein (GFP) are

<sup>4</sup><https://www.youtube.com/watch?v=kCOKtmNH9o>

put under the control of the transcriptional domain of the synNotch mechanism. When A-type and B-type cells are cocultured together, B-type cells are activated to express Ecad. Additionally, the cells self-sorted to form aggregates, where green activated B-type cells formed a tight inner core surrounded by blue A-type cells, as shown in Figure 2.15.

Additionally, Toda et al. (2018) developed a signalling cascade to form self-organising three-layered structures. An extra layer was put on top of the above two-layered circuit. The B-type cell was engineered as such that upon activation, additionally to E-cad, it also expresses a surface-tethered GFP as a secondary synNotch ligand. In addition to BFP and the C19 synNotch ligand, A-type cells also express an anti-GFP synNotch receptor. Upon activation, A-type cells are driven to express low amounts of E-cad as well as the mCherry fluorescent reporter gene. This way, a bidirectional signalling cascade is constructed. First, B-type cells are activated to express high amounts of E-cad and surface-tethered GFP. This GFP activates A-type cells, which express a low amount of E-cad and mCherry. Since only A-type cells, that are close to activated B-type cells, become activated, this mechanism gives rise to the formation of three distinct layers: green activated B-type cells, red activated A-type cells and blue non-activated A-type cells. A three-layered aggregate is formed, with a green internal core with the highest homotypic adhesion, an outer layer of blue non-activated A-type cells and a middle-interface layer of red activated A-type cells with intermediate levels of homotypic adhesive cell-cell interactions, as shown in Figure 2.15.

The method of synNotch is highly scalable and accurate. Toda et al. (2018) showed that more complex three-layered aggregates can be formed by altering signalling cascades. Artificial cell-cell signalling is a promising method for producing biobots at a large scale, which would be needed for real-world applications.

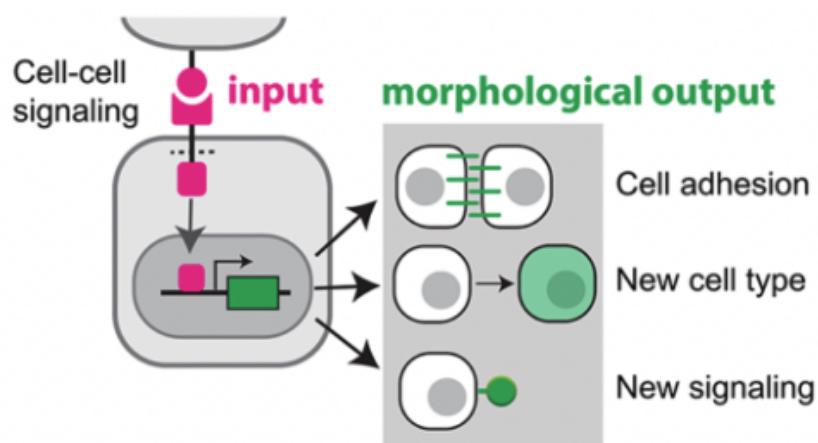


Figure 2.14: A schematic representation of the synNotch Mechanisms. Once the extracellular recognition domain recognises the proper ligand on an adjacent cell, the intracellular domain gets cleaved and can enter the nucleus. In the nucleus, it acts as a transcription factor (TF) and alters the expression of user-defined genes.

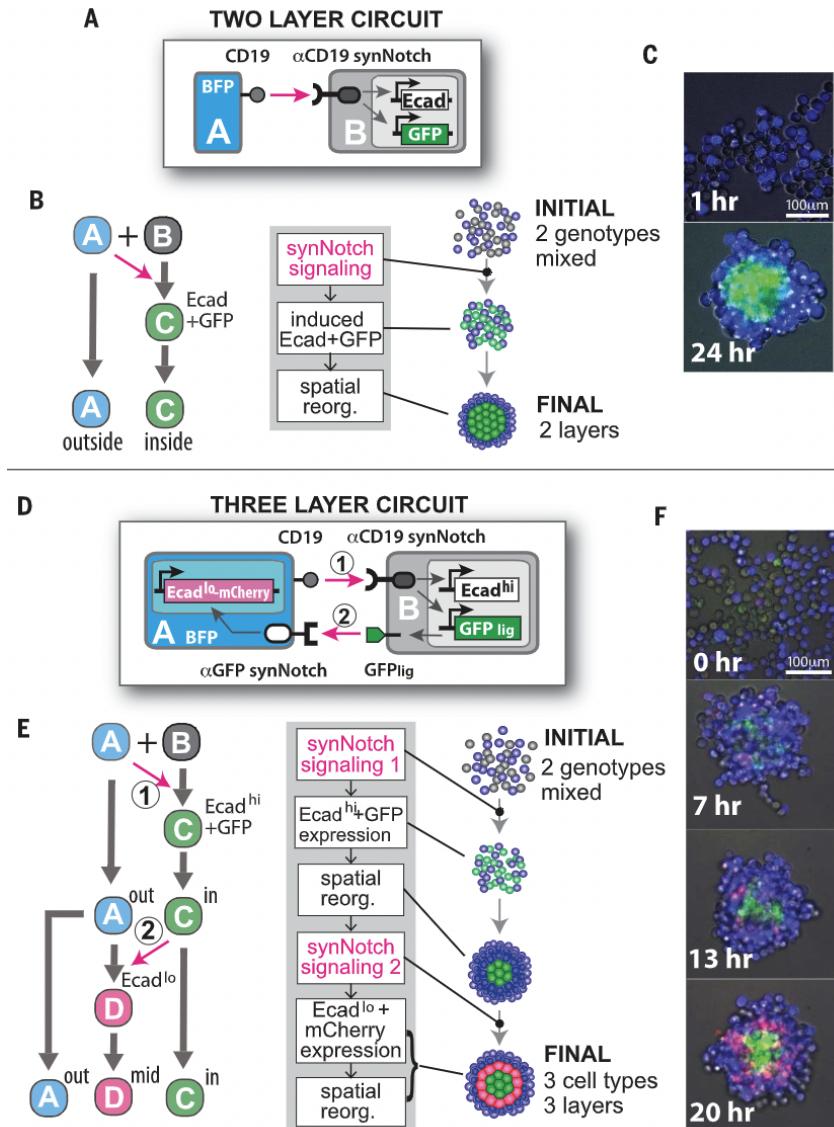


Figure 2.15: (A) A schematic overview of the implemented synNotch tool. Blue type-A cells activate the expression of E-cadherins (Ecad) and green fluorescent protein (GFP). (B) SynNotch cell-cell signals induce the expression of both Ecad and GFP. The higher expression of Ecad leads to homotypic cell-cell interactions, leading to the formation of an aggregate with spacial compartmentalisation. Green type-B cells make up the inner core, surrounded by blue type-A cells. (C) Microscopic images of the self-organisation of cells in a structured aggregate. (D) Schematic overview of the bidirectional synNotch construct. A-type cells express the CD19 ligand, BFP and an anti-GFP synNotch receptor. B-type cells express an anti-CD19 synNotch. Upon recognition, activated B-type cells express high amounts of E-cad and surface-tethered GFP, leading to strong homotypic adhesion and spatial reorganisation. A-type cells recognise the GFP produced by activated B-type cells and become activated themselves, expressing low amounts of E-cad and mCherry (Toda et al., 2018).

### **3. DICTYOSTELIUM DISCOIDEUM**

*Dictyostelium discoideum* is a remarkable microorganism and emerges as an interesting cell donor organism and model system for the development and manufacturing of a new type of biobots.

#### **3.1 Relevance of Dictyostelium as a model organism**

*Dictyostelium discoideum* is the best-studied species of all social amoeba, commonly referred to as *slime moulds*. *Dictyostelium* is a soil-dwelling amoeba of, around 10-20  $\mu\text{m}$ , with a unique life cycle. In nature, *Dictyostelium discoideum* is found in deciduous forest soil and decaying leaves, where the amoebae feed on bacteria through phagocytosis and multiply by asexual mitotic division (Eichinger and A.Noegel, 2003). It is a well-known model organism for studying the actomyosin cytoskeleton and chemotactic motility in non-muscle cells (Annesley et al., 2009). Its 34 Mb haploid genome, containing six chromosomes, is publicly available, along with a variety of tools and information at the dictybase.org website (Chisholm et al., 2006). The genome of *Dictyostelium discoideum* contains many genes homologous to higher eukaryotes, which are missing in other microscopic model organisms such as *Saccharomyces cerevisiae*. The homologous genes, combined with a relatively short life cycle, make *Dictyostelium discoideum* a valuable model organism. It can be observed at the organismal, cellular and molecular levels and is well suited to study biological processes such as cytokinesis (Noegel and Schleicher, 2000; Nagasaki et al., 2008), cell motility (Carnell and Insall, 2011), phagocytosis (Junemann et al., 2016), chemotaxis (Scavello et al., 2017), signal transduction (Suess et al., 2017), quorum sensing (Singh et al., 2017) and cell differentiation during development (Chattwood et al., 2013).

With its remarkable life cycle of *Dictyostelium discoideum* shows to be an interesting cell donor organism for the development and manufacturing of biobots. In order to discover the potential of *Dictyostelium* as a cell-donor organism, the next section gives an overview of the life cycle of *Dictyostelium*. Section 3.3 discusses how different features of *Dictyostelium* could be exploited in order to derive biobots from it.

#### **3.2 Life cycle**

Under normal conditions *Dictyostelium discoideum* is found in a vegetative state. At this stage, *Dictyostelium discoideum* is a free-living unicellular soil amoeba, preying on bacteria and constantly proliferating through asexual mitotic division. Cells can go through many generations whilst constantly competing for food (Kawabe et al., 2019). Individual cells will remain to be in this vegetative state for as long as the environment will allow them and sufficient amounts of food is present. However, when all

prey is consumed, the developmental cycle will continue. A schematic overview of the life cycle of *Dictyostelium* is given by Figure 3.1. If the amoebae undertake no action, they will die due to starvation (Kessin, 1975). In response to this starvation stress, three distinct responses can occur. *Dictyostelium discoideum* can either form microcysts, macrocysts or fruiting bodies. In the scope of this thesis, the last response is by far the most interesting.

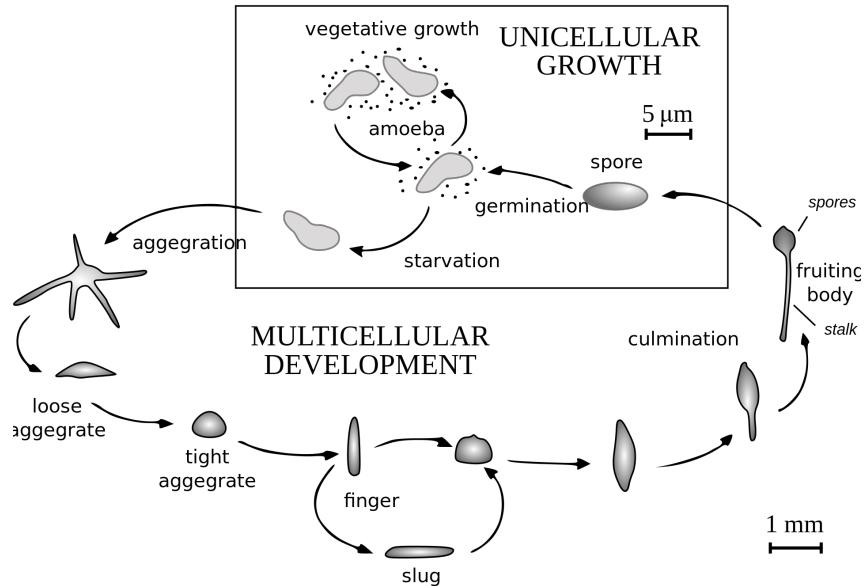


Figure 3.1: The cell cycle of *Dictyostelium discoideum* starts at the vegetative state. Unicellular amoebae dwell the soil looking for bacteria to prey on. During this stage, cells proliferate by means of mitotic fission. Under starvation stress cells become susceptible to cAMP and start secreting cAMP in their environment. Unicellular amoebae start actively moving up the gradient of cAMP and aggregate to form a pseudoplasmodium or *slug*. At early stages of aggregation, cells develop to either prespore or prestalk cells. Within the aggregate, cells perform a salt and pepper sort to form strictly regulated patterns of different cell types. As a result a migratory *slug* is formed which is phototactically, chemotactically and thermotactically sensitive. When culmination criteria are met, the prestalk cells will mature in stalk cells, lifting up the spores. The spores are then released into the environment and will germinate under the appropriate conditions. Germinated spores will develop as vegetative unicellular amoebae, completing the life cycle of *Dictyostelium discoideum*.

### 3.2.1 Aggregation

In order to form this multicellular fruiting body or sorocarp, spatially close individuals need to find one another. This phase is called the aggregation phase. Starvation induces the transcription of a variety of new genes, whose products are required for aggregation. The amoebae become aggregation competent and start secreting and chemotactically responding to some kind of pheromone. This pheromone was named *acrasin*, a generic name for the chemoattractants produced by a member of the order Acrasiales<sup>1</sup>. Two decades later Konijn et al. (1968) proved that in fact the *acrasin* in case of *Dictyostelium discoideum* was cyclic adenosine monophosphate (cAMP). After a couple of hours some amoebae begin to secrete cAMP into the population. Upon detection of this signal, other cells will start to produce enzymes which can

<sup>1</sup>A family of cellular slime moulds belonging to the Percolozoa

produce, secrete, detect and modulate the cAMP signal (Kessin, 1975). When no cAMP is present, *Dictyostelium* cells perform a random walk by extending and retracting pseudopods in all directions. When cells sense a cAMP gradient, they become polarised and start extending and retracting pseudopods up in the direction of rising cAMP concentrations. In the mean while, cAMP represses the formation of lateral and posterior pseudopods, resulting in locomotion up the cAMP gradient (Soll et al., 2003).

The cAMP signal, used by *Dictyostelium discoideum* is relayed, which allows *Dictyostelium* aggregates to collect cells from a wide area and form groups of over 100,000 cells. A starving cell releases pulses of cAMP, and other cells detect the molecule as it binds to their specific receptors. Upon detection, cells respond in two different ways. They release a secondary pulse of cAMP in the environment and move up the gradient of cAMP released by the first cell. These mechanisms result in an outwardly propagating wave of cAMP and an inward movement of chemotactically motile cells. Both the cAMP propagation and the cell movement occur in distinct steps. Central cells release a pulse of cAMP around every six minutes (Kessin, 1975) and only moving cells proceed to migrate towards increasing concentrations of cAMP. Whenever the cAMP wave fades, cells will cease their movement. These wave-like patterns allow *Dictyostelium discoideum* to form aggregates with cells from aggregation territories as much as one cm across. As the cells arrive at the aggregate, they become adhesive and form three-dimensional structures. At this stage, the loose aggregate is formed. Figure 3.3 shows an electron microscopic image of the different post-aggregation stages of *Dictyostelium discoideum*.

### 3.2.2 Pseudoplasmodium formation

Already in early stages of the development of the aggregate, cells begin to differentiate into either prespore or prestalk cells. In the early stages both types of cells are homogeneously mixed, which results in a salt and pepper pattern (Weijer, 2004). After a couple of hours, the aggregate tightens up and some prestalk cells sort out and form a tip. The tip is a new element that consists of an essential group of cells that organize the behaviour of the cells behind it. Under control of the tip, the aggregate elongates, and a new finger-like structure develops. At this stage, the aggregate can directly mature into a fruiting body. However, the finger can also fall over and form a migration slug or pseudoplasmodium. During the formation of the pseudoplasmodium cells differentiate into different types of subtypes and rearrange within the pseudoplasmodium. As a result highly regulated tissue patterns are formed as shown by Figure 3.2. Cells are able to self-organise by specifically regulating and modulating different types of cell-cell adhesion mechanisms (Williams, 2006).

### 3.2.3 Migratory slug

This multicellular slug thanks its name to its superficial resemblance of a regular garden slug and is around 2-4 mm long, and contains an estimate of 100,000 cells (Srinivasan et al., 2000). This migrating slug leaves behind its extracellular matrix, as it moves. The slug is remarkably phototactically and thermotactically sensitive as it is able to move upwards gradients as shallow as  $0.04\text{ }^{\circ}\text{C}/\text{cm}$  and  $1\text{ }\mu\text{W}/\text{m}^2$  of white light, the equivalent of a bright star (Annesley et al., 2009). These mechanisms serve to take the

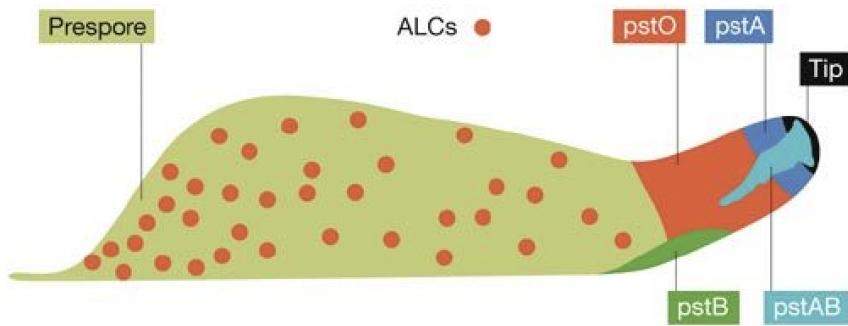


Figure 3.2: A schematic representation of different subtypes in the *Dictyostelium* migratory slug. Prestalk cells (pst) migrate to the front, while prespore cells form the back of the slug. Anterior-like cells (ALCs) are a type of prestalk cells that form a structure that cradles the spore head during the final culmination stages Williams (2006).

slug to the soil surface, which is typically hotter and lighter than soil itself. Once the *slug* arrives at the soil surface, the *slug* will start to culminate.

*Dictyostelium discoideum* cells, once they formed aggregates, strongly resemble the features of an embryo of higher vertebrates. They show polarity, along with exquisite proportionality. They regulate and have an organising centre, the anterior tip. During culmination or maturation of the fruiting body, the formation or extension occurs in a manner that is dependent on the position. This process strongly resembles the pattern formation in non-regulative animal embryos and higher plants. By the time the aggregate forms the tip, cells have already differentiated into one of two major cell types: prespore cells and prestalk cells. Prestalk cells (of which several subtypes exist) are located in the first one-fifth of the *slug* (including the tip), while prespore cells make up the rest of the *slug* (Williams, 2006). As their name implies, they are fated to eventually form either the stalk of the fruiting body or the spores. Neither of these cell types is irreversibly committed. Until the final stages of development, one type can still convert into the other.

In addition, when given food, both cell types can revert to the unicellular amoeboid state (Kessin, 1975). During the aggregation phase, essential prespore or prestalk genes have been expressed in both the precursor populations, and by the tipped stage, most of the internal cell sorting has already taken place. A particular class of prestalk cells migrate through the substratum and start the formation of the stalk. During this period, stalk cells become highly vacuolated and undergo a sort of autophagic, programmed cell death (Golstein et al., 2003). Prespore cells are gradually lifted up, as shown in the last stages of Figure 3.3. The most left stage is called the Mexican hat. The following stages are early to late culminants and lead to the final morphology of the fruiting body.

### 3.2.4 Culmination

As the prespore cells get lifted by the stalk cells, they undergo encapsulation. The final structure of the fruiting body now contains a spore mass, supported by a stalk, where the stalk to spore ratio is about 1:4, leading to a total of around 80,000 spores. At this stage, spores are released into the environment.

ment. When the right conditions are met, spores will germinate, completing the cycle at the unicellular vegetative state.

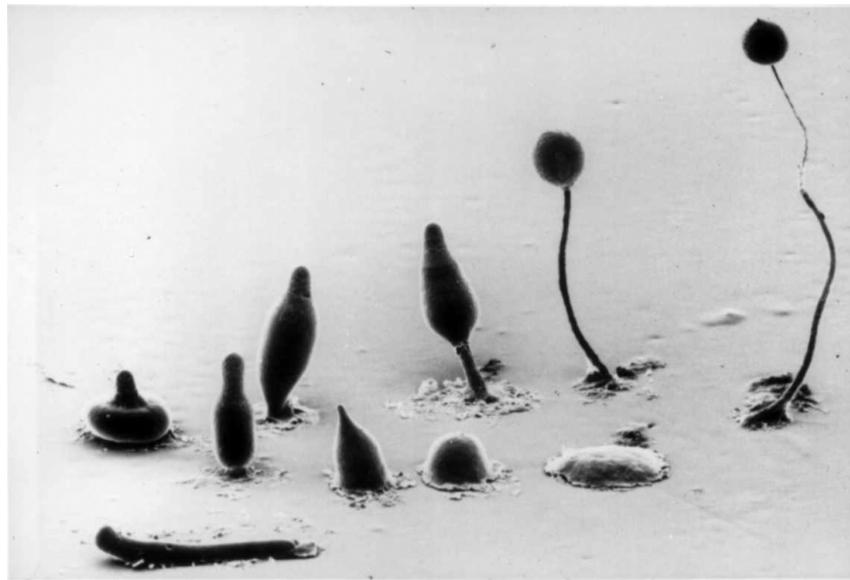


Figure 3.3: The different stages of *Dictyostelium discoideum* after aggregation. Counter-clockwise from the bottom left are the loose aggregate, the tight aggregate, the tipped aggregate, the finger, the slug and then the stages of culmination from the Mexican hat stage to the mature fruiting body Kessin (1975).

The remarkable life cycle of *Dictyostelium discoideum* has intrigued scientists for over a century by now. Looking at *Dictyostelium discoideum* through the lens of evolutionary theory, some remarkable features can be observed. Nearly all studied multicellular organisms result from a blastocyst and a zygote. As a result, all cells in the body are genetically identical, neglecting somatic mutations, the immune system and post-meiotic cells. Since all cells are identical, genetic drift does not take place until after meiosis; however, since *Dictyostelium discoideum* does not evolve from a blastocyst but rather through the aggregation of individuals in spatial proximity *Dictyostelium slugs* are genetical chimera, making them interesting model organisms when studying the origins of multicellularity.

### 3.3 Dictybots?

*Dictyostelium discoideum* has a remarkable, yet highly regulated life cycle. This section discusses three particular events from its life cycle that make *Dictyostelium discoideum* an interesting cell donor organism for the manufacturing of dictybots.

#### 3.3.1 Unicellular aggregation

During the aggregation phase, unicellular *Dictyostelium* are able to chemotactically move up a gradient of cAMP. Using *Dictyostelium* cells as a building material for the development of biobots, this feature could be exploited. *Dictyostelium* can be cultured in high densities both as surface cultures, as well as in suspension (Fey et al., 2007). Since *Dictyostelium* is relatively easy to culture, producing *Dictyostelium*

---

cells is more scalable than those of higher vertebrates, like *Xenopus laevis*, since it does not involve any harvesting of embryos or dissociation of embryonic cells. After culturing in either suspension or on a surface, cells can be made susceptible to cAMP by methods described by Fey et al. (2007). Cell density on a Petri dish can be controlled, therefore, also the size of the resulting aggregates could be controlled. Dropping small amounts of cAMP on a Petri dish full of aggregation competent *Dictyostelium* cells would give rise to the formation of multicellular pseudoplasmodia. Making the production of dictybots a highly scalable process, while sidestepping the ethical hurdles, when working with embryonic cells. In addition *Dictyostelium* can be easily genetically manipulated, in order to develop real-world applications for these dictybots (Paschke et al., 2019).

### 3.3.2 Salt and pepper sorting

*Dictyostelium* are natively able to self-organise into strongly regulated patterns. The specific sorting is obtained by modifying cell-cell adhesion mechanisms. During the *in vivo* manufacturing of xenobots, Kriegman et al. (2020) layered different types of tissue by means of micro surgery, which involves a lot of manual labour and is hard to scale. *Dictyostelium* cells natively possess over the ability to self-organise into specific layers of tissue. Exploiting this feature can drastically improve the throughput for the production of biobots. Different types of cells can be engineered to sort together, obtaining complex morphology's.

### 3.3.3 Migratory slug

After aggregation of unicellular *Dictyostelium* amoebae, a multicellular pseudoplasmodium or *slug* is formed. The migratory *slug* is phototactically, chemotactically and thermotactically sensitive. As a result, dictybots are already capable of directed locomotion. By engineering *Dictyostelium* cells, dictybots could potentially perform some real-world applications. Since the *slug* is natively chemotactically sensitive, it shows as a promising method for environmental biosensing and remediation of chemical pollutants in soil. *Dictyostelium* cells could be equipped with the genetic machinery to detect and break down different chemical compounds.

Another application of dictybots could be in the counteracting the problems associated with micro plastic in the oceans. Recently, insect cells and their symbionts have proven to be able to degrade plastics. Introducing the genetic machinery into dictybots, might enable them to perform the same plastic bioremediation. Another option would be to introduce insect cells directly into the dictybots, making chimeras.

Various aspects about the *Dictyostelium* life cycle show as promising features for the development of dictybots, however this would involve a lot of additional advances in metabolic engineering, synthetic biology and culturing of *Dictyostelium*.

### 3.4 Experiment

In order to gather data on the aggregation phase of *Dictyostelium discoideum* an experiment was conducted. A *Dictyostelium discoideum* kit was obtained from Carolina Biological Supply Company<sup>2</sup>. For 72 hours, every eight hours, two plate cultures were started up according to the manual included in the kit. *Dictyostelium* cells were cultured together with *E. coli* to prey on. Figure 3.4 shows an ented plate after 48 hours. After 72 hours, the eighteen resulting plates were placed under an upright light-microscope<sup>3</sup>, equipped with a high resolution camera in order to record the aggregation phase of *Dictyostelium* over time. However, when analysing the plates, they showed to be contaminated and no *Dictyostelium* cells had grown on them. As there was only one growing kit available and there was no sufficient time to repeat the experiment, no results were obtained from this experiment.

The cause of the contamination is difficult to exactly point out. Contamination could be caused by a bad quality of the kit, the use of contaminated lab materials or during the manipulation by the operator.



Figure 3.4: *Dictyostelium* cells were cultured together with *E. coli* to prey on. The *E. coli* was ented in a North-South orientation and *Dictyostelium* was plated perpendicular to the *E. coli* in an East-West orientation. This allows for some *Dictyostelium* cells to proliferate in the vegetative state, while others start the aggregation phase.

---

<sup>2</sup>[https://www.carolina.com/slime-molds/dictyostelium-culture-kit/FAM\\_155824.pr](https://www.carolina.com/slime-molds/dictyostelium-culture-kit/FAM_155824.pr)

<sup>3</sup><https://www.leica-microsystems.com/products/light-microscopes/p/leica-dm6-b/specification/>



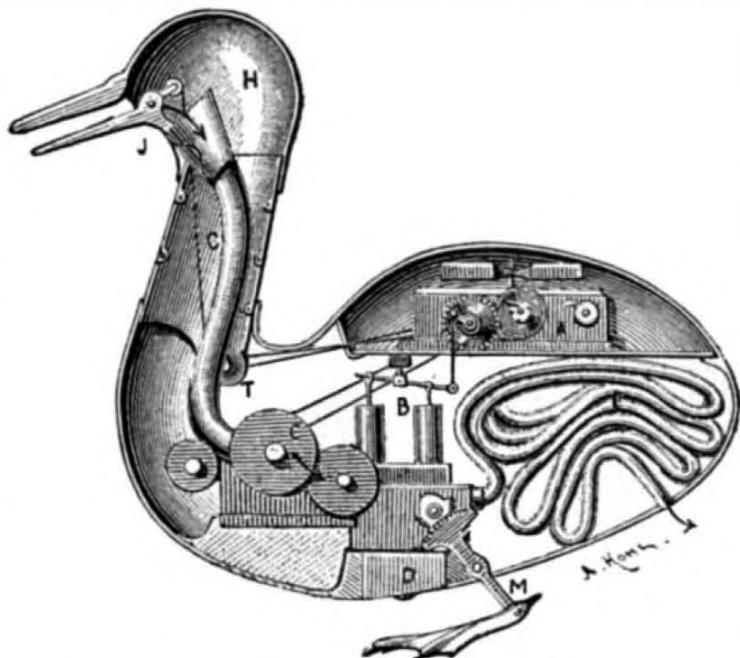
## **4. INFORMATIONAL INDIVIDUAL**

### **4.1 Complexity of biology**

At the beginning of this century, when Stephen Hawking was asked what he thought of the common opinion that the twentieth century was that of physics and the twenty-first century would be that of biology. Hawking replied that, in his opinion, the twenty-first century would be the "century of complexity". In a way, when studying biology, one is always studying the complexity associated with it. Over the last couple of decades, advances in Artificial Intelligence (AI) have allowed scientists to untangle the biology partly. Scientists make use of sophisticated machine learning techniques. The need for these vast and complex algorithms implies that the enormous amount of complexity is a couple of orders of magnitude larger than the human brain would comprehend, thus making any attempts at fully understanding biological life futile.

Although, in the seventeenth century, with essential contributions from the renowned polymath René Descartes, efforts were made towards studying this complexity. He argued that the world could be described as a machine, like a clockwork mechanism, and that it could be understood by taking its pieces apart, studying them piece by piece. After the exact working mechanisms of each piece are thoroughly understood, one could put them back together and understand the machine as a whole. Later, this way of thinking would become known as reductionism and would be one of the main methods in philosophy and science in the centuries to come. The fundamental idea behind the method of reductionism is that the whole is nothing more but the sum of all its parts, and one can extrapolate the derived features of the parts to the whole in order to understand the way it operates and behaves under different circumstances. Figure 4.1 depicts a conceptualisation of the method of reductionism.

However, in the case of biological life, this approach is not sufficient to capture the full complexity of nature. When breaking down biology in its fundamental processes as translation, transcription and cell-cell signalling, no physics or chemistry remains unaccounted for (Krakauer et al., 2020). Throughout the history of biology, this has been observed several times. The most famous instance probably is the complete discovery of the human genome. After completing the human genome project (HGP) in 2001, researchers from all over the world were able to sequence the entire human genome and locate the majority of gene encoding regions. Researchers hoping to fully understand biological life by cracking the genomic code would be left disappointed. Over two decades later, mechanisms behind the phenotype-genotype coupling are yet to be fully understood. Although fundamental biological processes are fully understood, biological life is not. Suggesting that biology is an emergent property, where the whole is greater than the sum of all its parts. This emergent property can be observed on all levels of biological organisation, from cell organelles to populations, to the biosphere.



**INTERIOR OF VAUCANSON'S AUTOMATIC DUCK.**

A, clockwork; B, pump; C, mill for grinding grain; F, intestinal tube; J, bill; H, head; M, feet.

Figure 4.1: Descartes argued that all non-human animals could be explained as automata. This figure depicts the *Canard Digérateur*, an interpretation of this idea by the French inventor and artist de Vaucanson. The duck as a whole can be explained by a series of simple processes. The Cloaca machine by Wim Delvoye is a contemporary conceptualisation of this idea.

Another interesting observation is that physics and chemistry are universal. The entire galaxy is subject to the same fundamental laws of physics and chemistry. However, as far as we know, biology is solely reserved for the planet Earth (Krakauer et al., 2020). All chemical phenomena arise from quantum mechanical laws. Even though biology directly stems from these perfectly predictable laws, biological life remains to surprise. This chapter explores the surprises and proposes a framework to quantify them.

A crucial concept in the studies of genetics, immunology, evolution, development, embryogenesis, anatomy and physiology is that of the biological individual (Gilbert et al., 2012). Depending on the field, the definition of individuality can be filled in different. This phenomenon reflects one of the long-standing criticisms of biology being an *ad hoc*, qualitative science, lacking mathematical foundations or principles (Thompson, 1945). The concept of individuality acts as a basic container or unit of life, and despite the near-universal assumption of it, there is little agreement on what individuals are (Krakauer et al., 2020). At first glance, this concept does not seem very complicated or important. However, defining what is and what is not an individual could have significant implications in different areas of science. Current criteria, based on features like cell membranes and genetics for defining individuality, do not allow to describe the complexity of organisms.

Additionally, Krakauer et al. (2020) argue that the individuality of a system should not be limited to a binary 1 or 0. Complex emergent systems can display a certain degree of individuality. For example the *Dictyostelium* amoebae discussed in Section 3. At the unicellular stage, these amoebae would be classified as individuals. Classifying the multicellular aggregate or pseudoplasmodium would cause typical criteria for individuality to contradict. A common approach for defining individuality is through genetics. Since multicellularity did not arise through mitotic division, cells are not genetically identical, supporting the claim of identifying each cell as an individual. However, the multicellular aggregates show some form of individuality as well. Cells voluntarily sacrifice themselves by going into apoptosis in order for the species to survive. In this example, the common criteria for describing the individuality of a system fail to draw a clear line to determine at what point the feature of individuality is transferred from the unicellular amoeba to the migratory slug. Boundaries of individuality show to be more fluid than the current framework for defining it allows. Given that an average human body hosts around as many bacterial cells as cells it is built from, a decision has to be made to take those into account or not. Another criterion for assuming individuality is that individuality is presupposed by replication (Wilson and Barker, 2007). Work on social insects, such as ants, shows that this idea fails to describe individuality properly. Physically distinct ants aggregate into colonies. The vast majority of worker ants in a colony will never replicate, nor does the colony as a whole replicate, suggesting that individuality without replication is possible. In biology, viruses have always been a point of discussion. Defining non-living particles as individuals seems counterintuitive, yet viruses can replicate, adapt and alter their environment (Krakauer et al., 2020). In order to formalise the concept of individuality and decouple it from, what seems to be an arbitrary or convenient choice, Krakauer et al. (2020) aimed to develop a mathematical framework in order to quantify the *degrees* of individuality in different branches of biological life.

In order to formalise the concept of individuality, Krakauer et al. (2020) first proposed some assumptions that universally apply:

- individuality appears as a spectrum, and some systems can possess greater amounts of individuality than others;
- individuality can emerge at any level of organisation and is not solely reserved for the organismal level;
- individuality can be nested. A system can possess individuality on different levels of its internal organisation. In the *Dictyostelium* example, both the multicellular pseudoplasmodium as a whole and the cells of which it is made can assert a certain degree of individuality.

According to Krakauer et al. (2020), the critical aspect which defines individuality is the way information is propagated from the past to the future and how a system can act on this information. Krakauer et al. (2020) propose a coarse-grained mathematical frame based on information theory in order to quantify the individuality of a system. Revealing the patterns of information streams in a system allows one to calculate its individuality.

---

## 4.2 A digression into information theory

The field of information theory was founded by Claude Shannon in 1948, with the publication of his paper called *A Mathematical Theory of Communication* (Shannon, 1948). In his paper, Shannon clearly defined the concept of information and, more importantly, how to quantify it. As of today, it is still Shannon's theory that allows us to quantify information and the rate at which it can be transferred. Information can flow within any system, man-made or biological (Stone, 2015).

### 4.2.1 What is information

Information is usually expressed in *bits*, where one *bit* of information stores the information present in a choice between two equally probable alternatives. Figure 4.2 clearly illustrates this idea and gives an intuitive interpretation of this measurement. A person standing at the crossroads at point A needs to find his way to point B. He can either decide to take the left path or the right path. When no prior knowledge is given, and the chances of either going left or right are identical, he would arrive at his destination with a probability of exactly 50 percent. However, if one were to tell him which road to take, he would always arrive at the right point, and there would be no more uncertainty in the system. This instruction can be represented by a binary digit, commonly known as a bit, where 0 represents left and 1 represents right. The information for navigating from point A to B can be represented by exactly one bit of information. Suppose an extra step is added to this example, and the person has to make two consecutive correct decisions to arrive at his correct destination. In that case, the information of his trajectory could be stored in two bits, and four possible outcomes would be achievable. If yet an extra step were to be added, the person has to navigate towards point D in Figure 4.2, eight possible outcomes would be possible. However, this can be stored in just three bits. From this toy example, a key aspect can be derived: a bit of information is exactly the amount of information that is required to choose between two equally probable alternatives (Stone, 2015).

### 4.2.2 Shannon's entropy

When a coin is flipped, it is supposed to land heads up 50% of the time. In other words, the probability of observing the event of a coin landing on tails is  $p = 0.5$ . Whether the coin lands heads up or tails up, the amount of surprise is equal. However, if an unfair coin that lands heads up 90% of the time lands tails up, the amount of surprise is larger. Thus, the surprise of observing a certain event is inversely proportional with its probability of occurring and can be defined as  $1/p(x)$ . In order to express this in bits, the surprise is defined as the logarithm to the base 2 of  $1/p(x)$ :

$$h(x) = \log_2 \frac{1}{p(x)} . \quad (4.1)$$

The Shannon information of a particular outcome  $x$  is denoted by  $h(x)$  and measured in bits. Equation 4.1 can be rewritten and results in:

$$h(x) = -\log_2 p(x) . \quad (4.2)$$

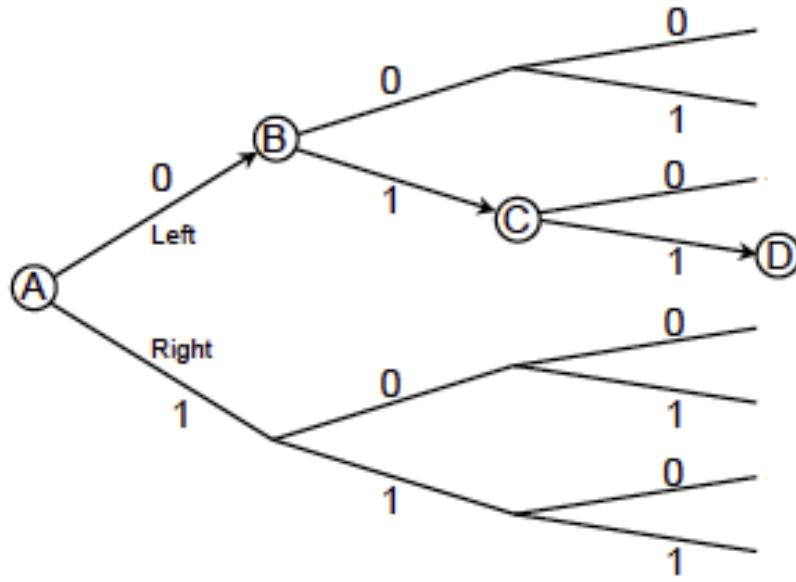


Figure 4.2: In order to arrive at the targeted location, a decision has to be made at every crossroads. The information required for making the correct decision at every decision point requires exactly one bit of information, which can be expressed as either a 1 or a 0. A bit can be interpreted as the amount of information required to make the correct choice between two equally probable options (Stone, 2015).

In order to calculate the surprise of a given outcome, its probability has to be known. For each possible outcome of the variable  $X$ , the corresponding probability can be defined to build the probability distribution  $p(X)$ . Not only does this allow for the calculation of the surprise in each observation, but it also allows for the calculation of the average amount of surprise in the entire set of possible outcomes of the variable  $X$ . Shannon called this average amount of surprise in a variable the entropy of  $p(X)$  defined as  $H(X)$ . Where  $H(X)$  is the average of the Shannon information in a series of  $n$  observed outcomes of  $X$ , given by:

$$H(X) = E[h(x)] . \quad (4.3)$$

The number of times a certain event will be directly proportional to its probability, which allows to rewrite 4.3 to:

$$\begin{aligned} H(X) &= \sum_{i=1}^n p(x_i) \log_2 \frac{1}{p(x_i)} \\ &= -\sum_{i=1}^n p(x_i) \log_2 p(x_i) . \end{aligned} \quad (4.4)$$

This formula calculates the average amount of surprise or entropy for a given variable  $X$ , given the probability distribution over that variable  $p(X)$ . Suppose  $X$  is the observation of a fair coin landing on either tail or heads up. The corresponding probability distribution is given by  $p(X) = (0.5, 0.5)$ . By using Equation 4.4 the entropy of a coin flip is calculated to be:

---


$$\begin{aligned} H(X) &= -0.5 \log_2(0.5) - 0.5 \log_2(0.5) \\ &= 1. \end{aligned} \tag{4.5}$$

Suppose a biased or unfair coin is flipped, which lands heads up with a probability of 0.9. The average surprise when flipping this coin would equate to:

$$\begin{aligned} H(X) &= -0.9 \log_2(0.9) - 0.1 \log_2(0.1) \\ &= 0.469. \end{aligned} \tag{4.6}$$

The average uncertainty of a biased coin is significantly lower than that of an unbiased one. The information when flipping a coin is maximised when both outcomes are equally probable. Suppose a coin that always lands heads up, there would be no extra information gathered when actually flipping the coin since the outcome is already determined. This equates to:

$$\begin{aligned} H(X) &= -1 \log_2(1) - 0 \log_2(0) \\ &= 0, \end{aligned} \tag{4.7}$$

where according to L'Hopital's rule,  $-0 \log(0)$  equals zero:

$$\begin{aligned} \lim_{x \rightarrow 0} x \log(x) &= \lim_{x \rightarrow 0} \frac{\log(x)}{1/x} \\ &\stackrel{\text{Hopital}}{\rightarrow} \lim_{x \rightarrow 0} \frac{-x^2}{x} \\ &= 0. \end{aligned} \tag{4.8}$$

The average uncertainty of a variable  $X$  is summarised by its entropy  $H(X)$ . Every observation of  $X$  provides, on average, an amount of information equal to its entropy (Stone, 2015).

### Mutual information and conditional entropy

The mutual information of two variables is a general measure of association between them (Stone, 2015). Given two variables,  $X$  and  $Y$ , the mutual information  $I(X, Y)$  is the average amount of information gained about  $Y$  after observation of  $X$ . Since mutual information is a symmetric property, this is also true the other way around. In other words,  $I(X, Y)$  is the average reduction in uncertainty or  $H(Y)$ , when the outcome of  $X$  is known. The mutual information  $I(X, Y)$  equals:

$$I(X; Y) = H(X) + H(Y) - H(X, Y). \tag{4.9}$$

where  $H(X, Y)$  is the joint entropy of the two variables,

$$H(X, Y) = -\sum_i \sum_j P(x_i, y_j) \log_2 P(x_i, y_j). \tag{4.10}$$

The joint entropy reaches a maximum when there is no relation between the variables  $X$  and  $Y$  and they are statistically independent. Since an observation of  $X$  conveys no information about  $Y$  their mutual information equals zero.

Consequently, the uncertainty left in  $Y$  once a value of  $X$  is given is the conditional entropy  $H(Y|X)$  and equates to:

$$H(Y|X) = H(Y) - I(X; Y). \quad (4.11)$$

If  $X$  and  $Y$  are fully dependent, and  $X$  is known, there is no uncertainty left in  $Y$  and  $H(Y|X) = 0$ .

Suppose two consecutive fair coin flips. Where  $p(X) = p(Y) = \{p(\text{heads}), p(\text{tails})\} = \{0.5, 0.5\}$ . Using Equation 4.4, the entropy or uncertainty  $H(X) = H(Y) = 1$  bits. Moreover, the joint probability distribution is given by:

$$P(X, Y) = \{p(x_1, y_1), p(x_1, y_2), p(x_2, y_1), p(x_2, y_2)\}. \quad (4.12)$$

Since two consecutive coinflips are statistically independent this equates to:

$$\begin{aligned} P(X, Y) &= \{p(x_1)p(y_1), p(x_1)p(y_2), p(x_2)p(y_1), p(x_2)p(y_2)\} \\ &= \{0.25, 0.25, 0.25, 0.25\}. \end{aligned} \quad (4.13)$$

Therefore, the mutual information of two consecutive coinflips equates to

$$I(X, Y) = H(X) + H(Y) - H(X, Y) = 0, \quad (4.14)$$

and

$$H(Y|X) = H(Y) - I(X, Y) = 1. \quad (4.15)$$

Since an observation of  $X$  conveys no information about  $Y$ , there is no information shared between both variables and the conditional entropy of  $Y$ , given  $X$  is equal to the entropy in  $Y$ .

Suppose a fair coin, but for whatever reason, the second flip is three times more likely to land on the same side as the first. Under these conditions, the entropy in  $X$  and  $Y$  is still equal to 1 bit, as the chances of landing heads or tails are unchanged. However, the joint distribution has changed and is given by:

$$\begin{aligned} P(X, Y) &= \{p(x_1)p(y_1), p(x_1)p(y_2), p(x_2)p(y_1), p(x_2)p(y_2)\} \\ &= \{0.5 \times 0.75, 0.5 \times 0.25, 0.5 \times 0.25, 0.5 \times 0.75\} \\ &= \{0.375, 0.125, 0.125, 0.375\}. \end{aligned} \quad (4.16)$$

And  $H(X, Y) = 1.81$  bits. Therefore:

$$I(X, Y) = H(X) + H(Y) - H(X, Y) = 0.19 \text{ bits} \quad (4.17)$$

and  $H(Y|X) = H(Y) - I(X, Y) = 0.81$  bits. Mutual information is metric for quantifying how much information is shared by variables. When two variables are highly associated, they will have a large amount of shared information. When an observation of one of the two is known, the average amount

of uncertainty left in the other will be lowered. Figure 4.3 gives a schematic summary of how entropy, joint entropy, conditional entropy and mutual information relate to each other.

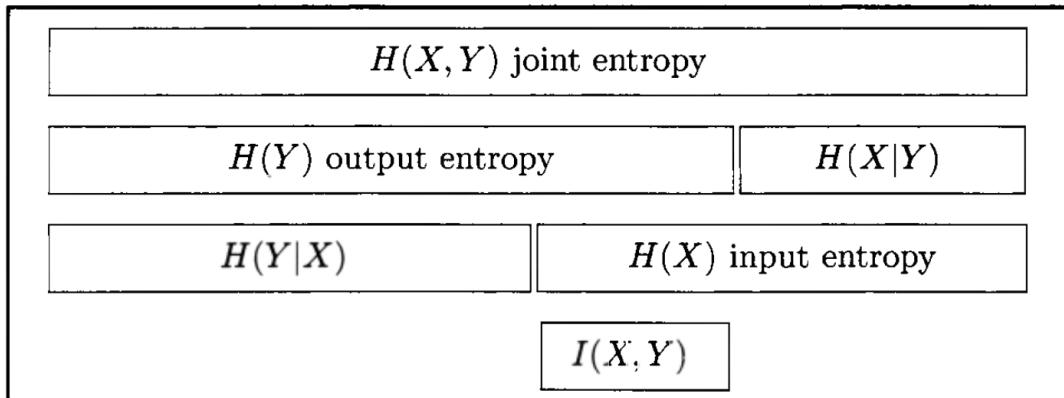


Figure 4.3: The relationship between input entropy  $H(X)$ , output entropy  $H(Y)$ , joint entropy  $H(X, Y)$ , mutual information  $I(X, Y)$  and conditional entropy  $H(X|Y)$  and  $H(Y|X)$  (Stone, 2015).

### 4.3 Quantifying individuality

Krakauer et al. (2020) propose a mathematical framework based on information theory in order to determine the amount of individuality present in a system. The main criterion for individuality is the capability to propagate information from the past to the future over time. Biological systems are constantly interacting with their environment. For simplicity, both the system and environment are described as discrete states, where the next state of the system depends on its previous state and the state of the environment. Similar to the system, the environment is not static, and its state is dependent on its previous state and the systems it interacts with. Modelling the interactions between a system and its environment requires two channels. Both the system and its environment can be represented as a time series. At every given point, the current state of both the system and its environment depends on the system's previous state and environment. Figure 4.4 gives a schematic overview of this abstraction.

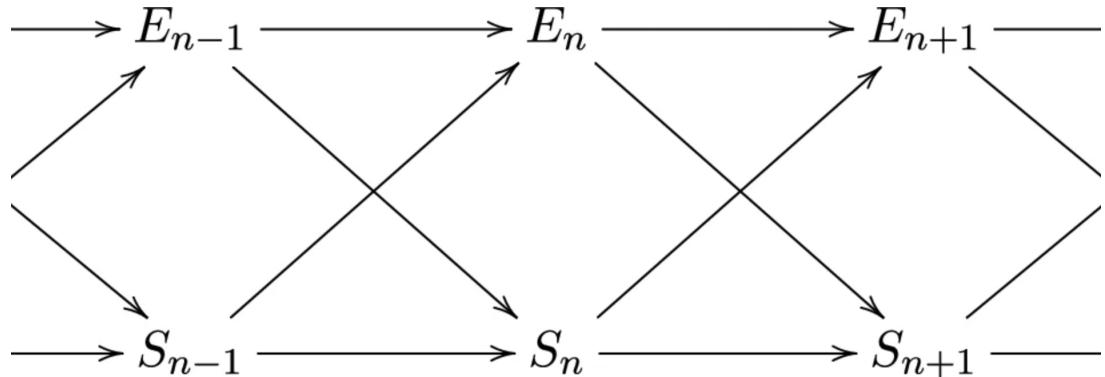


Figure 4.4: A schematic representation of the system-environment abstraction. At any given point in time, the current state of the system ( $S_n$ ) and environment ( $E_n$ ) depends on the state of both the system ( $S_{n-1}$  and its environment ( $E_{n-1}$ ) in the previous step (Krakauer et al., 2020).

The next state of a system is dependent on its previous state and on the state of the environment. Therefore, the probability of observing the next state of a system  $s'$ , given its previous state  $s$  and that of the environment  $e$  is given by  $\phi(e, s; s')$ . Similarly the probability for observing the next state of the environment is given by  $\psi(s, e; e')$ , where:

$$\sum_{i=1}^n \phi(s, e; s'_i) = 1, \quad (4.18)$$

and

$$\sum_{i=1}^m \psi(s, e; e'_i) = 1, \quad (4.19)$$

for all  $s$  and  $e$ , where  $n$  and  $m$  denote the number of possible states of the system and environment, respectively.

This dynamic process can be modelled through a Markov kernel, where the transition probabilities are given by  $\phi$  and  $\psi$ . This abstraction allows for the quantification of information flows between the two variables.

The next step consists of formulating how patterns in information flow between a system and its environment define the system's relevant degree of individuality. A biological individual can be usefully understood as an informational individual Krakauer et al. (2020), where typical ideas of individuality do not always apply.

First of all, the system en environment decomposition must be justifiable. Biology is hierarchically structured, making it hard to distinguish between a system and its environment. When studying organelles in a cell, the organelles could be considered the system and the cell as their environment. However, When studying cells in the human body, the body becomes the environment and the cells, the individuals. Next, the formalisation should be generally applicable. The system's state at any given point in time can be predicted by its previous state and that of the environment. The degree of predictability of the next system's state is given by the mutual information shared between the previous state of the system and the environment and the next state of the system. Using Equation 4.9 this equates to:

$$I(S_n, E_n; S_{n+1}) = H(S_{n+1}) - H(S_{n+1}|S_n, E_n). \quad (4.20)$$

This expression can be decomposed as:

$$\begin{aligned} I(S_n, E_n; S_{n+1}) &= I(S_{n+1}; S_n) + I(S_{n+1}; E_n|S_n) \\ &= I(S_{n+1}; E_n) + I(S_{n+1}; S_n|E_n). \end{aligned} \quad (4.21)$$

The mutual information is a measure of how much information is shared between the next state of a system and its previous state and the previous state of the environment.

The first term in Equation 4.21,  $I(S_{n+1}; S_n)$ , denotes the amount of information that a system shares with itself. In other words, it measures the amount of influence the system state has when transitioning from one state to the next. High values of  $I(S_{n+1}; S_n)$  suggest that the system is highly in control of

---

its next stage. In literature, this metric is called autonomy and is denoted as  $A^*$  (Krakauer and Zanotto, 2013).

The second term,  $I(S_{n+1}; E_n|S_n)$ , measures the amount of influence the environment has over the next state of the system, i.e. the information flowing from the environment into the system.

If  $I(S_{n+1}; E_n) + I(S_{n+1}; S_n|E_n)$  is considered, the first term denotes the influences of the environment on the next state of the system. The remaining influence, or the second term, is the influence the system has on its next state, given its environment in the previous and is valid when all dependencies between the states of the system are fully a consequence of its environment and are denoted by  $A$  (Krakauer et al., 2020). These three types of information flows can be summarised as follows:

- *colonial individuality*:  $A = I(S_{n+1}; S_n|E_n)$ ;
- *organismal individuality*:  $A^* = I(S_{n+1}; S_n)$ ;
- *environmental determined*: individuality  $nC = I(S_{n+1}; E_n|S_N)$ .

By introducing the concepts of shared, complementary and unique information as proposed by Williams and Beer (2010), can be decomposed as:

$$I(S_{n+1}; S_n, E_n) = \underbrace{SI(S_{n+1}; S_n, E_n)}_{\text{Shared}} + \underbrace{CI(S_{n+1}; S_n, E_n)}_{\text{Complementary}} + \underbrace{UI(S_{n+1}; S_n \setminus E_n)}_{\text{Unique to } S_n} + \underbrace{UI(S_{n+1}; E_n \setminus S_n)}_{\text{Unique to } E_n}. \quad (4.22)$$

These four terms appear in the pairwise and mutual information obtained by the chain rule:

$$\begin{aligned} I(S_{n+1}; S_n) &= SI(S_{n+1}; S_n, E_n) + UI(S_{n+1}; S_n \setminus E_n); \\ I(S_{n+1}; E_n|S_n) &= CI(S_{n+1}; S_n, E_n) + UI(S_{n+1}; E_n \setminus S_n); \\ I(S_{n+1}; E_n) &= SI(S_{n+1}; S_n, E_n) + UI(S_{n+1}; E_n \setminus S_n); \\ I(S_{n+1}; S_n|E_n) &= CI(S_{n+1}; S_n, E_n) + UI(S_{n+1}; S_n \setminus E_n). \end{aligned} \quad (4.23)$$

These terms can be interpreted as:

- $SI(S_{n+1}; S_n, E_n)$  the shared information between the system and environment;
- $CI(S_{n+1}; S_n, E_n)$  the complementary or synergistic information present in the interactions between the system and environment;
- $UI(S_{n+1}; S_n \setminus E_n)$  the unique information, maintained by the system;
- $UI(S_{n+1}; E_n \setminus S_n)$  the unique information from the environment, influencing the system.

### 4.3.1 Types of individuality

The previous decomposition allows for a more precise definition of the three different types of individuality, measured by shared information between the system and environment and information that is either unique to the system or environment and information in interactions between them.

#### Organismal individuality $A^*$

$$A^* = SI(S_{n+1}; S_n, E_n) + UI(S_{n+1}; S_n \setminus E_n). \quad (4.24)$$

The organismal individuality  $A^*$  reflects the amount of information that the system shares with its previous states and that of the environment. The second term measures the amount of information shared solely within the system. Since the informational individual is characterised by its ability to propagate information from its past to its future, this metric should be maximised for systems with a high amount of organismal individuality.

#### Colonial individuality $A$

$$A = CI(S_{n+1}; S_n, E_n) + UI(S_{n+1}; S_n \setminus E_n). \quad (4.25)$$

Bacteria in a biofilm, for example, share little to no information with their environment. However, the biofilm as a whole is constantly adapting to its environment. This metric quantifies the colonial individuality of an aggregate regulated by its environment.

#### Environmental determination $nC$

$$\begin{aligned} nC &= I(S_{n+1}; E_n | S_n) \\ &= CI(S_{n+1}; S_n, E_n) + UI(S_{n+1}; E_n \setminus S_n). \end{aligned} \quad (4.26)$$

This metric aims to quantify the environment's influence on the system's future state. Low measures of environmental determination suggest that the system is rather insensitive to the environment. In this case, the system cannot adapt to its environment and thus does not possess a high degree of organismal or colonial individuality.

#### Environmental encoding $NTIC$

$$NTIC = SI(S_{n+1}; S_n, E_n) - CI(S_{n+1}; S_n, E_n). \quad (4.27)$$

This metric quantifies the difference between colonial and organismal individuality. It can be seen as how much information is encoded about the environment within the system minus the amount of information that flows through system-environment interactions. When this measure is large, the system largely depends on its environment. As this measure decreases, the system largely controls itself.

---

## 4.4 Individuality of aggregating *Dictyostelium* cells

Conventional criteria for determining individuality fail to unambiguously describe the individuality of *Dictyostelium discoideum* aggregates. Section 3 discusses the remarkable life cycle of *Dictyostelium* cells. The unicellular amoebae combine to form a multicellular aggregate under starvation stress during its life cycle. This multicellular entity differs from other forms of multicellularity since it did not emerge from mitotic cell division or cloning but rather by the chemotactic gathering of a group of physically distinct cells under the control of cAMP. As a result, the pseudoplasmodium or slug is considered a genetic chimaera. During the culmination, stalk cells voluntarily go into apoptosis, sacrificing themselves for the cause of survival and reproduction, suggesting that some individuality can be attributed to the slug rather than the unicellular amoebae. This behaviour is reminiscent of the altruism described in social insects. In case of danger, worker bees will defend their colony by stinging possible intruders, dying in the process. This behaviour goes against the Darwinian approach of survival of the fittest. Worker bees lower their fitness in order for the colony to survive. Several explanations, such as that of *kin selection* ought to describe this phenomenon (West-Eberhard, 1975). However, today, assigning individuality to either the unicellular amoebae or the whole migratory slug has remained an arbitrary choice.

### 4.4.1 Abstraction

Under normal circumstances, *Dictyostelium discoideum* amoebae dwell in the soil preying on soil bacteria. In this unstimulated state, *Dictyostelium* cells continuously change their morphology by extending and retracting their pseudopods in all directions, resulting in a random walk (Soll et al., 2003). Under starvation stress, the unicellular *Dictyostelium* cells become sensitive to chemoattractants such as cAMP. In the presence of a cAMP gradient, cells become polarised. They actively start extending and retracting their pseudopods in the direction of rising cAMP while suppressing the formation of pseudopods on the lateral sides (Weijer, 2004). This mechanism results in guided locomotion up the cAMP gradient. At any given point in time, the movement of *Dictyostelium* cells can either be random or directed up the cAMP gradient.

Suppose a set-up of unicellular *Dictyostelium* amoebae randomly placed in a circular Petri dish with radius  $r_{dish} = 1$ . A small amount of cAMP solution is delivered by a pump in the centre of the dish. Dropping small amounts of cAMP in the centre of the dish will cause the formation of a concentric gradient due to diffusivity. The edge of the gradient is considered the lowest concentration of cAMP that *Dictyostelium* cells can generally cover. Therefore, the area of the dish, covered by the gradient  $\theta$  equates to  $\theta = S_{gradient}/S_{dish} = r_{gradient}^2/r_{dish}^2 = r_{gradient}^2$ .

### Environment

Since a cell is either in a cAMP gradient or not, the environment can be encoded as a binary state, where  $E_n = \{\text{no cAMP gradient, cAMP gradient}\}$ , or  $E_n = \{-1, +1\}$ . As a result, when an amoeba is randomly placed in the Petri dish, the probability of landing in the gradient is given by  $\theta$  or  $p(E_n =$

$+1) = \theta$ . Since  $\sum_{i=1}^n p(E_N) = 1$ , the probability of landing in an area with no cAMP gradient should be  $p(E_N = -1) = 1 - \theta$ . Hence, the probability distribution of the environment is given by  $p(E_N) = \{1 - \theta, \theta\}$ .

### System

Under normal vegetative circumstances, *Dictyostelium* cells perform a random walk. Under starvation stress, the cells become sensitive to low cAMP concentrations and sense a gradient over their plasma membrane. Small differences in cAMP over the plasma membrane lead to a directed motion up the cAMP gradient. Consider the state of the system. In this case, the *Dictyostelium* cell is either random walking or directed motion. Thus,  $S_N = \{\text{random, directed}\}$  or  $S_N = \{-1, +1\}$ .

Defining both the environment and the system as binary units, results in a total of four distinct states:  $\{+1, +1\}$ ,  $\{-1, +1\}$ ,  $\{+1, -1\}$  and  $\{-1, -1\}$ . Both channels  $S_n$  and  $E_n$  are synchronously updated according to the schema depicted in Figure 4.4. Since the next state of the system and the environment are only dependant of the previous states of the system and environment, therefore the conditional distribution is given by  $p(S_{n+1}, E_{n+1}|S_n, E_n)$ . Since both channels are synchronously updated, the predictions of  $E_{n+1}$  and  $S_{n+1}$  are only dependant on previous states of the model, the conditional transformation distribution can be rewritten as:

$$p(s_{n+1}, e_{n+1}|s_n, e_n) = p_s(s_{n+1}|s_n, e_n) \times p_e(e_{n+1}|s_n, e_n). \quad (4.28)$$

Where  $p_s(s_{n+1} = +1|s_n, e_n)$  is given by the vector  $\{a_S, b_S, c_S, d_S\}$ , resulting in:

$$\begin{aligned} p_s(+1|+1, +1) &= a_S \\ p_s(+1|-1, +1) &= b_S \\ p_s(+1|+1, -1) &= c_S \\ p_s(+1|-1, -1) &= d_S. \end{aligned} \quad (4.29)$$

Similarly for  $E_N$ , the conditional probability is given by  $p_e(e_{n+1} = +1|s_n, e_n)$ :

$$\begin{aligned} p_e(+1|+1, +1) &= a_E \\ p_e(+1|-1, +1) &= b_E \\ p_e(+1|+1, -1) &= c_E \\ p_e(+1|-1, -1) &= d_E. \end{aligned} \quad (4.30)$$

$a_S$  Should be interpreted as the probability that, given the cell is located in a cAMP gradient and is actively moving up that gradient at time  $n$   $\{S_n, E_n\} = \{+1, +1\}$ , the cell is active at the next observation  $n + 1$ , or  $\{S_{n+1} = +1\}$ . Correspondingly, the probability of observing  $\{E_{n+1} = +1\}$  from  $\{S_n, E_n\} = \{+1, +1\}$  is given by  $a_E$ . Given that transitions are independent the probability of observing  $\{S_{n+1}, E_{n+1}\} = \{+1, +1\}$  after  $\{S_n, E_n\} = \{+1, +1\}$  equals  $p(+1, +1|+1, +1) = a_S a_E$ . Both channels can be merged into a single Markov chain resulting in the four by four conditional transition matrix given by Table 4.1.

$(S_{n+1}, E_{n+1})$				
	$(+1, +1)$	$(-1, +1)$	$(+1, -1)$	$(-1, -1)$
$(+1, +1)$	$a_S a_E$	$(1 - a_S) a_E$	$a_S (1 - a_E)$	$(1 - a_S)(1 - a_E)$
$(-1, +1)$	$b_S b_E$	$(1 - b_S) b_E$	$b_S (1 - b_E)$	$(1 - b_S)(1 - b_E)$
$(+1, -1)$	$c_S c_E$	$(1 - c_S) c_E$	$c_S (1 - c_E)$	$(1 - c_S)(1 - c_E)$
$(-1, -1)$	$d_S d_E$	$(1 - d_S) d_E$	$d_S (1 - d_E)$	$(1 - d_S)(1 - d_E)$

Table 4.1: The transition matrix  $p(S_{n+1}, E_{n+1}|S_n, E_n)$ . The rows of the matrix denote  $S_n, E_n$ , while  $S_{n+1}, E_{n+1}$  is shown by the columns. The conditional probability of observing a transition from row  $i$  to column  $j$  is given by  $M_{ij}$ .

In order to validate the mathematical framework for quantifying different forms of individuality, values for  $a_S, b_S \dots$  should be determined. This work aims to fill in the values according to their biological meaning. However, since these values are only rationalised guesses, this example serves as a proof of concept, and the obtained results should be interpreted with caution.

A *Dictyostelium* cell, placed in a cAMP gradient  $E_n = +1$  and actively moving up that gradient  $S_n = +1$  is will remain in that state  $\{+1, +1\}$  with a high probability. Therefore, both  $a_S$  and  $a_E$  should be a rather high value.

Cells that are performing a random walk while sensing a cAMP gradient  $\{S_n, E_n\} = \{-1, +1\}$  should become activated and start moving up the gradient with a reasonably high probability, suggesting a relatively high value for  $b_S$ . However, since they only become activated in the next step, they might also lose the cAMP trail, which suggests a moderate value for  $b_E$ . This probability depends on the value  $\theta$  or the surface area covered by the cAMP gradient.

Cells that are moving up a cAMP gradient, but no longer sense the gradient  $\{S_n, E_n\} = \{+1, -1\}$  are likely to become deactivated in the next step, thus low values of  $c_S$  are expected. The probability of a retrieving the cAMP gradient,  $c_E$  are again dependant on  $\theta$ .

In the last case, cells that perform a random walk and do not sense a cAMP gradient are likely to keep performing their random walk. However, they could find a cAMP gradient with a probability relational to  $\theta$  in the next step. Therefore,  $d_S$  should be a low value and,  $d_E$  is relational to  $\theta$ .

The transition probabilities for the environment channel are clearly dependant on the surface, covered by the gradient  $\theta$ . When a cAMP gradient is present in the entire Petri dish  $\theta = 1$ , the transition probability should equate to  $p_e(E_{n+1}|S_n, E_n) = \{1, 1, 1, 1\}$ . This example supposes a linear relation between  $b_E, c_E, d_E$  and  $\theta$  and  $b_E = c_E = d_E = \theta$ . If a cell is actively moving up the cAMP gradient the probability of sensing a cAMP gradient in the next step should be higher than  $\theta$ . By using a logarithmic scale,  $a_S$  is set to  $a_S = \sqrt{\theta}$ , as shown in Figure 4.5.

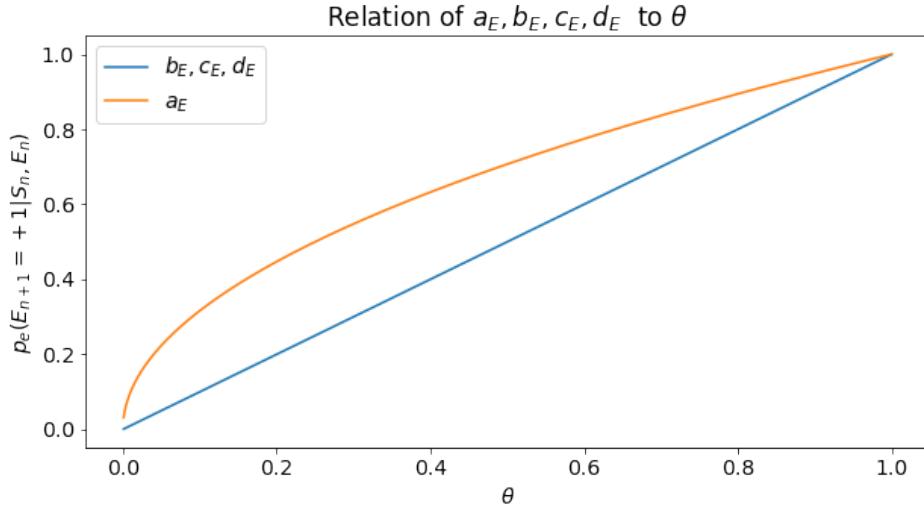


Figure 4.5: The relation between  $p_e(E_{n+1} = +1 | S_n, E_n)$  and  $\theta$ . The probability of sensing the environment while performing a random walk is directly proportional to  $\theta$ . However, when moving up the gradient, the probability of sensing a cAMP gradient in the next step should be higher, and is given by  $a_E = f_s(\theta) = \sqrt{\theta}$ .

$a_S, b_S, c_S$  and  $d_S$  are assigned values according to their biological meaning. Suppose:

$$\begin{aligned}
 p_s(S_{n+1} = +1 | S_n, E_n) &= (a_S, b_S, c_S, d_S) \\
 &= (0.95, 0.5, 0.3, 0.05) \\
 p_e(E_{n+1} = +1 | S_n, E_n) &= (a_E, b_E, c_E, d_E) \\
 &= (\sqrt{\theta}, \theta, \theta, \theta)
 \end{aligned} \tag{4.31}$$

Then, the conditional transition matrix  $p(S_{n+1}, E_{n+1} | S_n, E_n)$  is given by Table 4.2, where  $f_s(\theta) = \sqrt{\theta}$ .

$(S_{n+1}, E_{n+1})$				
	$(+1, +1)$	$(-1, +1)$	$(+1, -1)$	$(-1, -1)$
$(+1, +1)$	$0.95 \times f_s(\theta)$	$0.05 \times f_s(\theta)$	$0.95 \times (1 - f_s(\theta))$	$0.05 \times (1 - f_s(\theta))$
$(-1, +1)$	$0.5 \times \theta$	$0.5 \times \theta$	$0.5 \times (1 - \theta)$	$0.5 \times (1 - \theta)$
$(+1, -1)$	$0.3 \times \theta$	$0.7 \times \theta$	$0.3 \times (1 - \theta)$	$0.7 \times (1 - \theta)$
$(-1, -1)$	$0.05 \times \theta$	$0.95 \times \theta$	$0.05 \times (1 - \theta)$	$0.95 \times (1 - \theta)$

Table 4.2: Transition matrix filled in according to Table 4.1 for  $p_s = (0.95, 0.5, 0.3, 0.05)$  and  $p_e = (\sqrt{\theta}, \theta, \theta, \theta)$ .

#### 4.4.2 Forms of individuality

##### Organismal individuality

The organismal individuality determines at what rate the system shares information about its state to its next state. As discussed in Section 4.3, this equates to  $A^* = I(S_{n+1}; S_n)$ . The shared information

---

between the state of a system can be derived from Equation 4.9 and results in:

$$I(S_{n+1}; S_n) = H(S_{n+1}) + H(S_n) - H(S_{n+1}, S_n). \quad (4.32)$$

Where  $H(S_{n+1})$  denotes the uncertainty in observing either  $S_{n+1} = +1$  or  $S_{n+1} = -1$ .

For  $\theta \in ]0, 1[$  the Markov chain is irreducible and aperiodic, therefore a unique steady-state distribution can be calculated for the transition matrix where:

$$\lim_{k \rightarrow \infty} \mathbf{P}^k = \boldsymbol{\pi}. \quad (4.33)$$

$\boldsymbol{\pi}$  Equates to the stationary distribution of  $\{S_n, E_n\}$ . Where  $p(\{S_n, E_n\})$  denotes the probability of observing each individual state. Therefore:

$$\begin{aligned} p(S_n) &= (p(S_n = -1), p(S_n = +1)) \\ &= (p(-1, +1) + p(-1, -1), p(+1, +1) + p(+1, -1)). \end{aligned} \quad (4.34)$$

Since  $\boldsymbol{\pi}$  gives the steady-state distribution of the Markov chain, the probability of observing each state becomes independent of time, thus  $p(S_n) = p(S_{n+1})$  and  $H(S_n) = H(S_{n+1})$ . This allows to calculate the average uncertainty in the system's state. Figure 4.6 shows the relation between  $p(S_n)$ ,  $H(S_n)$  and  $\theta$ .

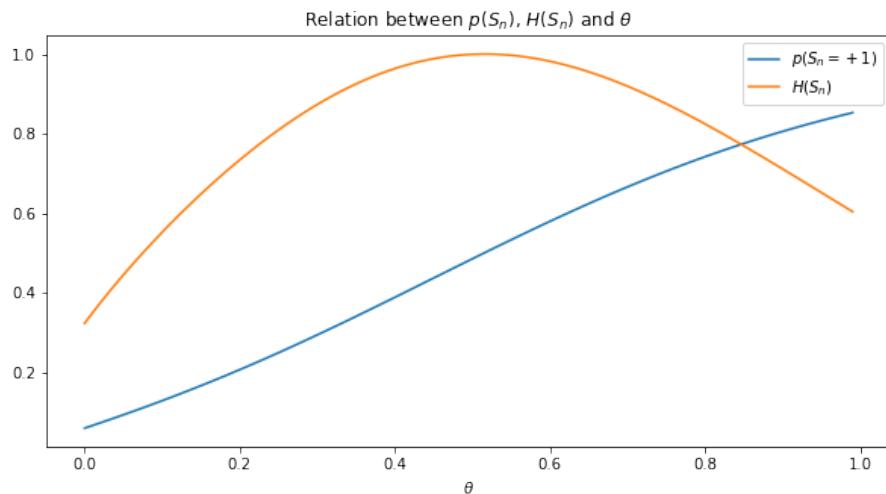


Figure 4.6: The relation between  $p(S_n)$ ,  $H(S_n)$  and  $\theta$ . As the surface of the cAMP gradient increases, cells are more likely to be in an active state, and move up the gradient. The uncertainty  $H(S_n)$  reaches the maximum value of 1 bit, when the uncertainty in  $S_n$  is maximised, or  $p(S_n = \pm 1) = 0.5$ .

In order to calculate the amount of information a system shares with its next state, the joint distribution  $p(S_{n+1}, S_n)$  needs to be defined. Since  $S_n$  is defined as a binary variable, the joint distribution is denoted by a two-by-two matrix. Where  $p(S_{n+1}, S_n)$  denotes the probability of each transition occurring. The joint distribution can be retrieved from the overall transition matrix in Table 4.1. Where

$p(S_{n+1} = +1, S_n = +1)$  is calculated according to:

$$\begin{aligned}
 p(S_{n+1} = +1, S_n = +1) &= p(\{+1, +1\} \rightarrow \{+1, +1\}) \\
 &\quad + p(\{+1, +1\} \rightarrow \{+1, -1\}) \\
 &\quad + p(\{+1, -1\} \rightarrow \{+1, +1\}) \\
 &\quad + p(\{+1, -1\} \rightarrow \{+1, -1\}).
 \end{aligned} \tag{4.35}$$

Note that, the transition matrix in Table 4.1 denotes the conditional probability distribution. Therefore  $p(\{+1, +1\} \rightarrow \{+1, +1\}) \neq a_S a_E$ . According to Bayes' theorem, absolute probability for each transition is obtained by multiplying each row of the transition matrix with its corresponding eigenvalue. As a result,  $p(\{+1, +1\} \rightarrow \{+1, +1\}) = a_S a_E \pi_1$ .

This allows for the state transition matrix  $p(S_n \rightarrow S_{n+1})$  to be built. Note that the elements of the transition matrix denote the absolute probability of a transition, and not the conditional probability  $p(S_{n+1}|S_n)$ . Since both  $p(S_n)$  and  $p(S_{n+1})$  as well as  $p(S_n, S_{n+1})$  are known, the mutual information  $I(S_n; S_{n+1})$  can be retrieved according to Equation 4.9.

The graph in Figure 4.7 shows the calculation for different values of  $\theta$ . The mutual information reaches a maximum of approximately 0.146 bits for  $\theta = 0.6$ . In other words, the uncertainty of  $S_{n+1}$  is on average reduced by 15 percent, when  $S_n$  is given.

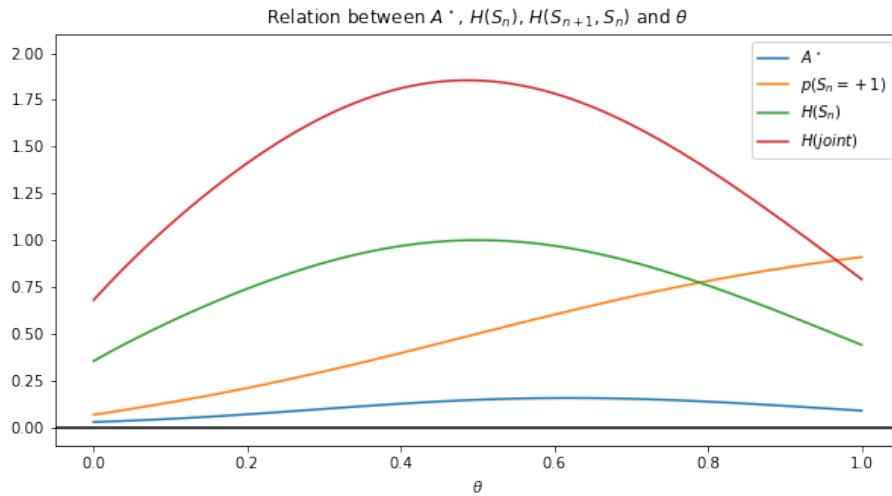


Figure 4.7: The relation between  $p(S_n)$ ,  $H(S_n)$ ,  $H(S_{n+1}, S_n)$ ,  $I(S_{n+1}, S_n)$  and  $\theta$ , for  $(a_S, b_S, c_S, d_S) = (0.95, 0.5, 0.3, 0.05)$  and  $(a_E, b_E, c_E, d_E) = (f_s(\theta), \theta, \theta, \theta)$ . The mutual information reaches a maximum value of 0.146 bits, for  $\theta \approx 0.6$ . At these parameters, the uncertainty of  $S_{n+1}$  is on average, reduced by 0.146 bits, if  $S_n$  is given.

### Colonial individuality

The colonial individuality of a system  $A$  equates to:

$$A = I(S_{n+1}; S_n | E_n). \tag{4.36}$$

---

A Aims to quantify the information that is passed through from the current state of the system, given its environment  $p(S_n|E_n)$  to the system's state in the next step  $S_{n+1}$ .

According to Mackay (1995),  $I(S_{n+1}; S_n|E_n)$  can be decomposed as:

$$I(S_{n+1}; S_n|E_n) = H(S_{n+1}|E_n) - H(S_{n+1}|S_n, E_n). \quad (4.37)$$

This can be composed using the relations:

$$\begin{aligned} I(X; Y|Z) &= H(X|Z) - H(X|Y, Z); \\ H(X|Z) &= H(X, Z) - H(Z); \\ H(X|Y, Z) &= H(Y, X|Z) - H(Y|Z); \\ I(X; Y|Z) &= H(X|Z) - H(X|Y, Z) \\ &= H(X|Z) - H(Y, X|Z) + H(Y|Z) \\ &= H(X, Z) - H(Z) - H(Y, X, Z) + H(Z) + H(Y, Z) - H(Z) \\ &= H(X, Z) + H(Y, Z) - H(Y, X, Z) - H(Z). \end{aligned} \quad (4.38)$$

Therefore the colonial individuality can be decomposed as:

$$I(S_{n+1}; S_n|E_n) = H(S_{n+1}, E_n) + H(S_n, E_n) - H(S_{n+1}, S_n, E_n) - H(E_n). \quad (4.39)$$

For different values for  $\theta$  the colonial individuality  $A$  is given in Figure 4.8. The colonial individuality of the system reaches a maximum of 0.165 bits for  $\theta \approx 0.6$  for the parameters of  $p_S$  and  $p_E$  according to Table 4.2. In other words, the average uncertainty in  $S_{n+1}$  is reduced by 0.166 bits when  $p(S_n|E_n)$  is known.

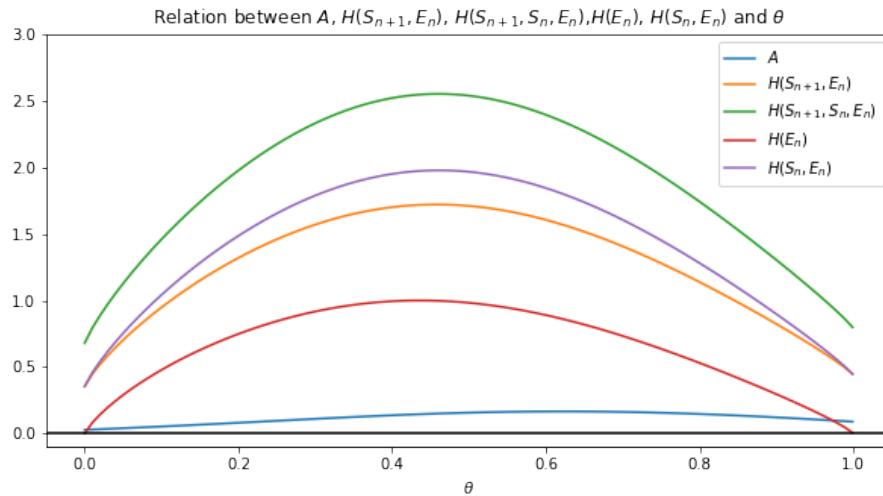


Figure 4.8: Relation between  $A$ ,  $H(S_{n+1}, E_n)$ ,  $H(S_{n+1}, S_n, E_n)$ ,  $H(E_n)$ ,  $H(S_n, E_n)$  and  $\theta$ , for  $(a_S, b_S, c_S, d_S) = (0.95, 0.5, 0.3, 0.05)$  and  $(a_E, b_E, c_E, d_E) = (f_s(\theta), \theta, \theta, \theta)$ . The information, passed from the system's state to its, next state, given its environment reaches a maximum of 0.166 bits, when  $\theta \approx 0.6$ .

### Environmental determined individuality

The environmental determined individuality  $nC$  denotes the average amount of shared information between a system's next state and its current environment, given its current state, or:

$$nC = I(S_{n+1}; E_n | S_n). \quad (4.40)$$

$nC$  Quantifies the amount of information that flows from the environment to a system and measures how the system is influenced by its environment. When  $nC = 0$ , the system is informationally closed from its environment. Therefore, it loses its adaptability to its environment, which is a critical feature for both a colonial and organismal individual.

According to Equation 4.38,  $nC$  can be decomposed as:

$$\begin{aligned} nC &= I(S_{n+1}; E_n | S_n) \\ &= H(S_{n+1} | S_n) - H(S_{n+1} | E_n, Z) \\ &= H(S_{n+1}, S_n) + H(E_n, S_n) - H(S_{n+1}, S_n, E_n) - H(S_n). \end{aligned} \quad (4.41)$$

The metrics of  $nC$  in function of  $\theta$  is given by Figure 4.9. The metric is maximised for  $\theta \approx 0.57$  and reaches a maximal value of 0.80 bits. The average amount of uncertainty in  $p(S_{n+1})$  is, on average, reduced by 0.80 bits when the environment, given the previous state, is provided.

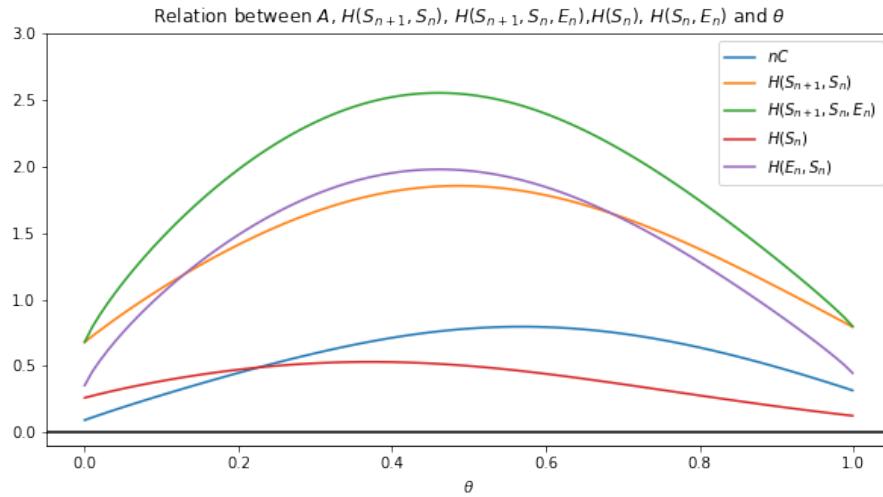


Figure 4.9: Relation between  $nC$ ,  $H(S_{n+1}, S_n)$ ,  $H(S_{n+1}, S_n, E_n)$ ,  $H(S_n)$ ,  $H(S_n, E_n)$  and  $\theta$ , for  $(a_S, b_S, c_S, d_S) = (0.95, 0.5, 0.3, 0.05)$  and  $(a_E, b_E, c_E, d_E) = (f_s(\theta), \theta, \theta, \theta)$ . The information, passed from the state of the environment to the next system's state, given its current state reaches a maximum of 0.80 bits, when  $\theta \approx 0.57$ .  $nC$  is a metric, which denotes the amount of influence a state has on its environment

---

## Summary

Information theory allows the application of the mathematical framework for quantifying different forms of individuality. Figure 4.10 shows the calculations for the organismal  $A^*$ , colonial  $A$  and environmental determined  $nC$  individuality for different values of  $\theta$ . Since this is an illustrative example with simulated data, one should be cautious when making assumptions about the obtained results. High values for  $nC$  suggest that the *Dictyostelium* cells in this example are strongly influenced by their environment. Therefore, the observation of  $p(E_n|S_n)$  shares a large amount of information about  $p(S_{n+1})$  and increases predictions about  $p(S_{n+1})$ .

The organismal and colonial individuality is fairly lower than their environmental counterpart, suggesting that  $S_n$  and  $p(S_n|E_n)$  have less influence over the system. The similar distribution of  $A$  and  $A^*$ , suggest that  $p(S_n)$  and  $p(S_n|E_n)$  are similar; again, since this is an illustrative example, assumptions should be made with caution.

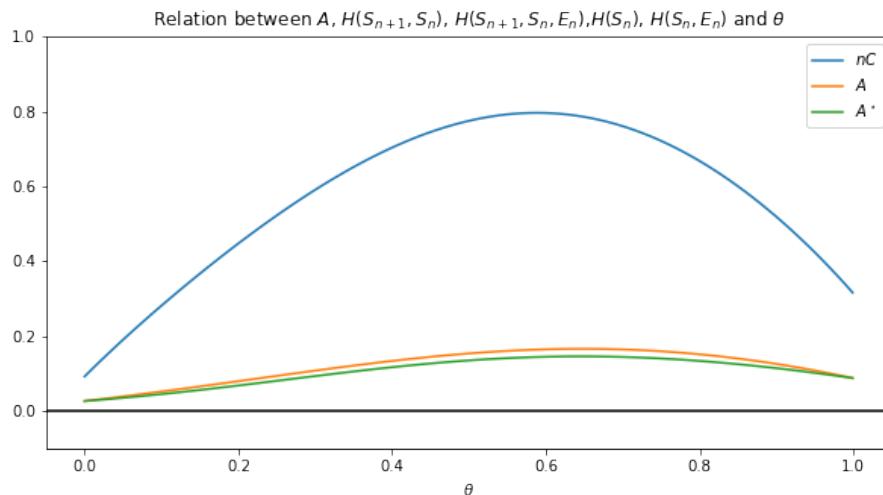


Figure 4.10: The three different forms of individuality for different values of  $\theta$ . For  $(a_S, b_S, c_S, d_S) = (0.95, 0.5, 0.3, 0.05)$  and  $(a_E, b_E, c_E, d_E) = (f_s(\theta), \theta, \theta, \theta)$ , the organismal  $A^*$  and colonial  $A$  individuality are similar and reach a maximum at  $\theta = 0.6$ . The environmental determined individuality  $nC$  is higher than both  $A$  and  $A^*$ , suggesting that the system is influenced by its environment to a larger extend than its previous state.

## **5. CONCLUSION**

In the future, biobots will become valuable instruments for various real-world applications. Despite being a relatively new field of research, it is evolving at a rapid pace. The first proof-of-concept was developed by Kriegman et al. (2020) at the University of Vermont. As of today, two years later, this group has already developed three generations of xenobots, where the third is already capable of kinematic self-replication. Currently, xenobots mainly serve as a scientific tool. Xenobots are the first organisms whose evolutionary history completely occurred in a virtual world under the control of man-made algorithms, allowing scientists to create new organisms from scratch. By doing so, the effect of each perturbation can be assessed, making biobots a valuable complementary tool for gaining additional information in a variety of research fields, including embryology, tissue engineering, evolutionary biology, and regenerative biomedicine, among others.

Microrobots show potential both for medicinal and environmental applications. As discussed in this thesis, miniaturising traditional rigid robots comes with challenges such as damage resistance and the ability to perform in highly unpredictable environments like the human body. Biobots, entirely made from living cells, embody control mechanisms into their materials and are less prone to damage. Additionally, biological life is able to self-organise into highly regulated tissue patterns, sidestepping hurdles involved in the manufacturing of large quantities of microrobots. Biobots could be employed for targeted drug delivery, abrasion of deposits in arteries or as biosensors for early disease detection and treatment. Environmental applications include the sensing and degradation of pollutants, as well as piling up microplastics in the oceans or even degrading them. Due to their biological nature, they are fully biodegradable and do not introduce any additional pollution.

The slime mould *Dictyostelium discoideum* emerges as an interesting donor organism and model system for the development and manufacturing of biobots. They are easily cultured, and methods of genetically engineering *Dictyostelium* have already been developed. At the unicellular stage, they can aggregate under specific conditions and sort out in tightly regulated tissue patterns during and after the aggregation. These features would enable scientists to scale up the production of biobots significantly. Lastly, the multicellular organism is natively sensitive to different chemical compounds, a valuable feature for environmental applications.

Using living, biological materials to build microrobots also comes with disadvantages. Conventional robots can be described as the sum of all their parts, whereas life cannot. By exploiting the adaptability and plasticity in nature, scientists have to compromise in control. In contrast to physics and chemistry, biology often does not rely on fundamental laws backed by mathematics. The final part of this thesis describes a mathematical framework applicable to both *Dictyostelium* and biobots. By inventing new forms, previously defined concepts, such as the concept of individuality, need to be revised.

---

To conclude this thesis: biobots show promising potential for a variety of different applications. However, in order to fully break through as a useful tool, a lot of research in the fields involved in the scalable development and manufacturing of biobots still has to be conducted.

## BIBLIOGRAPHY

- Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., and Walter, P. (2015). *Molecular Biology of THE CELL*. Garland Science.
- Alcântara, C. C., Kim, S., Lee, S., Jang, B., Thakolkaran, P., Kim, J. Y., Choi, H., Nelson, B. J., and Pané, S. (2019). 3d fabrication of fully iron magnetic microrobots. *Small*, 15:1805006.
- Annesley, S. J., Fisher, P. R., Annesley, S. J., and Fisher, A. P. R. (2009). Dictyostelium discoideum—a model for many reasons. *Molecular and Cellular Biochemistry* 2009 329:1, 329:73–91.
- Asimov, I. (1942). *I, robot*. Bantam Spectra.
- Bahrami, B., Hojjat-Farsangi, M., Mohammadi, H., Anvari, E., Ghalamfarsa, G., Yousefi, M., and Jadidi-Niaragh, F. (2017). Nanoparticles and targeted drug delivery in cancer therapy. *Immunology Letters*, 190:64–83.
- Beladi-Mousavi, S. M., Hermanová, S., Ying, Y., Plutnar, J., and Pumera, M. (2021). A maze in plastic wastes: Autonomous motile photocatalytic microrobots against microplastics. *ACS Applied Materials and Interfaces*, 13.
- Bernstein, B. E., Stamatoyannopoulos, J. A., Costello, J. F., Ren, B., Milosavljevic, A., Meissner, A., Kellis, M., Marra, M. A., Beaudet, A. L., Ecker, J. R., Farnham, P. J., Hirst, M., Lander, E. S., Mikkelsen, T. S., and Thomson, J. A. (2010). The nih roadmap epigenomics mapping consortium. *Nature Biotechnology* 2010 28:10, 28:1045–1048.
- Bilodeau, R. A. and Kramer, R. K. (2017). Self-healing and damage resilience for soft robotics: A review. *Frontiers Robotics AI*, 4:48.
- Blackiston, D., Lederer, E., Kriegman, S., Garnier, S., Bongard, J., and Levin, M. (2021). A cellular platform for the development of synthetic living machines. *Science Robotics*, 6:31.
- Bongard, J. and Levin, M. (2021). Living things are not (20th century) machines: Updating mechanism metaphors in light of the modern science of machine behavior. *Frontiers in Ecology and Evolution*, 9.
- Buss, L. W. (1988). *The Evolution of Individuality*. Princeton University Press.
- Carnell, M. J. and Insall, R. H. (2011). Actin on disease – studying the pathobiology of cell motility using dictyostelium discoideum. *Seminars in Cell & Developmental Biology*, 22:82–88.
- Chattwood, A., Nagayama, K., Bolourani, P., Harkin, L., Kamjoo, M., Weeks, G., and Thompson, C. R. (2013). Developmental lineage priming in dictyostelium by heterogeneous ras activation. *eLife*, 2013.

- 
- Cheney, N. and Clune, J. (2013). Unshackling evolution evolving soft robots with multiple materials and a powerful generative encoding. *GECCO'13 : proceedings of the 2013 Genetic and Evolutionary Computation Conference*.
- Chisholm, R. L., Gaudet, P., Just, E. M., Pilcher, K. E., Fey, P., Merchant, S. N., and Kibbe, W. A. (2006). Dictybase, the model organism database for *dictyostelium discoideum*. *Nucleic Acids Research*, 34:D423–D427.
- Coello, C. A., Lamont, G. B., and Veldhuizen, D. A. V. (2007). *Evolutionary Algorithms for Solving Multi-Objective Problems*. Springer.
- Coleman, O. J., Blair, A. D., and Clune, J. (2014). Automated generation of environments to test the general learning capabilities of ai agents. *GECCO '14: Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation*.
- Corucci, F., Cheney, N., Giorgio-Serchi, F., Bongard, J., and Laschi, C. (2015). Evolving soft locomotion in aquatic and terrestrial environments: effects of material properties and environmental transitions. *Soft Robotics*, 5.
- Cremaldi, J. C. and Bhushan, B. (2018). Bioinspired self-healing materials: lessons from nature. *Beilstein Journal of Nanotechnology*, 9:907.
- Cvetkovic, C., Raman, R., Chan, V., Williams, B. J., Tolish, M., Bajaj, P., Sakar, M. S., Asada, H. H., Saif, M. T. A., and Bashir, R. (2014). Three-dimensionally printed biological machines powered by skeletal muscle. *Proceedings of the National Academy of Sciences of the United States of America*, 111:10125–10130.
- Deblandre, G. A., Wettstein, D. A., Koyano-Nakagawa, N., and Kintner, C. (1999). A two-step mechanism generates the spacing pattern of the ciliated cells in the skin of *xenopus* embryos. *Development*, 126:4715–4728.
- Ebrahimkhani, M. R. and Levin, M. (2021). Synthetic living machines: A new window on life. *iScience*, 24:102505.
- Eiben, A. E., Raué, P. E., and Ruttkay, Z. (1994). Genetic algorithms with multi-parent recombination. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 866 LNCS:78–87.
- Eichinger, L. and A.Noegel, A. (2003). Crawling into a new era- the *dictyostelium* genome project. *The EMBO Journal*, 22:1941–1946.
- Fey, P., Kowal, A. S., Gaudet, P., Pilcher, K. E., and Chisholm, R. L. (2007). Protocols for growth and development of *dictyostelium discoideum*. *Nature Protocols*, 2:1307–1316.
- Gasparetto, A., Scalera, L., Gasparetto, A., and Scalera, L. (2019). A brief history of industrial robotics in the 20th century. *Advances in Historical Studies*, 8:24–35.
- Gilbert, S. F., Sapp, J., and Tauber, A. I. (2012). A symbiotic view of life: we have never been individuals. *The Quarterly review of biology*, 87:325–341.

## BIBLIOGRAPHY

---

- Golstein, P., Aubry, L., and Levraud, J. P. (2003). Cell-death alternative model organisms: why and which? *Nature Reviews Molecular Cell Biology* 2003 4:10, 4:798–807.
- Grosberg, R. K. and Strathmann, R. R. (2007). The evolution of multicellularity: A minor major transition? *Source: Annual Review of Ecology, Evolution, and Systematics*, 38:621–654.
- Junemann, A., Filić, V., Winterhoff, M., Nordholz, B., Litschko, C., Schwellenbach, H., Stephan, T., Weber, I., and Faix, J. (2016). A diaphanous-related formin links ras signaling directly to actin assembly in macropinocytosis and phagocytosis. *Proceedings of the National Academy of Sciences of the United States of America*, 113:E7464–E7473.
- Kawabe, Y., Du, Q., Schilde, C., and Schaap, P. (2019). Evolution of multicellularity in dictyostelia. *The International journal of developmental biology*, 63:359.
- Kessin, R. H. (1975). *Dictyostelium: Evolution, Cell Biology, and the Development of Multicellularity*. Cambridge University Press.
- Konijn, T. M., Barkley, D. S., Chang, Y. Y., and Bonner, J. T. (1968). Cyclic amp: A naturally occurring acrasin in the cellular slime molds. <https://doi.org/10.1086/282539>, 102:225–233.
- Krakauer, D., Bertschinger, N., Olbrich, E., Flack, J. C., and Ay, N. (2020). The information theory of individuality. *Theory in Biosciences*, 139:209–223.
- Krakauer, D. C. and Zanotto, P. (2013). Viral individuality and limitations of the life concept. *Protocells*, pages 513–536.
- Kriegman, S., Blackiston, D., Levin, M., and Bongard, J. (2020). A scalable pipeline for designing reconfigurable organisms. *PNAS*, 117.
- Kriegman, S., Blackiston, D., Levin, M., and Bongard, J. (2021). Kinematic self-replication in reconfigurable organisms. *Proceedings of the National Academy of Sciences*, 118:e2112672118.
- Kriegman, S., Cappelle, C., Corucci, F., Bernatskiy, A., Cheney, N., and Bongard, J. C. (2017). Simulating the evolution of soft and rigid-body robots. *Gecco*.
- Lange, K., Buerger, M., Stallmach, A., and Bruns, T. (2016). Effects of antibiotics on gut microbiota. *Digestive Diseases*, 34:260–268.
- Lee, C., Kim, M., Kim, Y. J., Hong, N., Ryu, S., Kim, H. J., and Kim, S. (2017). Soft robot review. *International Journal of Control, Automation and Systems*, 15:3–15.
- Levin, M. (2020). How groups of cells cooperate to build organs and organisms. *The Scientist Magazine*.
- Levin, M., Bongard, J., and Lunshof, J. E. (2020). Applications and ethics of computer-designed organisms. *Nature Reviews Molecular Cell Biology* 2020 21:11, 21:655–656.
- Li, J., Ávila, B. E. F. D., Gao, W., Zhang, L., and Wang, J. (2017). Micro/nanorobots for biomedicine: Delivery, surgery, sensing, and detoxification. *Science Robotics*, 2.
- Lodish, H., Berk, A., Kaiser, C. A., Krieger, M., Bretscher, A., Ploegh, H., Amon, A., and Martin, K. C. (2016). *Molecular Cell Biology*. W. H. Freeman and Company.

- 
- Lyons, N. A. and Kolter, R. (2015). On the evolution of bacterial multicellularity. *Current opinion in microbiology*, 24:21.
- Mackay, D. J. C. (1995). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.
- Manicka, S. and Levin, M. (2019). The cognitive lens: a primer on conceptual tools for analysing information processing in developmental and regenerative morphogenesis. *Philosophical Transactions of the Royal Society B*, 374.
- Mitchell, M. (1996). *An Introduction to Genetic Algorithms*. The MIT Press.
- Morsut, L., Roybal, K. T., Xiong, X., Gordley, R. M., Coyle, S. M., Thomson, M., and Lim, W. A. (2016). Engineering customized cell sensing and response behaviors using synthetic notch receptors. *Cell*, 164:780–791.
- Nagasaki, A., Kanada, M., and Uyeda, T. Q. (2008). Cell adhesion molecules regulate contractile ring-independent cytokinesis in *dictyostelium discoideum*. *Cell Research* 2009 19:2, 19:236–246.
- Noegel, A. A. and Schleicher, M. (2000). The actin cytoskeleton of *dictyostelium*: a story told by mutants. *Journal of Cell Science*, 113:759–766.
- Palagi, S. and Fischer, P. (2018). Bioinspired microrobots. *Nature Reviews Materials*, 3:113–124.
- Park, S. J., Gazzola, M., Park, K. S., Park, S., Santo, V. D., Blevins, E. L., Lind, J. U., Campbell, P. H., Dauth, S., Capulli, A. K., Pasqualini, F. S., Ahn, S., Cho, A., Yuan, H., Maoz, B. M., Vijaykumar, R., Choi, J. W., Deisseroth, K., Lauder, G. V., Mahadevan, L., and Parker, K. K. (2016). Phototactic guidance of a tissue-engineered soft-robotic ray. *Science*, 353:158–162.
- Paschke, P., Knecht, D. A., Williams, T. D., Thomason, P. A., Insall, R. H., Chubb, J. R., Kay, R. R., and Veltman, D. M. (2019). Genetic engineering of *dictyostelium discoideum* cells based on selection and growth on bacteria. *JoVE (Journal of Visualized Experiments)*, 2019:e58981.
- Patra, D., Sengupta, S., Duan, W., Zhang, H., Pavlick, R., and Sen, A. (2013). Intelligent, self-powered, drug delivery systems. *Nanoscale*, 5:1273–1283.
- Purcell, E. (1977). Life at low reynold number. *American journal of physics*, pages 3–11.
- Reardon, S. (2022). First pig-to-human heart transplant: what can scientists learn? *Nature*.
- Risi, S. (2012). Towards evolving more brain-like artificial neural networks. *UCF Libraries*.
- Rossi, G., Manfrin, A., and Lutolf, M. P. (2018). Progress and potential in organoid research. *Nature Reviews Genetics* 2018 19:11, 19:671–687.
- Scavello, M., Petlick, A. R., Ramesh, R., Thompson, V. F., Lotfi, P., and Charest, P. G. (2017). Protein kinase a regulates the ras, rap1 and torc2 pathways in response to the chemoattractant camp in *dictyostelium*. *Journal of Cell Science*, 130:1545–1558.

## BIBLIOGRAPHY

---

- Sebille, E. V., Wilcox, C., Lebreton, L., Maximenko, N., Hardesty, B. D., Franeker, J. A. V., Eriksen, M., Siegel, D., Galgani, F., and Law, K. L. (2015). A global inventory of small floating plastic debris. *Environmental Research Letters*, 10:124006.
- Secretan, J., Beato, N., D'ambrosio, D. B., Rodriguez, A., Campbell, A., Folsom-Kovarik, J. T., and Stanley, K. O. (2011). Picbreeder: A case study in collaborative evolutionary exploration of design space. *Evolutionary Computation journal*.
- Servant, A., Qiu, F., Mazza, M., Kostarelos, K., and Nelson, B. J. (2015). Controlled in vivo swimming of a swarm of bacteria-like microrobotic flagella. *Advanced Materials*, 27:2981–2988.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423.
- Sharma, A., Arya, D. K., Dua, M., Chhatwal, G. S., and Johri, A. K. (2012). Nano-technology for targeted drug delivery to combat antibiotic resistance. <http://dx.doi.org/10.1517/17425247.2012.717927>, 9:1325–1332.
- Sharma, S. and Chatterjee, S. (2017). Microplastic pollution, a threat to marine ecosystem and human health: a short review. *Environmental Science and Pollution Research* 2017 24:27, 24:21530–21547.
- Singh, S., Mohamed, W., Aguessy, A., Dyett, E., Shah, S., Khan, M., Baskar, R., and Brazill, D. (2017). Functional interaction of pkca and pldb regulate aggregation and development in dictyostelium discoideum. *Cellular Signalling*, 34:47–54.
- Soll, D. R., Wessels, D., Heid, P. J., and Voss, E. (2003). Computer-assisted reconstruction and motion analysis of the three-dimensional cell. *TheScientificWorldJournal*, 3:827–841.
- Soreni-Harari, M., Pierre, R. S., McCue, C., Moreno, K., and Bergbreiter, S. (2020). Multimaterial 3d printing for microrobotic mechanisms. <https://home.liebertpub.com/soro>, 7:59–67.
- Soto, F., Karshalev, E., Zhang, F., de Avila, B. E. F., Nourhani, A., and Wang, J. (2022). Smart materials for microrobots. *Chemical Reviews*, 122:5365–5403.
- Srinivasan, S., Alexander, H., and Alexander, S. (2000). The dictyostelium fruiting body – a structure of cells and cellulose. *Trends in Cell Biology*, 10:315.
- Stanley, K. O. (2007). Compositional pattern producing networks: A novel abstraction of development. *Genetic Programming and Evolvable Machines*, 8:131–162.
- Stanley, K. O. and Miikkulainen, R. (2002). Evolving neural networks through augmenting topologies. *The MIT Press Journals*.
- Stone, J. V. (2015). *Information Theory A Tutorial Introduction*. Sebtel Press, 1st edition.
- Suess, P. M., Watson, J., Chen, W., and Gomer, R. H. (2017). Extracellular polyphosphate signals through ras and akt to prime dictyostelium discoideum cells for development. *Journal of Cell Science*, 130:2394–2404.
- Thompson, D. W. (1945). *On Growth and Form*. Cambridge University Press.

- 
- Toda, S., Blauch, L. R., Tang, S. K., Morsut, L., and Lim, W. A. (2018). Programming self-organizing multi-cellular structures with synthetic cell-cell signaling. *Science*, 361:156–162.
- Tong, K., Bozdag, G. O., and Ratcliff, W. C. (2021). Selective drivers of simple multicellularity. *Preprints 2021*.
- Wallingford, J. B., Liu, K. J., and Zheng, Y. (2010). Xenopus. *Current Biology*, 20:R263–R264.
- Webster, V. A., Young, F. R., Patel, J. M., Scariano, G. N., Akkus, O., Gurkan, U. A., Chiel, H. J., and Quinn, R. D. (2017). 3d-printed biohybrid robots powered by neuromuscular tissue circuits from. *Biomimetic and Biohybrid Systems. Living Machines 2017*, pages 475–486.
- Weijer, C. J. (2004). Dictyostelium morphogenesis. *Current Opinion in Genetics and Development*, 14:392–398.
- Weiner, D., Neptune, E. R., Sedat, W., cariy H out chemotaxis Arai, Monteclaro, F. S., Bourne, C. L. H. R., Artavanis-Tsakonas, S., Rand, M. D., and Lake, R. J. (1999). Notch signaling: Cell fate control and signal integration in development. *Mol. Biol. Cell*, 10:25037–25073.
- West, S. A., Griffin, A. S., Gardner, A., and Diggle, S. P. (2006). Social evolution theory for microorganisms. *Nature Reviews Microbiology* 2006 4:8, 4:597–607.
- West-Eberhard, M. J. (1975). The evolution of social behaviour by kin selection. *The quarterly review of biology*.
- Williams, J. G. (2006). Transcriptional regulation of dictyostelium pattern formation. *EMBO reports*, 7:694–698.
- Williams, P. L. and Beer, R. D. (2010). Nonnegative decomposition of multivariate information. *Cornell University*.
- Wilson, R. and Barker, M. (2007). The biological notion of individual. *Stanford encyclopedia of philosophy*.
- Yu, J., Wang, B., Du, X., Wang, Q., and Zhang, L. (2018). Ultra-extensible ribbon-like magnetic microswarm. *Nature Communications* 2018 9:1, 9:1–9.
- Zhang, B., Korolj, A., Lai, B. F. L., and Radisic, M. (2018). Advances in organ-on-a-chip engineering. *Nature Reviews Materials* 2018 3:8, 3:257–278.