

THIS IS A PRE-PRINT, WEB VERSION OF THE  
PH.D. DISSERTATION:

“DIGITAL PHILOSOPHY”

THE TEXT IS RELEASED UNDER A STANDARD CC BY-NC-ND  
LICENSE.

FOR THE FULL, FINAL VERSION, PLEASE CONTACT ME.

**UNIVERSITA' VITA-SALUTE SAN RAFFAELE**

**PROGRAMMA DI DOTTORATO IN FILOSOFIA E SCIENZE  
COGNITIVE**

**DIGITAL PHILOSOPHY**

*Formal Ontology and Knowledge Representation in Cellular Automata*

First Supervisor: Prof. Michele Di Francesco

Second Supervisor: Prof. Francesco Berto

Tesi di DOTTORATO di RICERCA di Jacopo Tagliabue  
matr. 004943

Ciclo di dottorato: XXV

SSD: M-FIL/01

Anno Accademico 2011/2012

## **Abstract**

The core working assumption behind a “digital philosophy” is that digital universes may be helpful to systematic philosophy, and *vice versa*. We first introduce particular computational devices, known as cellular automata, and use them as intended models for a formal ontology of digital worlds – worlds that are discrete in space and time. *Chapter 1* and *Chapter 2* are devoted to develop a full-fledged formal theory, drawing from tools in mereology, topology and standard metaphysics; along the way, existing debates are framed in a digital context and old puzzles receive new insights. In *Chapter 3* we produce a digital, up-to-date version of the Leibniz-Carnap dream, i.e. a “theory of everything” allowing every statement about our digital world to be effectively computed. By implementing the ontology we are able to discuss the computational properties of our metaphysical assumptions and assess from a non-standard perspective the modeling capabilities of the theory. Finally, three appendices explore in depth particular topics at the intersection of cellular automata and philosophical subfields: philosophy of information, philosophy of psychology, philosophy of physics.

## Table of Contents

|                                                 |    |
|-------------------------------------------------|----|
| 0. Introduction                                 | 3  |
| 0.0 Philosophy in a Digital World               | 3  |
| 0.1 Cellular Automata                           | 6  |
| 0.2 Project Structure                           | 9  |
| 1. Parts and Places                             | 12 |
| 1.0 The Game of Life                            | 12 |
| 1.1 Toward a CA Proto-Geometry: Mereology       | 17 |
| 1.2 Toward a CA Proto-Geometry: Topology        | 25 |
| 1.3 The Meaning of Life                         | 27 |
| 1.4 Questions and Answers                       | 30 |
| 2. Time and Other Dimensions                    | 36 |
| 2.0 The Time of our Life                        | 36 |
| 2.1 Laws of Nature                              | 41 |
| 2.2 Layers                                      | 45 |
| 2.3 Emergence Reconsidered                      | 48 |
| 2.4 Possible Digital Universes                  | 50 |
| 2.5 Questions and Answers                       | 53 |
| 3. Calculemus                                   | 56 |
| 3.0 The Aufbau Digitally Remastered             | 56 |
| 3.1 Java Coding 101                             | 58 |
| 3.2 The Computational Cost of Ontology          | 61 |
| 3.3 Computational Ontology on the Cheap         | 69 |
| 3.4 The Reality of Time                         | 73 |
| 3.5 Layers in Computational Semantics           | 76 |
| 3.6 Lessons from <i>AufByte</i>                 | 79 |
| 3.7 Questions and Answers                       | 82 |
| 4. Conclusion                                   | 86 |
| 4.0 Looking Back, Going Forward                 | 86 |
| 4.1 <i>AufByte</i> and the Future of Philosophy | 88 |

|                                                   |     |
|---------------------------------------------------|-----|
| Appendix I: The Formal Theory                     | 91  |
| I.1 Introduction                                  | 91  |
| I.2 Axioms from <i>Chapter 1</i>                  | 91  |
| I.3 Axioms from <i>Chapter 2</i>                  | 95  |
| I.4 Concluding Remarks                            | 97  |
| Appendix II: Against ‘Against Digital Philosophy’ | 98  |
| II.1 Introduction                                 | 98  |
| II.2 Against Digital Philosophy                   | 98  |
| II.3 The Objection from Counting                  | 101 |
| II.4 The Objection from Mereology                 | 103 |
| II.5 Concluding Remarks                           | 103 |
| Appendix III: CA and the Philosophy of Psychology | 104 |
| III.1 Introduction                                | 104 |
| III.2 The Interface Problem vs. Anomalous Monism  | 105 |
| III.3 Anomalous Monism and Life                   | 108 |
| III.4 Emergent Computations and Explanations      | 109 |
| III.5 Concluding Remarks                          | 113 |
| Appendix IV: CA and the Philosophy of Physics     | 114 |
| IV.1 Introduction                                 | 114 |
| IV.2 A Philosopher’s View                         | 115 |
| IV.3 A Physicist’s View                           | 117 |
| IV.4 Concluding Remarks                           | 120 |
| Bibliography                                      | 122 |

## 0. Introduction

### 0.0 Philosophy in a Digital World

It is not worth an intelligent man's time to be in the majority.  
By definition, there are already enough people to do that.  
*Godfrey H. Hardy*

The world we live in is a very digital thing. Every stick and every stone you have ever seen are made of fundamental, discrete elements (“atoms” in the original sense of the word). And so are you and I. And so are the planet Earth, the solar system, the entire Milky Way, the remote galaxies we see through telescopes. There is nothing too complex that cannot be divided, at least in principle, in fundamental parts – time and space themselves constituting no exception<sup>1</sup>.

Or, maybe, this is absolutely *not* the case. But even if our world is not digital after all, there certainly are *some* worlds which are purely so, i.e. fundamentally discrete in space and time and any other primitive “physical” quantity. These worlds may be  $n$ -dimensional, have the craziest topology, but they all share the spirit (if not the letter) of what Ed Fredkin called the “Finite Nature Hypothesis”, that is the idea that ‘ultimately every quantity of physics, including space and time, will turn out to be discrete and finite; that the amount of information in any small volume of space-time will be finite and equal to one of a small number of possibilities’<sup>2</sup>. Moreover, even if our world is not digital, there are nonetheless very good digital approximations to its fundamental structure, whatever it is<sup>3</sup>.

The fancy label “digital philosophy” sounds perfect for *this* work. In a nutshell, *digital philosophy* is the claim that discrete universes can be useful to systematic philosophy and *vice versa* – and to see *why* we start by saying something about the latter.

As philosophers, we are committed to come up with a list of basic entities and simple rules out of which everything we “see” – atoms, people, galaxies, sets, *possibilia*, moral

---

<sup>1</sup> The philosophically sophisticated reader will recognize my small tribute to Lewis (1986b). Simply put, there is David K. Lewis and then there is the rest of us.

<sup>2</sup> Fredkin (1993), p. 116.

<sup>3</sup> See *Appendix IV* for a more detailed investigation of digital universes and philosophy of physics.

values – can be built. Non-philosophers can easily imagine our work as some sort of reverse-LEGO: you start with the whole model in front of you and the task is to compile the list of items that were in the LEGO kit in the first place<sup>4</sup>. The world we live in is a hell of a LEGO model, so it is no wonder that philosophers are fighting all the time about which items should make the final list. It is far too easy to joke about the results of this two millennia enterprise, but reverse-LEGO is *very* tricky: it is a fact obvious since Euclid that every clever argument must, at some point, rely on assumptions, which in turn are basically supported by “primitive intuitions” – and since reverse-LEGO does not allow us to touch the construction, this is an even greater challenge. Of course, if only we had reasons to think that our intuitions are designed to succeed in reverse-LEGO, arguments would be much easier to win; quite to the contrary, we actually have good evidence supporting the fact that our intuitions are just that, *intuitions*. There is a growing debate in contemporary philosophy<sup>5</sup> on the role of intuitions in philosophical enquiry; however, meta-philosophy is better left to older scholars. As far as *this* work is concerned, all we need is to acknowledge that a theory is the result of a “reflective equilibrium”, a sophisticated balance between intuitions and theoretical virtues (unification, simplicity, elegance, etc.) – which is pretty much the implicit standard for the best metaphysics of the last forty years, and the only way *I* can make sense of philosophical discussions<sup>6</sup>.

The fact that intuitions inform at least some part of our theorizing is the first reason for *this* essay: together with those that take themselves too seriously, philosophers without imagination are the most dangerous kind – as we sometimes declare from our armchair the metaphysical impossibility of *X* just because our imagination does not yet cover *X*, good science has always been doing a great job in broadening our horizons. Digital universes are *great* science and can do the job perfectly<sup>7</sup>.

The second rationale behind *this* work is to apply philosophical reasoning in a world whose “starting kit” is known *a priori* – yes, we are cheating at reverse-LEGO, but for good reasons: philosophical proposals have been developed in connection with our

---

<sup>4</sup> Throughout the history of philosophy this very project was explicitly stated several times: *aficionados* may easily recall Leibniz’s ‘De Alphabeto Cogitationum Humanarum’, Carnap’s *Aufbau* and, more recently, Chalmers’ ‘Constructing the World’. We shall come back later to these historical antecedents.

<sup>5</sup> See for example Williamson (2008), Hintikka (1999) – but see Cappelen (2012) for a somewhat critical perspective.

<sup>6</sup> More explicitly, the approach has been stated in Lewis (1983a), pp. ix-xi.

<sup>7</sup> For the pioneering use of CA in a philosophical context, see Dennett (1991). A non-technical introduction to this topic can be found in Tagliabue (2013).

everyday reality, but it is not always clear how to assess them. For example, the “best system theory of laws” (proposed by Frank Ramsey and David Lewis<sup>8</sup>) holds that natural laws are the true generalizations contained in the best deductive system describing our world: unfortunately, no one has the slightest idea of how this system may look like. So what happens when we apply this account to our digital universe? Does it have any unexpected or unwanted consequence?

A third reason for digital philosophy is the growing interest cognitive sciences (Artificial Intelligence in particular) are devoting to qualitative theories, a trend witnessed by the success of formal ontology in the IT business<sup>9</sup>. Exciting challenges lie at the crossroads of philosophy, logic and computer science: digital philosophy may have unexpected practical consequences by helping practitioners to address these issues from new perspectives.

A fourth reason (related to the third) is methodological: once you have your reverse-LEGO list, you can put it to work. How? For example, by teaching a computer how to reason about your toy universe and seeing if the results match your intuitions: if so, there are good chances that your ontology is not that bad after all. Digital universes are computer-friendly almost by definition, but what about ontology? I hold that any ontology that fits the bill should be expressible in a suitable formal language, an idea that may be traced back to Leibniz and that is nicely presented in many recent works in formal ontology (e.g. ‘Whatever the content of the theory of space one endorses, whatever the logic one subscribes to, whatever the limitations one sets by selecting certain cognitive facts as relevant, at some point the method of philosophy requires – as we may put it – that one calculate’<sup>10</sup>). If this constraint is enforced, digital universes allow for a somewhat new benchmark: by letting the computer interpret the universe through the lenses of our axioms, we can “experimentally” check if the theory is coherent and complete relatively to the domain of interest<sup>11</sup>. Finally, if none of the above is sufficient, a pragmatic – if not desperate – reason can be found in this deceptively simple argument: if we cannot reach some sort of philosophical consensus

---

<sup>8</sup> See Ramsey (1978), Lewis (1973a).

<sup>9</sup> See for example Borgo, Lesmo (2008), Ferrario, Oltramari (2009).

<sup>10</sup> Casati, Varzi (1999), p. 201.

<sup>11</sup> Grim, Mark, St. Denis (1998) has a different target, but certainly shares with *this* work the idea that computers can be very useful to model philosophical concepts. I’m not aware of any ontological project closely related to the current enterprise.

in such a small, controlled and *fundamental* environment, how can we hope that these tools will succeed when the real world is in front of us?

Of course, each of these threads of thought will be thoroughly explored in the following chapters: now that we have just introduced philosophy, it is time to present the remaining *dramatis personae*, cellular automata.

## 0.1 Cellular Automata

Perfection is achieved  
Not when there is nothing more to add,  
But when there is nothing left to take away  
*Antoine de Saint-Exupery*

There are many kinds of digital worlds but we shall focus on particular structures known as *cellular automata*<sup>12</sup>.

Cellular automata (henceforth: CA) are *discrete, abstract computational* systems that have proved useful both as general models of complexity and as more specific representations of non-linear dynamics in a variety of scientific fields. While CA can be specified in purely mathematical terms, having a concrete instance in mind can nevertheless help in the beginning: think of the following picture as standing for the front row of a high school classroom. Each box represents a student wearing (white) or not wearing (black) a hat:



Let us make the two following assumptions:

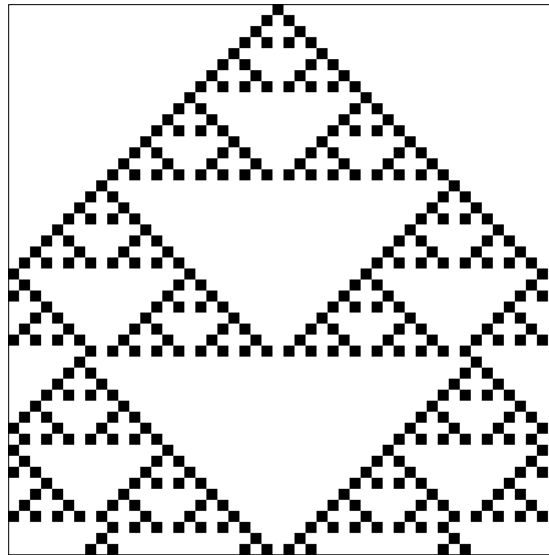
*Hat rule:* a student will wear the hat in the following class if one or the other - but not both - of the two classmates sitting immediately on her/his left and on her/his right has the hat in the current class (let us say that if nobody wears the hat, then a hat is for losers; but if both neighbors wear it, it is too popular to be trendy).

---

<sup>12</sup> The following material is readapted from Berto, Tagliabue (2012), which is a friendly introduction to cellular automata from a philosophical perspective. Another good introduction is Mainzer, Chua (2012).

*Initial class:* during the first class in the morning, only one student in the middle is wearing the hat.

The picture below shows what happens as time goes by (consecutive rows represent the evolution in time through subsequent classes):



The image should be surprising. The complex pattern displayed contrasts with the simplicity of the underlying law (the “Hat rule”): the scale at which the decision to wear the hat is made (immediate neighbors) is not the scale at which the interesting patterns become showy. While somewhat artificial, this example is a paradigmatic illustration of what makes CA appealing to a vast range of researchers: ‘even perfect knowledge of individual decision rules does not always allow us to predict macroscopic structure. We get macro-surprises despite complete micro-knowledge’<sup>13</sup>.

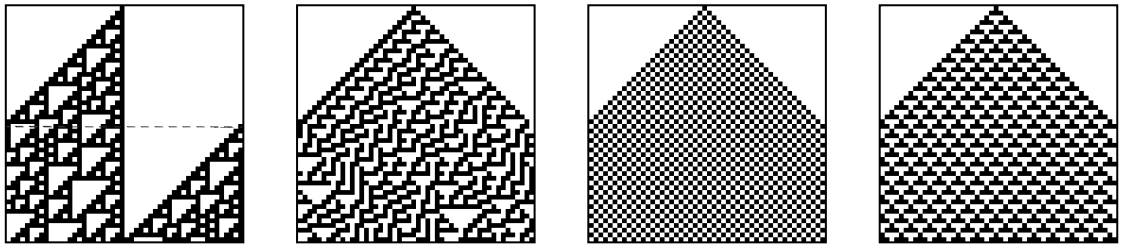
Of course, CA can be used to model much more interesting phenomena than our high-school class: in fact, the variety of applications to be found in CA literature is almost endless<sup>14</sup>. However, there is more to it than family resemblances, since CA share some basic, defining features: first of all, CA are constituted by a grid of  $n$  atomic cells – the grid may be one, two,  $k$ -dimensional, and atomic cells may come in different

---

<sup>13</sup> Epstein (1999), p. 48.

<sup>14</sup> To mention just a few: urban evolution (Batty (2005)), Ising models (Creutz (1986)), neural networks (Franceschetti, Campbell, Hamneken (1992) pp. 124–128), turbulence phenomena (Chen, Chen, Doolen, Lee, (1983)), battlefields (Ilachinski (2004)).

shapes (e.g. square, hexagon, cube, etc.), but they are finite in number and qualitatively indiscernible; second, time flows in *discrete* time-steps; third, each cell can be, at a given time-step, in one out of finitely many states (e.g. 1/0); fourth, at each time-step, all cells synchronously compute their next state using a rule such as ‘if, at the previous time-step, your neighboring cells  $c_1, \dots, c_n$  were in state  $s_1, \dots, s_n$ , then assume state  $s_k$ ’. In the previous example we had a one-dimensional grid with squared atomic cells, 50 time-steps, two possible states (hat/no hat) and a simple rule ‘if, at the previous time-step, your two neighboring cells were not in the same state, then assume state *hat*’. Although we will further explore CA immense variability in later sections, it is worth noting immediately how just changing the rule may dramatically affect the emergent behavior of the system, as exemplified by the four space-time diagrams below:



More formally we define a CA as any system satisfying the following requirements:

CA<sub>def.1</sub>) *Discrete space-time*: space is made of a finite number of qualitatively indiscernible atomic cells; time flows in discrete time-steps.

CA<sub>def.2</sub>) *Discrete states*: there is a finite number of possible states for each cell.

CA<sub>def.3</sub>) *Discrete dynamics*: at each time-step, each cell updates her state with a function mapping neighbors’ configurations into possible states. The update is *synchronous* and the function is such that it takes into account what were the neighbors’ states only at *the previous timestep*.

A CA behavior is thus *entirely determined* by these factors (plus, of course, the initial conditions) and this is what makes CA appealing for the whole reverse-LEGO game: the emergent behavior is so rich, yet heavily constrained by simple features and rules we ourselves set up. If you are not entirely sold on the amazing capabilities of these

systems, the CA presented in *Chapter 1* will clear any residual doubt: let us get to work.

## 0.2 Project Structure

Make something idiot-proof  
and someone will come up  
with a better idiot.  
*Anonymous*

So *this* work is about philosophy *and* digital universes: but what *exactly* are we going to do? The following is a first recognition of the exciting landscape we are about to explore.

We all remember the *Empire Strikes Back* memorable scene: Han Solo, Leila and Chewbacca are held captives by Lando Carlissian; Leila and Han profess their love – ‘I love you’, ‘I know’ – just before Han is sealed in carbonite by Empire’s soldiers. To make sense of what is happening we seamlessly integrate information from claims such as ‘Leila’s declaration *causes* Han’s response’, ‘Han is *part of* a carbonite block’, ‘Although Han is part of a carbonite block, he is *the same person as* before’, ‘Sealing innocent people in carbonite is *wrong*’ and many more. What all these claims have in common is that they make essential use of “universal” concepts (*parthood, causation, identity, morality*), concepts so crucial for our understanding of reality that they have attracted the attention of philosophers for millennia, concepts so powerful that they can be used to understand situations we have not been previously exposed to, such as a world governed by the Force.

The fundamental question we will try to answer is: ‘how are we to best understand these concepts?’. The question itself is *not* original at all, of course; what is unusual is that we are putting these concepts in a very specific context, cellular automata. CA offer one of the best environment to observe complex behavior and pattern formation: studying concepts in their *purest* form will hopefully challenge some of our firmest intuitions about the metaphysical structure of the world. Readers unfamiliar with formal ontology may well read the essay *as if* it was somehow an answer to the following question: ‘if we were to teach to a computer *how* to meaningfully talk about digital universe, what concepts should we introduce and how?’. Readers familiar with the

history of philosophy may read the essay *as if* it was a computational, CA-based version of the Leibniz/Carnap project of deducing (i.e. computing) all the truths about the world starting from a minimal set of primitive concepts and facts<sup>15</sup>.

From time to time, translating a typical philosophical question into our language will result in a somewhat *different* question, which of course has some bearing to the original issue, but that would nevertheless be unsatisfactory for some scholars: hopefully, what we lose in generality, we gain in precision and clarity. *Chapter 1* is devoted to lay down a sort of “proto-geometry”, a basic framework of mereological and topological notions that will form the core of our theory. In *Chapter 2* we will supplement the theory with substantial metaphysical notions, addressing the problems of identity through time, emergence and modality. In *Chapter 3* we will test our theory to see if the definitions we came up with are indeed capable of making sense of what happens in a digital world: in order to achieve this, we teach them to computers and see what *they* can understand through them. This is indeed a tough benchmark, since the distance between a standard metaphysics and the not-very-forgiving jargon of programming code is apparently huge, but, as we shall argue, it unveils very interesting consequences, both in methodology and content (and, by the way, we all know the old saying: ‘you do not really understand something unless you can explain it to your *laptop*’). At the end of each chapter a Q&A session elucidates specific aspects of the theory, replies to common objections and answers more technical worries that have been put aside in the main exposition.

*This* work grew out of three main passions: ontology, complex systems and computer science; as such, I would like practitioners with different backgrounds to be able to read it with the least possible effort (of course, the final result is unlikely to be entirely satisfactory for specialists in any of these subjects). As in every multidisciplinary essay a hard choice needs to be made along the continuum between assuming too much (making the text *hard*) and assuming too little (making the text *boring*). I hold as two of my deepest convictions that:

- 1) Everything that is said with symbols could be said, *given time, patience and ability*, without them.
- 2) Without symbols scientific progress is not *practically* possible.

---

<sup>15</sup> More recently, an epistemological version of the *Aufbau* has been proposed in Chalmers (2012). The link with these historical antecedents and David Lewis’ *Humean Supervenience* will be explored in later sections.

In the attempt to respect (1) I chose to write the main part of the essay as understandable as possible (working knowledge of first-order logic being the only prerequisite<sup>16</sup>); trying to respect (2) I chose to use a (generally) friendly language, add gentle introductions to unfamiliar concepts, point to educational references for every topic that cannot be adequately explained in *this* work (for example, *Chapter 1* will sound obvious to philosophers, but not so obvious to computer scientists, *vice versa* for *Chapter 3*). More technical material and in depth discussions of specific topics in the philosophical literature (analyzed, of course, from a digital perspective) can be found in the appendices.

As a final consideration, it should be noted that *this* dissertation comes with a substantial amount of lines of code. All the programs are freely available (until I realize I can sell them for millions, of course) – feel free to request them through my personal e-mail address: *tagliabue.jacopo@gmail.com*.

---

<sup>16</sup> The material covered by an introductory-to-intermediate volume (such as Barwise, Etchemendy (2002)) should be sufficient.

## 1. Parts and Places

### 1.0 *The Game of Life*

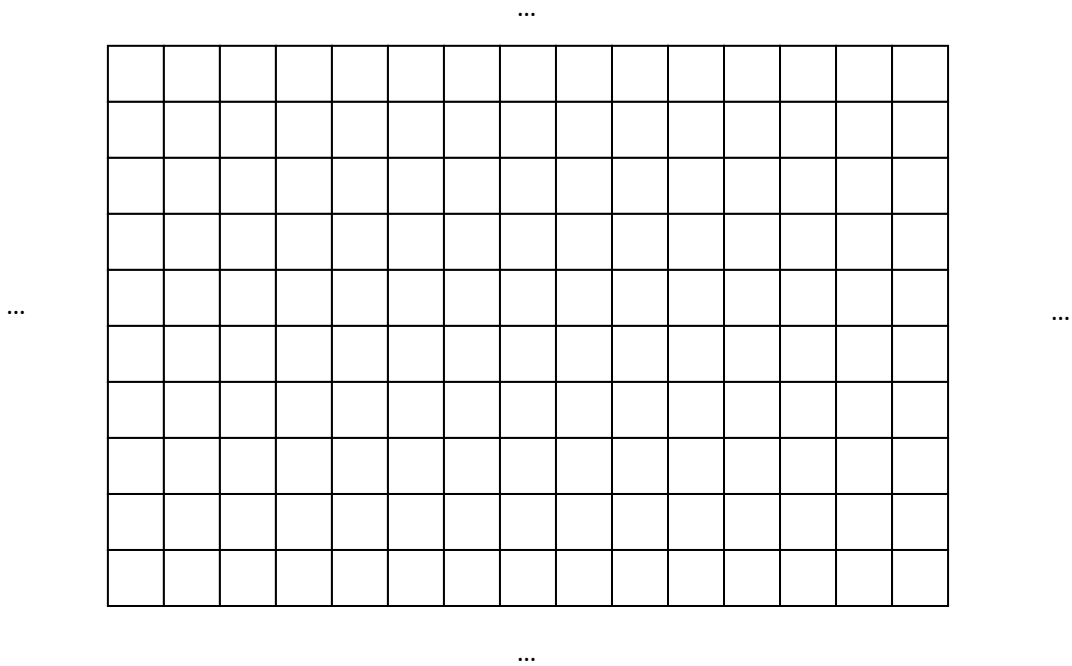
S.: This is the worst part. The calm before the battle.

F.: And then the battle is not so bad?

S.: Oh, right. I forgot about the battle.

*Futurama*

If there is such a thing as a true classic in the CA field, this is *The Game of Life*. Invented by John Conway<sup>17</sup> back in the Seventies, *Life* is by far the most popular CA ever created – the best proof that simple rules can indeed produce incredibly complex behavior. *Life* discrete space-time is a two-dimensional lattice of atomic square cells:



Each cell in *Life* can have one of two possible states, *dead* (black) or *alive* (white) – it will be obvious in a moment why they have such distinguished names and not just “1” and “0”.

---

<sup>17</sup> See Berkelamp, Conway, Guy (1982).



*Alive cell*



*Dead cell*

At each time-step, each cell updates its state according to the following rule:

*Birth*: if the cell at  $t - 1$  was dead, the cell becomes alive if exactly three neighbors were alive at  $t - 1$ .

*Survival*: if the cell at  $t - 1$  was alive, the cell state is still alive if either two or three neighbors were alive at  $t - 1$ .

*Death*: if the cell state at  $t - 1$  was alive, the cell state becomes dead if either fewer than two (it dies of loneliness) or more than three neighbors (it dies for overcrowding) were alive at  $t - 1$ .

In *Life* the contrast between the simplicity of local dynamics and the resulting emergent behavior is astonishing – as Conway himself summarized:

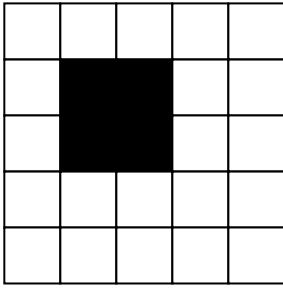
It's probable, given a large enough Life space, initially in a random state, that after a long time, intelligent self-reproducing animals will emerge and populate some parts of the space.<sup>18</sup>

Following the release of *Life*, a community of game's lovers was born and dozens of interesting patterns were discovered and shared<sup>19</sup>: what is now called *Life*'s “zoology” has plenty of “animals” – such as *Still Lifes* (i.e. immutable patterns):

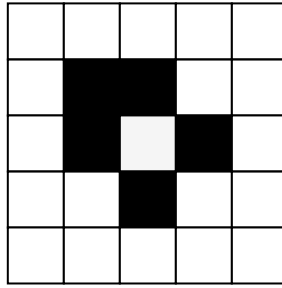
---

<sup>18</sup> Cited in Ilachinski (2001), p. 131.

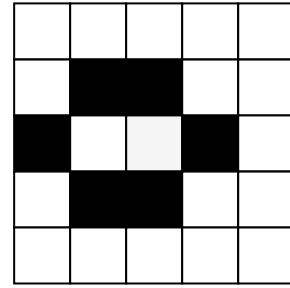
<sup>19</sup> As a small illustration of this fact, consider that the Life Wiki ([http://www.conwaylife.com/wiki/Main\\_Page](http://www.conwaylife.com/wiki/Main_Page)) offers a downloadable package of more than 3000 different patterns (last update, January 2012).



*Block*

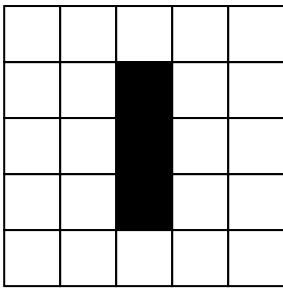


*Boat*

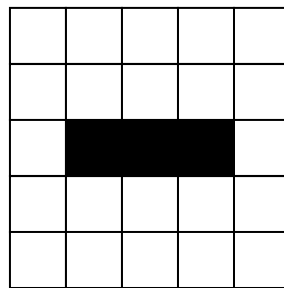


*Beehive*

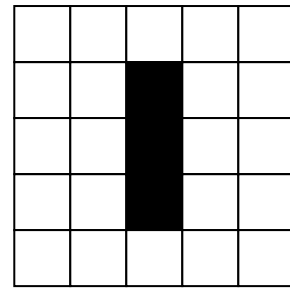
*Oscillators* (i.e. patterns cycling between two or more configurations):



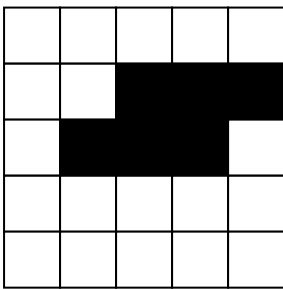
*Blinker  $t_0$*



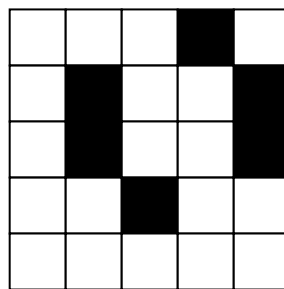
*Blinker  $t_1$*



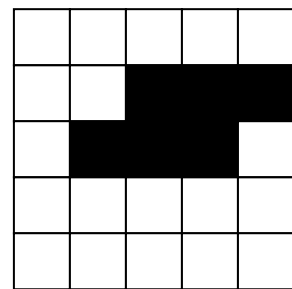
*Blinker  $t_2$*



*Toad  $t_0$*

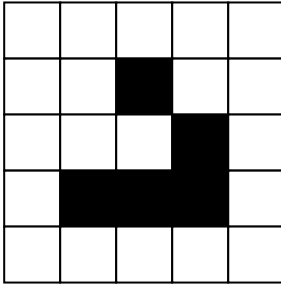


*Toad  $t_1$*

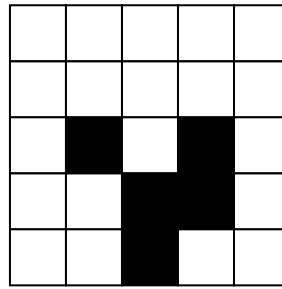


*Toad  $t_2$*

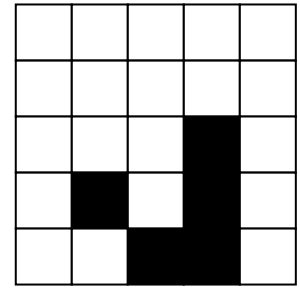
And *Spaceships* (i.e. patterns moving in the lattice):



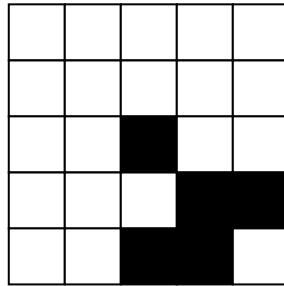
*Glider  $t_0$*



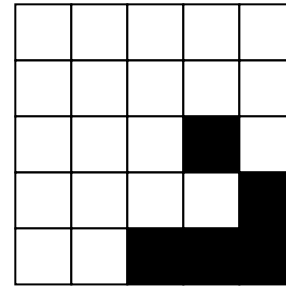
*Glider  $t_1$*



*Glider  $t_2$*



*Glider  $t_3$*



*Glider  $t_4$*

While it is hard to truly appreciate *Life*'s phenomenological complexity with pictures, we have collected snapshots of a “typical” model runs starting from random initial conditions – however, we strongly encourage the reader to explore *Life* with the code released with *this* work or one of the many simulators that Google easily provides.



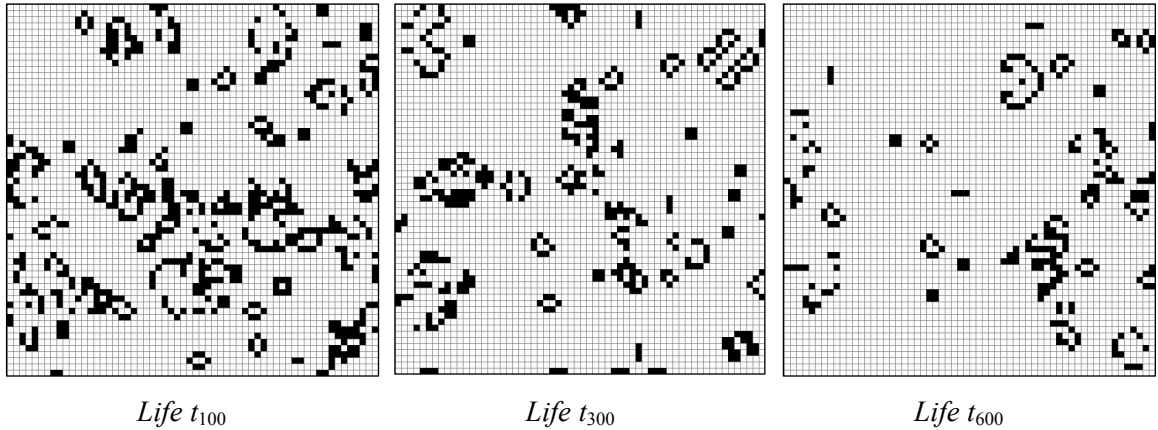
*Life  $t_0$*



*Life  $t_{20}$*



*Life  $t_{50}$*



We chose *Life* as our test-case for several reasons: granted, *Life* is the most popular CA out there and its dynamics (birth/survival/death) is very intuitive. However, we also chose *Life* because it is a paradigmatic example of two properties that makes digital universes worth exploring: first, all *Life* “physics” and “biology” are supported by a perfectly digital, *finite* world – the number of cells involved in a given simulation may be huge, but it is still finite and the cells’ states come from a finite set. In Ed Fredkin's worlds, it is a universe obeying the Finite Nature Hypothesis:

Ultimately every quantity of physics, including space and time, will turn out to be discrete and finite; that the amount of information in any small volume of space-time will be finite and equal to one of a small number of possibilities.<sup>20</sup>

Second, if you do not believe us when we say that *Life* displays complex behavior, we can always *prove* that to you: in fact, it is known that this CA is perfectly equivalent to a Universal Turing Machine<sup>21</sup> – so any algorithmic procedures your PC and your brain support, so does *Life*. Both these facts may not seem crucial right now, but they carry very interesting consequences we are going to exploit later (spoiler alert: if we prove that *our* universe shares these properties, we may generalize the consequences).

Now that *Life* has been introduced, we will take it to be *the intended model* for our analysis in this and later chapters: so, when we speak, say, in defense of unrestricted mereological composition, we are not so much concerned with its unconditional validity, rather we are interested in its usefulness for the universe we are investigating.

<sup>20</sup> Fredkin (1993), p. 116.

<sup>21</sup> See Berkelamp, Conway, Guy (1982) for the proof.

### 1.1. Toward a CA Proto-Geometry: Mereology

These are my principles.  
If you don't like them, I have others.  
*Groucho Marx*

To describe the metaphysical structure of digital universes we first need a formal theory describing the most basic feature of these worlds, their spatial structure.

In what follows we are going to introduce the first axioms of the theory using mereotopological notions: as already recognized by many scholars<sup>22</sup> the integration of mereological and topological concepts provide for a rich and flexible logical structure when dealing with spatial entities. As in other parts of *this* work, our strategy will be that of using “orthodox” material from the existing literature as starting point<sup>23</sup>, making “unorthodox” use of arguments or axioms whenever it is necessary to fully capture the peculiar features of CA. We start by introducing uncontroversial facts about parthood – the lexical part of the theory:

PL.1) Everything is part of itself.

PL.2) Two distinct things cannot be part of each other.

PL.3) Any part of any part of a thing is itself part of that thing.

In other words, the three principles together assert that *parthood* is a so-called partial ordering, i.e. a relation that is *reflexive*, *antisymmetric*, *transitive*. Armed with (PL.1)-(PL.3) we can also add definitions for other useful mereological predicates, such as *overlap*, *underlap*, *proper part*:

O<sub>def</sub>) Two objects *overlap* iff there is an object that is part of both.

U<sub>def</sub>) Two objects *underlap* iff there is an object of which they are both parts.

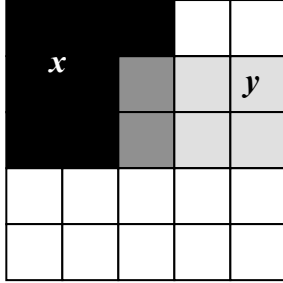
PP<sub>def</sub>) Any part of an object is a *proper part* iff it is not identical to that object.

---

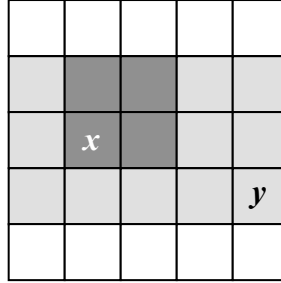
<sup>22</sup> See for example Casati, Varzi (1999) in philosophy, Cohn, Renz (2008) for Artificial Intelligence.

<sup>23</sup> The main reference is the excellent Casati, Varzi (1999). We will also heavily rely on arguments from Berto, Rossi, Tagliabue (2010): this chapter may be considered an extension and refinement of the formal sketch developed there. I discovered Galton (2000) almost at the end of the dissertation: the book contains smart discussions on many topics in spatial ontology – however, due to the particular nature of CA, most insights have only indirect bearing on what is discussed here.

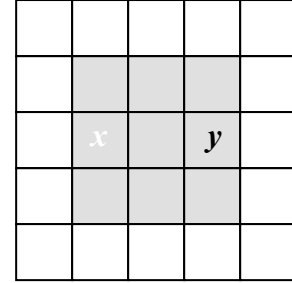
Although simple, the theory developed so far already allows us to model the basic patterns of mereological relationships (asterisks mark symmetric relations):



*x overlaps y\**



*x overlaps y\**  
*x underlaps y\**  
*x is part of y*  
*x is proper part of y*



*x overlaps y\**  
*x underlaps y\**  
*x is part of y*  
*y is part of x*  
*x is identical to y\**

However, to get a full-blown picture we need to supplement these core axioms with ‘principles asserting the (conditional) existence of certain mereological items given the existence of other items’<sup>24</sup>. In particular we start with a “supplementation principle”:

(PS) If an object is not part of another, some part of the former does not overlap the latter.

(PS) formally encodes an important intuition about spatially extended objects: if one thing is not part of another, there must be something in the universe that account for that fact; for example, if France is not part of Italy, there must be some part of France that does not overlap Italy. From a philosophical point of view, the important thing is that from (PS) (together with (P.1)-(P.3)) a form of extensionality can be derived, that is:

(PE) Two objects are identical iff they have the same parts.

What (PE) actually asserts is that objects are exhaustively defined by their parts, pretty much as sets are defined by their members<sup>25</sup>. For sure, some may resist the

<sup>24</sup> Casati, Varzi (1999), p. 38.

<sup>25</sup> Actually, mereology is even “more extensional” than set theory, since in the latter you can have different sets starting from the same object:  $\{x\} \neq \{\{x\}\} \neq \{\{\{x\}\}\} \dots$ .

temptation of considering (PE) a self-evident truth; for example, the following, almost embarrassingly simple argument threatens the principle:

- 1) (PE) implies that Omero the dog is identical to the thing that is the sum of its body and tail, call it Oremo.
  - 2) Omero the dog can survive the loss of its tail.
  - 3) Oremo cannot survive the loss of tail.
- C) So, Omero and Oremo are not identical *via* Leibniz's law, so (PE) is false.

However, as Casati and Varzi observe<sup>26</sup>, this may just well be an instance of a schema of defective arguments that build on the ambiguity of premise (3). In fact, (3) can be understood as a *de dicto* claim asserting that in any possible circumstance, the thing that is body+tail must have the tail as a proper part – which is a logical truth –, or as a *de re* claim, stating that the thing there on the carpet, which is actually composed by Omero's body and tail, must have the tail as a proper part in any possible circumstance – which does not imply anything about the identity expressed in (1). In other words, the argument cannot rule out that “Omero” and “Oremo” are indeed co-referential expressions, so (2) will be true just in case (3) is – true, it would be inappropriate to call *that* “Oremo” after the loss of the tail, just as it is wrong to use “the current US president” to refer to George Bush<sup>27</sup>.

Another sort of suspicion for (PE) comes directly from complexity science, the scientific field where CA were born (isn't it ironic?). As A.I. father Herbet Simon put it, ‘in such systems the whole is more than the sum of the parts’<sup>28</sup>. This and similar slogans are very much used in the field of complexity to characterize the emergent patterns displayed by complex systems. There are two reasons why ‘the whole is greater than the sum of the parts’ should not be taken seriously from a metaphysical point of view: the first, which is non-technical, is that scientists are interested in predicting *behavior*, so they just mean that the whole *looks* greater than the sum because you cannot anticipate global behavior of CA just by knowing the fundamental dynamics; as a matter of fact, Simon himself nicely explains this point of view: ‘given the properties of the parts and

---

<sup>26</sup> See for example Casati, Varzi (1999), p. 42.

<sup>27</sup> Of course, this does *not* solve the puzzle: does Omero survive the loss of his tail (or any other part)? We shall come back to the issue of mereological change in the next chapter.

<sup>28</sup> Simon (1962), p. 468.

the laws of their interaction, it is not a trivial matter to infer the properties of the whole<sup>29</sup>; the second, which is more technical, is that patterns *supervene* on the lattice of cell: there cannot be a difference in patterns without a difference in the underlying lattice - and if this is true, there are reasons to suspect that patterns may indeed be completely reducible to their supervenience base<sup>30</sup>.

As a second addition, we add closure principles (*Sum* and *Product*)<sup>31</sup> to the effect that:

PC.1) If two things underlap, there is a smallest thing of which they are parts.

PC.2) If two things overlap, there is a largest thing that is a part of both.

Even if (PC.1) may seem harmless, it has been contested on Ockhamian grounds: if (PC.1) is coupled with an axiom to the effect that any two things underlap (such axiom is the *Universe* axiom and we are going to introduce it in a moment), it implies that whenever you got, say, two objects (two bottles of beer), you actually got yourself another one (object, not beer) for free – the mereological sum of the two beers:

PC.1) For any two objects, there is a smallest thing of which they are parts.

Put in another words, there is a disagreement on the *counting* policy: whenever the enemy of (PC.1) sees  $n$  objects, the friend sees  $2^n - 1$  objects. While for some this disagreement is just a symptom of the fact that there is no definite, objective answer to the question ‘how many objects are there?’<sup>32</sup>, others takes this as a genuine metaphysical dispute.

A first reply available to the supporters of (PC.1) is that mereological sums are actually “no more reality” than one already accepts by including basic objects in her ontology<sup>33</sup>:

---

<sup>29</sup> Simon (1962), p. 468.

<sup>30</sup> We shall come back to supervenience and emergence in the next chapter.

<sup>31</sup> From a technical point of view, it may be worth noting that these axioms can be replaced by schemata that allow for infinitary operations (which of course are not needed for finite digital universes).

<sup>32</sup> See, for example, the arguments in Putnam (1982).

<sup>33</sup> This is discussed in literature as the alleged ontological “innocence” of mereology; see Lewis (1991), p. 81.

Given a prior commitment to cats, say, a commitment to cat-fusions is not a *further* commitment. The fusion is nothing over and above the cats that compose it. It just *is* them. They just *are* it. Take them together or take them separately, the cats are the same portion of Reality either way.<sup>34</sup>

The above suggestion can be taken to imply that composition literally *is* identity: however, this strong claim has been heavily criticized by many philosophers<sup>35</sup>; on the other hand, one can also read the “cats vs. cat-fusions” issue as a problem of epistemic relevance. According to this milder perspective, common sense counting is not “wrong”: we count things as we do because (not surprisingly) we are somehow hard-wired to select mostly middle-sized, self-connected pieces of Reality as objects worth of our cognitive attention; but the fact that some pieces are more psychologically *salient* than others should not mark a metaphysics distinction: what does it *mean* to say that the object composed by my left thumb and Michael Jordan’s right hand exists in a less metaphysically deserving way than, say, my laptop? What is the difference supposed to be?

In recent years, the so-called Minimalist program<sup>36</sup> tried to bridge the gap between mereology and common sense on the whole counting issue: while it is true that cats and cat-fusions are *strictly speaking* different objects, we can draw a non-redundant, yet complete inventory of the world by choosing one or the other. According to Minimalism, any satisfactory inventory may include atoms of a table *or* the table itself, but not both: in a slogan, you should include an object *x* if and only if *x* does not overlap with some other object already in the inventory. It is hard to understand the real import of the suggestion: once it is acknowledged that, “strictly speaking”, the domain of quantification includes both cats and cat-fusions, the only sensible reading of the Minimalist idea seems epistemic (i.e.: any satisfactory, non-redundant, complete “carving of reality” should satisfy the constraint, but of course the *real*, unrestricted catalogue of the world should not). We leave aside this topic for now, but we shall come back to the general issue of counting and metaphysical commitment in *Chapter 2* and *Chapter 3*.

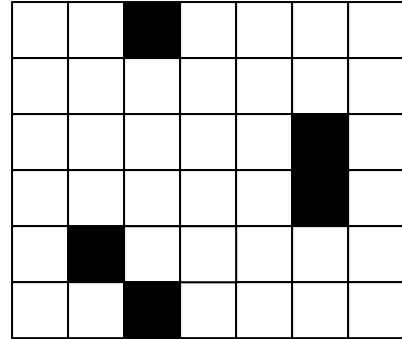
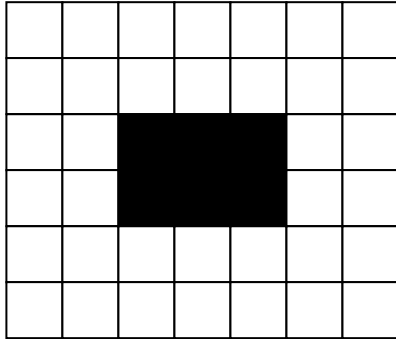
---

<sup>34</sup> Lewis (1991), p. 81.

<sup>35</sup> See for example van Inwagen (1994).

<sup>36</sup> See, for example, Varzi (2000).

We can push further this intuition playing directly with our *Life* universe. Consider the mereological sum of the black cells in the two situations below:



Even granting that the first sum is cognitively more “salient”, that does not mean that the second is less real: why should the first be counted as a full-fledged object but not the second?

A second independent reply is also available to the friend of (PC.1), a direct argument by David Lewis to the effect that there is no reasonable restriction to mereological composition. The argument runs as follows<sup>37</sup>:

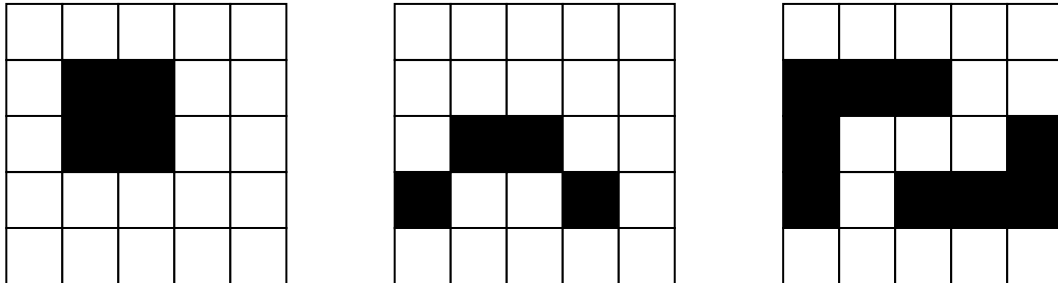
- 1) Any restriction to (PC.1) is bound to be vague, i.e. it implies the existence of situations where it is *vague* whether composition obtains.
- 2) The relevant vagueness in (1) is either semantical or metaphysical.
- 3) The vagueness cannot be semantical, since the language of mereology is *not* vague.
- 4) *Therefore*, the vagueness in (1) is metaphysical.
- 5) *But* metaphysical vagueness is absurd.
- C) *So*, no restriction to (PC.1) is acceptable.

The argument is certainly sound, but it also relies on certain premises which are a bit technical and not self-evident – let us go through it using once again *Life* for illustrative purposes. When stating (1), Lewis has in mind the “standard” restrictions to mereological fusion, such as ‘things should be topologically and/or functionally

---

<sup>37</sup> See Lewis (1986b), pp. 211-213.

connected to be part of a fusion’<sup>38</sup>. If we accept this restriction, we will have cases such the one below:



Given that the cells in the first picture are “connected enough” to be part of a bigger object (the black block constituting their fusion), what about the cells in the second? Are they still “connected enough”? And what about the third picture? One may be tempted to think that restricting composition to cells that are *alive* and *connected* would solve the problem: however, dead cells play a fundamental role in *Life* dynamics and cannot be ignored when pointing, for example, to oscillators, whose characteristic behavior is not captured by connected, alive cells. In the end, it seems that no matter how you specify the restriction, any proposal built to fit our intuitive *desiderata* is bound to create cases where it is vague whether composition obtains. So, what is the matter with *vagueness*?

The study of vagueness is an hot topic in the philosophy of language and metaphysics<sup>39</sup>; vague statements are in fact an essential part of our everyday talk: we use vague predicates (Jim is tall, but how tall is *tall*?), vague nouns (I love climbing mountains, but where exactly do *mountains* begin?), vague proper names (the Sahara is part of Africa, but what are the precise boundaries of the *Sahara*?). From the philosophical perspective, two main strategies are available to explain the phenomenon<sup>40</sup>: the first account attributes vagueness to our *words*, the second to our *world*. According to the semantical account, the world is made of perfectly determined entities: while there are no vague objects nor vague properties, our language has not evolved with sufficient precision to rule out vague statements. So, when we say that ‘Sahara’ is vague we are actually saying that there are *many* possible, massively

<sup>38</sup> See for example van Inwagen (1990) for the suggestion that only objects that are ‘part of a life’ can be mereologically summed together.

<sup>39</sup> For an excellent philosophical introduction see Keefe (2000).

<sup>40</sup> This is indeed an oversimplification: an increasingly popular third option is the *epistemic view* exposed in Williamson (1996).

overlapping, candidates in the world as referents for that name: each candidate is a perfectly respectable, definite desert, but our linguistic practices never required us to choose *one* as *the* referent for ‘Sahara’; in other words, vagueness reflects our indecision to choose a non-arbitrary denotation for our words. According to the metaphysical account, sometimes the world itself contains vague entities, such as clouds: we have precise words denoting truly not-definite objects. In this perspective, when we analyze the vagueness of the sentence containing ‘Sahara’<sup>41</sup> we actually consider the referent of the word as being *one* particular object that is, in some sense, indefinite. With this notion at hand, we can now close Lewis’ argument by showing the truth of (3) and (5).

As for (3), it should be clear from our discussion that if vagueness is a problem with words, we *need* different, equally respectable, possible referents for our words to generate vagueness: however, as we have seen introducing mereology, we can state the theory using only perfectly definite predicates<sup>42</sup> – so (3) must be true. Surveying the immense debate on the thesis stated in (5) goes far beyond the scope of this section<sup>43</sup> and it is indeed unnecessary, since one could easily point out that in digital worlds, after all, every object is perfectly defined (as in a computer memory any bit is either definitely 1 or 0). Since no compelling argument has been found against the two closure principles (PC.1)-(PC.2), we are going to include them without further ado in our theory.

Before leaving mereology, there is a couple of orthodox axioms that need to be discussed:

U) There is a maximal element of which everything is part.

N) There is a null element which is part of everything.

*Universe* is certainly a sound axiom for *Life*: as we have seen, any simulation of *Life* will have a constant, finite number of atomic cells whose fusion may well be regarded as the *universe*. To the contrary, we shall *not* adopt the *Null Element* axiom since it

---

<sup>41</sup> Assuming, for the sake of exposition, that the Sahara may indeed be considered one such indefinite entity.

<sup>42</sup> When all the axioms will be formally stated in first-order logic in later sections, this point will be more evident. However, as argued in Sider (2001) pp. 125-132, (3) may be considered a *petitio principii*: a more careful formulation of the same point can be obtained within a language whose primitives are only logical symbols.

<sup>43</sup> The starting point of the discussion is the (in)famous argument by Gareth Evans, Evans (1978). See also Lewis (1988) for further discussion.

appears metaphysically ungrounded (the axiom is in fact included in mereological systems mainly for algebraic reasons<sup>44</sup>).

As our final mereological principle we list what is actually a distinctive point of CA geometry, i.e. its atomistic structure:

A<sub>def</sub>) Any object is *atomic* iff it has no proper parts.

AT) Everything is ultimately composed by atomic objects.

### 1.2. Toward a CA Proto-Geometry: Topology

There are important notions of a “proto-geometry” that are still not captured by the theory developed so far, the topological concept of *connection* being the most important: since we are willing to accept (PC.1), we are left with no predicates to distinguish between scattered, gerrymandered objects and “good”, self-connected, wholes.

As usual, we start by laying down the lexical core of our topological primitive<sup>45</sup>:

TL.1) Everything is connected to itself.

TL.2) If one thing is connected to another, then also the latter is connected to the first.

Of course (TL.1)-(TL.2) are uncontested, but how exactly are we to understand their relationship with (PL.1)-(PL.3)? The first and most obvious suggestion is that *parthood* is a form of monotonicity, so:

TL.3) If one thing is a part of another, everything connected to the first is connected to the second.

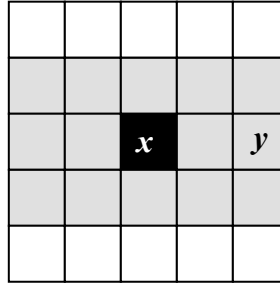
(TL.3) doesn’t look troublesome: since Nevada is part of the U.S.A., everything connected to Nevada (Utah, California, Arizona, etc.) is connected to the U.S.A.. The converse of (TL.3) is however much more controversial; assuming it is tantamount to

---

<sup>44</sup> See, for example, the considerations in Varzi (2009), Section 4.

<sup>45</sup> Generally speaking, *this* topology bears only a vague resemblance to standard point-set topology. For a nice and gentle introduction to orthodox topology, see Flegg (2001).

reducing mereology to topology – a spectacular move from the point of view of conceptual economy, but a right one? While some argued that the move is indeed warranted<sup>46</sup>, it does not fit our digital universe, as the following picture clearly shows:



In this counterexample, everything connected to  $x$  is connected to  $y$  even though  $x$  is not part of  $y$  – pretty much as everything connected to Vatican City is connected to Italy, but Vatican City is not (at least *officially*) part of Italy. Therefore, whatever the merits of the proposal in specific settings, we are not going to assume the converse of (TL.3) in what follows.

Armed with (TL.1)-(TL.3) we can define *enclosed* and other important topological predicates:

$E_{\text{def}}$ ) One thing is *enclosed* in another iff everything connected to the first is also connected to the second.

$IPP_{\text{def}}$ ) One thing is an *internal proper part of* another iff the first is a proper part of the second and everything connected to the first overlaps the second.

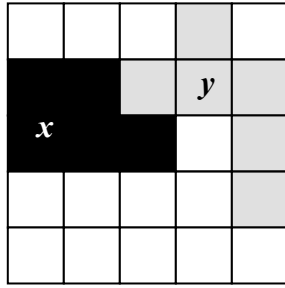
$TPP_{\text{def}}$ ) One thing is a *tangential proper part of* another iff the first is a proper part of the second and something connected to the first does not overlap the second.

$SC_{\text{def}}$ ) One thing is *self-connected* iff any two parts that make up the whole of it are connected to each other.

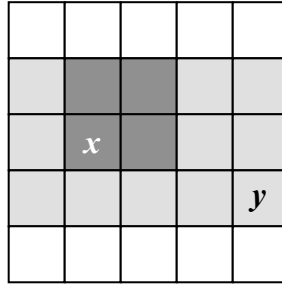
The enriched vocabulary allows us to finally express more complex spatial relationships between entities inhabiting *Life* world:

---

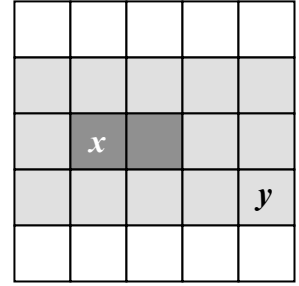
<sup>46</sup> See for example De Laguna (1922) and Clarke (1981). See also Casati, Varzi (1999), pp. 64-66 for general misgivings about the strategy.



*x is connected to y\**  
*x is self-connected*  
*y is self-connected*



*x is connected to y\**  
*x is self-connected*  
*y is self-connected*  
*x is enclosed in y*  
*x is a tangential proper part of y*



*x is connected to y\**  
*x is self-connected*  
*y is self-connected*  
*x is enclosed in y*  
*x is an internal proper part of y*

### 1.3. The Meaning of Life

As we have seen, the proto-geometry developed so far allow us to express a quite rich set of properties and relations between entities in *Life*. However, the theory lacks any expressive power as far as *objects* themselves are concerned: not only it is impossible to describe *gliders* (since moving objects require explicit reference to time, which will be introduced in the next chapter), it is impossible to talk about individual cells as well. When we introduced the automaton in *Section 1.0* we did not list predicates (and corresponding axioms) governing the peculiar features of *Life*: on the one hand, we need to talk about cells being alive/death; on the other, it would be useful to characterize the concept of *neighborhood* of a given cell.

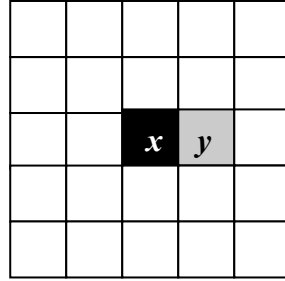
As it turns out, we already have in our theory a predicate that is equivalent to *being a cell*, that of *being atomic* expressed by ( $A_{\text{def}}$ ), so we just need to introduce the properties of *being alive* and *being dead*. Since we are now only concerned with “snapshots” of the world, we start by illustrating the basic lexical fact about the two properties:

LL.1) *Being alive* and *Being dead* exclusively and exhaustively define each cell's state.

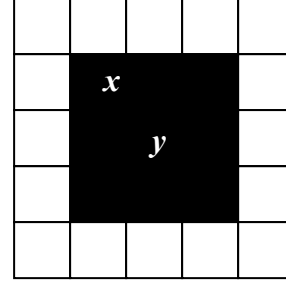
In other words, each cell is either black *or* white, but not both. Finally, we can appreciate the power of our proto-geometry by seeing how easy it is to define the neighborhood of a given cell:

$N_{\text{def}}$  One cell is another's neighbor iff they are connected.

$NH_{\text{def}}$  A cell's *neighborhood* is the mereological sum of its neighbors.



*x is neighbor of y*



*x is the neighborhood of y*

To further constrain *Life* topology, we also add the following:

NA) Each cell has exactly nine neighbors.

(NA) constraints the topology in a non-obvious way: since the universe is finite, we can get a model with (NA) only by “wrapping up” the edges of the grid (in literature, this is called “periodic boundary conditions”). The notion of *neighbor* let us easily introduce other standard topological notions; even if they are not very important in a digital setting, we list them here for the sake of completeness<sup>47</sup>:

$I(s)_{\text{def}}$  The interior of a region (any sum of cells) is the sum of the cells whose neighborhood is itself in the region.

$CL(s)_{\text{def}}$  The closure of a region is the sum of the cells whose neighborhood overlaps the region<sup>48</sup>.

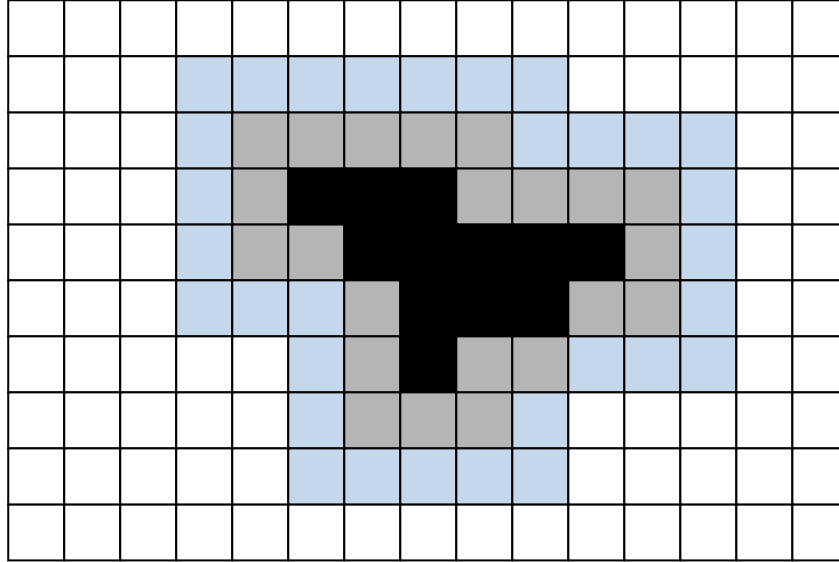
$X(s)_{\text{def}}$  The exterior of a region is the sum of all the cells not in the region.

$B(s)_{\text{def}}$  The boundary of a region is the sum of the cells that are in the closure but not in the interior.

<sup>47</sup> In particular, only some of the usual equivalences and properties hold for the digital counterparts of point-set topology: for example, idempotence does not hold for digital closures. See the comments in Galton (2000), pp. 92-93.

<sup>48</sup> If you allow numerals in the theory you can use nested closures to characterize the distance between a given region and an external cell, and the interior operator to quantify the “bulkiness” of a region. See the suggestions in Galton (2000), p. 94.

These notions can be better appreciated with the following visualization, where black cells are the interior of the region, blue cells are the closure, white cells the exterior, gray cells the boundary:



As it should be clear, the definition of boundaries in a digital universe is pretty straightforward and does not involve any of the subtleties needed for dense spaces<sup>49</sup>.

The ontological importance of the concept of *neighborhood* in *Life* can be further appreciated by noting the following fact: even if – following the “orthodox” approach – we have been using *connection* as the main topological notion, a closer look at  $(N_{\text{def}})$  reveals that we could have instead just taken *neighbor* as the primitive (a reflexive, symmetric relation holding between *atoms*) and defined *connection* accordingly<sup>50</sup>:

$C_{\text{def}}$ )  $x$  and  $y$  are *connected* iff there is one  $x$ -atom which is the *neighbor* of one  $y$ -atom.

Obviously enough,  $(C_{\text{def}})$  implies that *connection* is symmetric and reflexive, as it should be.

<sup>49</sup> For a discussion of the “boundary paradox”, see Casati, Varzi (1999), pp.74-75.

<sup>50</sup> Interestingly, this holds true in *Life* where the technical notion of *neighbor* matches the intuitive one. Generally speaking, however, arbitrary neighborhood’s structures will not warrant the reduction; moreover, we shall see in Chapter 2 that adding a temporal dimension makes the reduction impossible.

#### 1.4 Questions and Answers

Q<sub>1</sub>) The theory developed so far is too weak to be interesting: how can it be extended to deal with objects and events, that is, things extending *through time*? More importantly, some of the principles you use may turn out invalid if you are considering objects, not just spatial structures!

A<sub>1</sub>) There is one basic ingredient missing from the framework to be considered a *minimal* theory of objects: we need a way to express relations between objects and time, since the current theory only applies to single “snapshots” of *Life* – we shall deal with this extension in the next chapter. However, before moving on, it may be useful to address here two arguments proposed by Kit Fine to prove the inadequacy of extensional mereology to represent the qualitative structure of material objects. The first objection goes as follows<sup>51</sup>: according to the theory, a ham sandwich is the mereological sum of two slices of bread and the ham; but mereological sums exist *whenever* any of the components exists, so the ham sandwich existed even *before* I actually made it by placing the ham between the slices of bread – which is absurd, so standard mereology must be rejected. Even if it is true that we have not discussed mereological sums and time indexes, nothing in the reply to the argument really depends on technicalities. So, what is wrong with Fine’s intuitive reasoning? According to our theory, whenever you have three objects in your domain you can take their mereological sum: so, at  $t_0$  you have an object, call it  $X$ , that is the mereological sum of the two slices and the ham, wherever they may be in the fridge. From the fact that  $X$  is an object it does *not* follow that  $X$  is a ham sandwich, because being a sandwich involves certain properties that  $X$  does not have. When, at  $t_n$ , you “make the sandwich” you are just spatially rearranging the objects composing  $X$ : *nothing* comes into existence – it is just that now the elements are such that their sum,  $X$ , may be rightly called a sandwich.  $X$ ’s history is timelessly present in space-time whenever its parts are, but only in some portions of this history  $X$  is a ham sandwich.

The second argument is more directly related to the temporal aspect of the parthood relation. Standard mereology – Fine argues – analyzes ‘The carburetor is now part of my car’ as ‘The current time-slice of the carburetor is timelessly part of the current time-slice of the car’. More generally ‘given any two objects  $x$  and  $y$  that exist at a given

---

<sup>51</sup> See Fine (1999), p. 62.

time  $t$ , we may say that  $x$  is a part of  $y$  at  $t$  if the time-slice of  $x$  at  $t$  is a timeless part of the time-slice of  $y$  at  $t$ .<sup>52</sup> Natural as the picture may look, it runs quickly into troubles:

Consider the aggregate of the current time-slice of the carburetor with Cleopatra or, if you like, with all things whatever that do not currently exist. The current time-slice of this monster aggregate is the same as the current time-slice of the carburetor. Given that the carburetor is currently part of the car, it follows, on the present account, then the monster is as well. But surely this is absurd. How can an object that contains Cleopatra as a part - not to mention all past and future galaxies - currently be a part of my car?<sup>53</sup>

The first thing to note is that the analysis described by Fine is by no means part of standard mereology, at least in the strict, formal sense of *this* work: no axioms adopted so far make any commitment whatsoever to time-slices of objects or the analysis of temporal parthood; therefore, even if Fine's suggestion is indeed natural in the context of contemporary use of mereology in metaphysics, there are no reasons to think that other "time-sensitive" extensions of the theory would look very different from that one<sup>54</sup>.

Let us try to look at the argument from a different perspective: if someone asked 'what are the parts of the car at  $t$ ?' (that is 'what are the timeless parts of the time-slice  $t$  of the car?'), an obvious answer would be: the  $t$ -slice of the steering wheel, the  $t$ -slice of the wheels, the  $t$ -slice of the carburetor, the  $t$ -slice of the pistons, etc.. What would the inclusion of Fine's monster add to this list? Ontologically speaking, nothing: with or without it, the portion of reality that is singled out by the list would be exactly *the same* – in a digital universe, this amounts to say that the set of atoms we select would be the same. So, saying that the monster is part of the car does not add anything to the claim that the carburetor is part of the car: we may note here a departure from common-sense, no more severe than the disagreement on counting that classical mereology already accepts – since it was no reason to reject the theory before, the same reasoning also applies here. In other words, innocence in a general extensional atomistic mereology is

---

<sup>52</sup> Fine (1999), p. 64.

<sup>53</sup> Fine (1999), p. 64.

<sup>54</sup> See for example the discussion in Hovda (*forthcoming*).

not to be found in some double-counting policy (as per the Minimalist suggestion), but in the supervenience of properties of complex objects upon properties of their *atomic parts* only. According to this view, the monster is part of the car but it does not contribute, metaphysically, to the properties of the car (if not for the properties of the carburetor) – and that is exactly what justifies the sloppy way of counting things typical of common sense (i.e. ‘count only relevant parts of an object’).

In the light of these considerations, Kathrin Koslicki’s assessment of Fine’s arguments is surprising:

I take these two considerations to be fatal for the standard conception of mereology as it applies to ordinary material objects.<sup>55</sup>

But Fine’s arguments do not really show any more counterintuitive consequences of standard mereology than what supporters already acknowledge; moreover, we have seen that, by adding a topological layer, we can make the needed distinctions without touching the extensional, neat core of the basic theory of parts and wholes.

Q<sub>2</sub>) How is the theory related to the *real* world?

A<sub>2</sub>) Granted, the proto-geometry we are building is still pretty weak compared to other theories in the market: taking *Parts and Places* as a reference, our theory does not contain the concept of *location*. Why? For a start, a theory of location needs to be developed in a world where objects *occupy* regions of space, but it is not a primitive relation in a CA, where objects *collapse* into regions of space. How bad is that? Casati and Varzi list basically two reasons to believe that objects and regions should be completely distinct ontological kinds: first, some entities appear to be located in other entities’ spatial regions (angels, shadows, ghosts<sup>56</sup>); second, events and objects share the same spatial locations, but they are nonetheless distinct things (say, Cesar’s body and its death<sup>57</sup>). Whatever the merit of the first argument, it does not apply to *Life* for obvious reasons; the second argument, instead, points out an important feature of *Life*: in CA the distinction between objects and events blurs – we may say that in CA you

---

<sup>55</sup> Koslicki (2008), p. 75.

<sup>56</sup> Casati, Varzi (1999), p. 17.

<sup>57</sup> Casati, Varzi (1999), p. 17.

have “processes”, dynamical patterns spreading in the space-time which are then *seen as* objects or events depending on their qualitative features. For professional philosophers this is less exotic than what it seems at first: hardcore four-dimensionalists actually believe that material objects are basically well localized events whose temporal parts are qualitatively homogenous. Of course, this is a serious metaphysical assumption that *Life* forces upon us from the start – one that many philosophers would not like: we shall come back to the issue of generality when addressing the problem of objects and persistence.

Q<sub>3</sub>) Ok, but what *if* the theory turns out to be incomplete in other respects?

A<sub>3</sub>) These misgivings provide us with a chance to discuss one of the rationale behind the whole CA-and-philosophy project. Notwithstanding the efforts in producing a theory as general as possible by using CA, *if* indeed Hamlet persuades us that there are more things in heaven than are dreamt of in our *Life*, the project’s results will still be interesting: at the deepest level of metaphysical seriousness, we are not concerned with the truth of the proposed picture, but with its *tenability*. There are striking analogies between *this* project and the *Humean Supervenience* (hence HS) thesis as proposed by David Lewis, ‘the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another’<sup>58</sup>: at its bottom, the world is just the arrangement in the space-time of point-size objects with *perfectly natural* properties<sup>59</sup>. HS is a *supervenience* thesis since it states that anything else (e.g. modality, causality, laws of nature, moral values, mental states) supervenes on this arrangement. When Lewis put forward HS he was actually proposing an attitude, not just a metaphysical claim:

What I uphold is not so much the truth of Humean supervenience as the *tenability* of it (...). When philosophers claim that one or another common-place feature of the

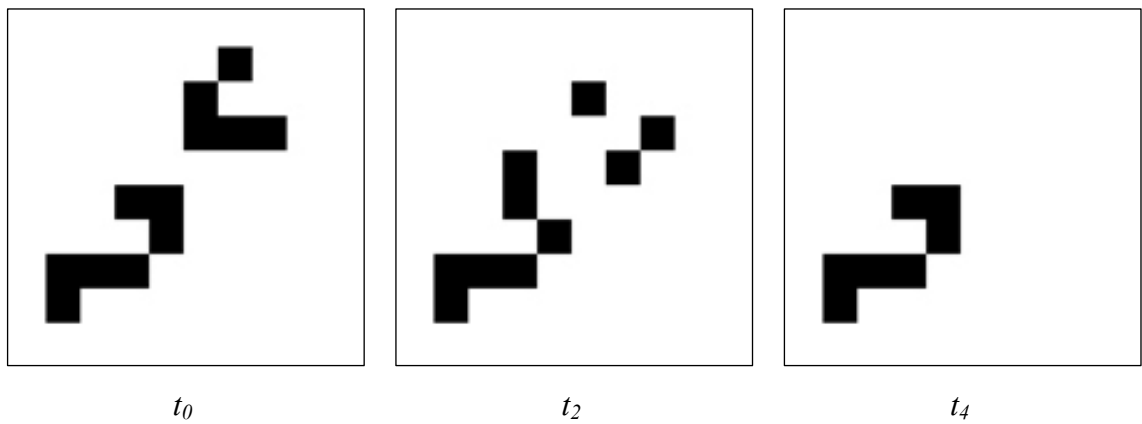
---

<sup>58</sup> See Lewis (1986a), p. ix.

<sup>59</sup> Natural properties are those properties that somehow “carve nature at the joints”. They are what grounds resemblance between things that have them. Naturalness comes in degree: some physical properties are perfectly natural, but many other properties can be considered natural to some extent. For a lengthy discussion of the role of natural properties in metaphysics, see Lewis (1983b): interestingly enough, the explicit formal framework of this project could be also use to test Lewis’ idea on how to quantify the naturalness of properties (see Lewis (1999), p. 66).

world cannot supervene on the arrangement of qualities, I make it my business to resist.<sup>60</sup>

The very sparse ontology of HS can readily be compared to the primitives of *Life*: spatio-temporal relations between simple objects instantiating perfectly intrinsic properties. Sure, there are differences<sup>61</sup>, but our enterprise is close enough, in letter and in spirit, to deserve the name of *Wolframian Supervenience* (hence WS). Consider the following *Life* scene:



An *eater* is devouring a *glider*, i.e. a mereologically complex entity is *causing* the death of another mereologically complex entity – and this phenomenological miracle is happening in a world whose basic primitives do not contain complex objects, persistence, causality. If someone claims that one of these notions cannot supervene on the primitive ontology of WS (and who the hell put it there, since we *programmed* the world ourselves?), we make it our business to resist: even if it turns out that *some* concepts cannot be expressed in *Life*, it is still impossible to believe that understanding complexity, persistence, causality within this framework may not bring fruitful results. In other words, just as we have said that non-philosophers may wish to interpret *this* work as an attempt to teach computers how to reason about a special domain, professional philosophers may wish to interpret it as a *computational, testable* version of the Humean Supervenience hypothesis. In particular, *Chapter 3* is entirely devoted to

---

<sup>60</sup> See Lewis (1986a), p. xi.

<sup>61</sup> For example, Lewis' space-time is continuous and it is not clear what the laws of nature are in a CA - see Berto, Tagliabue (2012) for some preliminary considerations. We shall come back to this comparison in later chapters.

actually *implement* a computational ontology and discuss in depth the methodological issues the project raises.

Finally, there is another related worry that we have not answered: recently, Floridi (2009) and Floridi (2011) claimed that ‘digital ontology’ is a categorical mistake, since the digital/analogue distinction is a consequence of our conceptualization, not a property of reality *per se*. Berto, Tagliabue (*forthcoming*) is a thoroughly argued answer to Floridi’s worries: the interested reader is encouraged to read *Appendix II* to further explore this debate.

Q<sub>4</sub>) What about the *finite* nature of the model?

A<sub>4</sub>) Cardinality is always a tricky issue in first order theories: indeed, those (like myself) who naively thought that atomistic mereology could somehow escape the curse of *Lowenheim-Skolem*-like results, are going to be disappointed. As proved in Hodges, Lewis (1968), there is no purely mereological formula in atomistic theories that may distinguish between finite and infinite models. Since we have been stressing the importance of finitude to get consistency between digital universes implemented in finite computers and the formal, armchair theory we are developing, this limitative theorem is unwelcomed news: there is just no way of saying, in full generality, that the universe *is finite*. However, since nothing important (from a formal point of view<sup>62</sup>) is gained or lost by specifying a particular size for the universe, we shall use the symbol # to denote an arbitrary large integer representing the number of atoms in the domain – as in the following *Finite Nature* axiom:

FN) There are # atoms.

---

<sup>62</sup> As we shall see in *Chapter 3*, the number of atoms makes a *huge* difference in the practical implementation of the theory.

## 2. Time and Other Dimensions

### 2.0 The Time of our Life

Because things are the way they are,  
things will not stay the way they are.

*Bertolt Brecht*

Up until now we have only dealt with snapshots of *Life*, developing an ontology tailored to single instants of the universe evolution. However, it is definitely *change* that makes *Life* interesting (pun intended): introducing a temporal dimension in the framework is the task of this section.

As highlighted at the end of *Chapter 1*, there is no real difference between objects and events in *Life* space-time: we shall introduce a new concept, *process*, to deal unambiguously with temporal entities in our universe. Naturally extending the ideas developed so far, a first tempting definition of a *process* is that of a mereological sum of cells existing at a given timespan. To make the definition more precise we need to define *i*) what conditions cells must satisfy to be included in the summation (e.g.: are dead cells eligible to be part of a process?); *ii*) what temporal restrictions are imposed over *timespans* (e.g.: are non-contiguous instants allowed in a timespan?). However, a much more challenging (and potentially far-reaching) modification is needed *before* addressing these issues: since there is nothing in our proto-geometry allowing references to timespans, how should the basic theory be changed to accommodate this further dimension?

Given that atomic cells are the building blocks of *Life* it is no surprise that the first issue to address is exactly the persistence of cells: are cells bi-dimensional (they occupy space but travel through time) or three-dimensional (they occupy space and time in the same fashion) objects (echoing the more familiar metaphysical discussion of three-dimensionalism vs. four-dimensionalism)? A second problem concerns the reality of the past (and future): at the end of the universe, can we directly quantify over what happened at the “Big Bang” (echoing the familiar metaphysical discussion of presentism vs. eternalism)? Leaving the second dilemma to *Chapter 3*, the first issue boils down to

the following question: do we want to tamper with the logic of simple predication? Let us suppose a bi-dimensional account of cells, i.e. let us say that cell  $c$  “wholly” exists at time  $t$  and  $t + 1$ ; furthermore, let us say that  $c$  is alive at  $t$  but dead at  $t + 1$ . Given that the cell is one and the same, the only way to make sense of this change is to say that the predicate ‘isAlive’ is indeed a two-places predicate, and one of the place is occupied by a time instant, i.e. ‘ $c$  isAlive at  $t$ ’. Applying the same reasoning to other predicates in the theory, we end up modifying mereology accordingly, so that *parthood* becomes a three-places relation and crucial axioms, like extensionality, should be changed in non-obvious ways (e.g., are two objects identical if they have the same parts at the same times?). The alternative is to embrace a three-dimensional account of cells, according to which any cell exists just at one time and have its properties timelessly; in the case of temporal predication, ‘ $c_t$  isAlive’ and ‘ $c_{t+1}$  isDead’ would describe the previous scenario, where  $c_t$  is not identical to  $c_{t+1}$  but has the same “coordinates” in the lattice. Given that a three-dimensional account does not require any modification to the formal machinery developed so far and that, computationally speaking, the two ideas are almost equivalent<sup>63</sup>, we shall stick to a three-dimensional account of change in *Life*<sup>64</sup>. The formal theory should therefore include axioms for time and time instant:

TIL.1) No time instant precedes itself.

TIL.2) If  $t$  precedes  $t'$  and  $t'$  precedes  $t''$ , then  $t$  precedes  $t''$ .

TIL.3) If  $t$  and  $t'$  are distinct, either  $t$  precedes  $t'$  or  $t'$  precedes  $t$ .

IS<sub>def</sub>)  $t'$  is the *immediate successor* of  $t$  iff  $t$  precedes  $t'$  and  $t$  does not precede any other instant preceding  $t'$ .

IP<sub>def</sub>)  $t$  is the *immediate predecessor* of  $t'$  iff  $t'$  is the immediate successor of  $t$ .

TIF) There is an instant with an immediate successor but no immediate predecessor (i.e. the first instant) and there is an instant with an immediate predecessor but no immediate successor (i.e. the last instant); any other instant has both an immediate predecessor and an immediate successor.

---

<sup>63</sup> In fact in both cases we need to introduce indexes in the computational semantics to allow for temporal predication: in one case indices relate to different time-dependent properties inside a cell, in the other to different cells living in different instants.

<sup>64</sup> There is also another reason to prefer the account: given the similarities between *this* theory and *Humean Supervenience*, accepting bi-dimensionality requires us to consider probable natural properties (in the sense of Lewis (1983b)) as extrinsic properties.

All the above axioms should be obvious<sup>65</sup>: time in *Life* is a strict linear ordering; moreover, in each simulation, time is finite, since there is a first and a last time instant separated by finitely many instants (TIF). If we adopt a three-dimensionalist perspective there are just some more axioms that we need to add:

(CT) Any atomic cell exists at one instant of time.

(CT) guarantees that each cell has a fixed, “eternal” temporal collocation.

Moreover, we need to update the Finite Nature axiom to reflect the new temporal dimension of the model: instead of requiring the existence of # cells in the lattice, we now require the existence of # cells for each  $t$ .

(FN) For each  $t$ , there are # atomic cells at  $t$ .

(FN) guarantees that, no matter how many instants of time there are, there is always a “full lattice” in each of them.

Finally, we need to introduce a new notion to capture the idea that, at each instant of time, the updating rule does not change a cell, it sets eternally a new state. A first approximation may be the following:

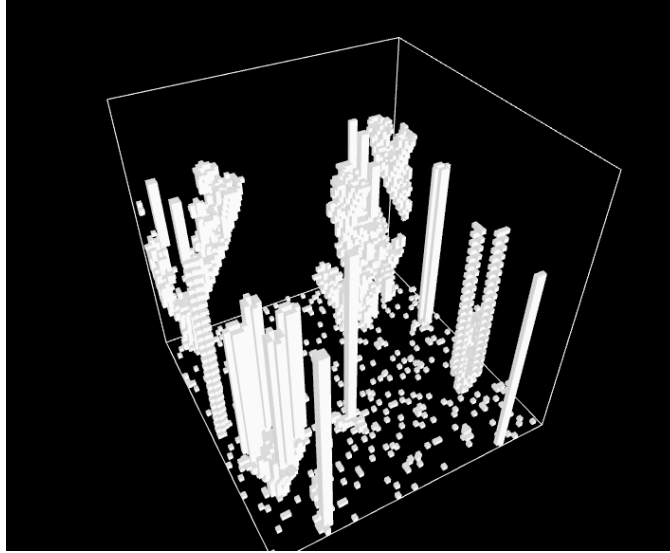
(TCS) Any atomic cell has one *immediate predecessor/successor* (except for cell living at the last/first instant of time).

(TCS) assures that, for each cell existing at  $t$ , there is one “companion” cell at  $t - 1$  and one at  $t + 1$ : intuitively, if we use  $Cell_{x,y,t}$  to designate a cell in the three-dimensional universe, (TCS) states that for every pair  $\langle x, y \rangle$  in the lattice and every time instant  $t$ ,  $Cell_{x,y,t}$  has  $Cell_{x,y,t+1}$  as successor and  $Cell_{x,y,t-1}$  as predecessor. This constraint is crucial when we consider the dynamics of *Life*, since nothing, in the theory developed so far, makes the state of  $Cell_{x,y,t}$  dependent upon  $Cell_{x,y,t-1}$  and its neighborhood, as it should be in a three-dimensional *Life*. For the moment, we rely on an intuitive grasp of the notions *predecessor* and *successor*: we shall come back to (TCS) in the next section.

---

<sup>65</sup> Compare, for example, Galton (2000), pp. 215-226.

Before looking in the detail at processes, it could be useful to observe a three-dimensional version of *Life* to visually appreciate the complexity of the space-time as the universe evolves<sup>66</sup>:



What are the processes of this world? As we have said in the introduction of this section, we need to specify two kinds of restrictions to narrow down the intended interpretation of our concept: first, we need to say if any mereological summation is admitted, second, we need to say if some temporal restrictions over timespans apply. We can frame this problem using insights from Ted Sider's Diachronic Composition Question<sup>67</sup>:

Given a class of time instants,  $I$ , and a function  $f$  assigning a non-empty class of objects,  $f(t)$ , to each  $t$  in  $I$ , under what conditions will there be an object,  $x$ , that exists exactly at the times in  $I$  and that at each such time  $t$  is composed exactly by the objects in  $f(t)$ ?

If such an object existed, it would be the mereological composition of the fusions specified by the function  $f$  for the instants in  $I$ . Sider's own conclusion is that  $x$  will exist in any condition whatsoever; as a matter of fact, Sider's argument is the diachronic version of the Lewis' argument for unrestricted composition we met in *Chapter 1*: since

<sup>66</sup> Obviously, the fact that, for drawing purposes, different cells represent different time instants is irrelevant for the metaphysical question of the persistence of cells though time (picture generated with a custom *NetLogo* model).

<sup>67</sup> See Sider (2001), Section 4.9. For an excellent discussion of the argument, see Varzi (2005a).

before we had no reasons to abandon unrestricted composition, by the same reasoning (i.e., that any restriction would amount to be *vague* and digital universes are precise entities) we should not abandon it now. The upshot of the discussion is thus that *processes* are indeed just mereological fusions: once we extend to the diachronic case the principle of unrestricted composition, any “three-dimensional worm” in the universe is in the domain of quantification, so it is eligible to be a process. Of course, for the sake of completeness, we could introduce a special class of processes/objects, *alive/dead objects*, by “extending” the property of *being alive* to aggregates:

GA) An object is alive/dead iff every atomic part of the object is alive/dead.

Processes have boundaries in space *and* time. Luckily, our three-dimensional approach enables us to extend the mereotopological notions of interest to temporally extended objects, insofar as we update our conception of *neighbor* to reflect the new dimension: while in *Chapter 1* we reduced the topological notion of *connection* to the CA-based notion of *neighbor*, we are no longer allowed to make this spectacular move in a temporal setting. In particular, we should introduce a separate notion of *connection*, intuitively holding between any two atoms in the lattice that are “in contact”: in the case of contact between atoms living at the same  $t$ , *connection* and *neighbor* will still be related (since two atoms will be connected just in case they are neighbors); however, we also have contact between atoms living at different times – in that case we can introduce the notion of *temporal connection*, i.e. connection without being neighbors.

TC<sub>def</sub>)  $x$  existing at  $t_x$  and  $y$  existing at  $t_y$  are *temporally connected* iff an  $x$ -atom is connected to an  $y$ -atom and one of  $\langle t_x, t_y \rangle$  is the immediate successor of the other.

N<sub>def</sub>) Cell  $x$  existing at  $t_x$  is *neighbor* of cell  $y$  existing at  $t_y$  iff  $x$  is connected to  $y$  and  $t_x = t_y$ .

By interpreting *neighbor* has a special case of connection, we can preserve much of the theory we already developed: the neighborhood once again is the mereological sums of neighboring cells, the mereotopological operators work as expected, and the concepts

defined in *Chapter 1* can be easily applied to the new dimension (e.g. process A and process B are connected if an A-atom and a B-atom are connected<sup>68</sup>).

Finally, we wish to point out that introducing time into the equation highlights another very important issue, i.e. the interplay between semantics and ontology. The language we use to talk about the world is, of course, constrained by the structure of the world, so that, for example, the domain of quantification of a semantics for *Life* should reflect the mereological choice we made on the ontological side. However, when we go beyond spatial reasoning and introduce a temporal dimension in the semantics, it turns out that the link with ontology is not as strict: given an ontology we all agree upon, there are different ways to setup a semantics mimicking everyday temporal predication (e.g. ‘ $x$  is  $F$  at  $t$ ’). In other words, given the ontology we just sketched, it is highly non-trivial to establish how parts of the language (names, predicates) are linked to parts of the world: if we wish to include temporal predication in our computational semantics, such semantic subtleties cannot be ignored.

## 2.1 Laws of Nature

Gravity is not responsible for people falling in love.

*Albert Einstein*

Adding a temporal dimension requires us to specify what is the relation between different “slices” of our three-dimensional universe; in other words, we should finally add axioms to make the state of the lattice at  $t$  dependent upon the state of the lattice at the previous time instant. The first notion we need is the concept of *temporal predecessor/successor* for atoms: using the handy notation  $Cell_{x,y,t}$  as we did above,  $Cell_{x,y,t}$  has  $Cell_{x,y,t+1}$  as successor and  $Cell_{x,y,t-1}$  as predecessor; the notion plays a crucial role in stating the laws of *Life*, since the basic CA rule is defined as a function from the neighborhood of  $Cell_{x,y,t}$  to the state of  $Cell_{x,y,t+1}$ . To define this concept in our framework, we shall use *temporal connection* as introduced in the previous section: a necessary condition for  $x$  at  $t$  being the *immediate successor* of  $y$  is that  $x$  and  $y$  are temporally connected and that  $y$  exist at  $t + 1$ ; however, to obtain sufficient conditions we also have to rule out  $Cell_{x+1,y,t+1}$  as the immediate successor of  $Cell_{x,y,t}$ , so we will

---

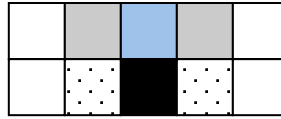
<sup>68</sup> Of course, at this point we may wish to further distinguish between *temporal* and *spatial* connection.

require that every neighbors of  $y$  is temporally connected to  $x$ : with this constraint,  $Cell_{x+1,y,t+1}$  will no longer count as immediate successor, since it has a neighbor (e.g.  $Cell_{x+2,y,t+1}$ ) that is not a neighbor of  $Cell_{x,y,t}$ . Putting the two conditions together we obtain:

ICS<sub>def</sub>)  $x$  existing at  $t_n$  is *the immediate successor* of  $y$  existing at  $t_m$  iff  $t_n$  is the immediate successor of  $t_m$  and every neighbor of  $x$  is temporally connected to  $y$ .

ICP<sub>def</sub>)  $x$  existing at  $t_n$  is *the immediate predecessor* of  $y$  existing at  $t_m$  iff  $y$  is the immediate successor of  $x$ .

The following picture illustrates the point in two-dimensions (the vertical dimension represents consecutive time instants,  $t$  and  $t + 1$ ):



The blue cell in the first row is  $Cell_{x,t}$ , the gray cells are the neighboring cells (i.e.  $Cell_{x+1,t}$  and  $Cell_{x-1,t}$ ); intuitively, we want our definition to select the black cell (i.e.  $Cell_{x,t+1}$ ) to be the immediate successor of the blue cell, ruling out the two cells with black dots (i.e.  $Cell_{x+1,t+1}$  and  $Cell_{x-1,t+1}$ ). Let us take for example  $Cell_{x-1,t+1}$ : the cell passes the first test, since it is temporally connected to the blue cell and lives at  $t + 1$ ; however, the cell does not pass the second test, since one of its neighbor (i.e. the white cell at its right) is *not* temporally connected to the blue cell. By analogous reasoning it can be shown that  $Cell_{x+1,t+1}$  cannot be the immediate successor of  $Cell_{x,t}$ , leaving us with just the black cell satisfying the conditions, as required. On the basis of (ICS) we can now state the basic rule of *Life* in our framework:

LN) If  $x$  is the immediate successor of  $y$ ,  $x$  is alive  
     if  $y$  is alive and two or three neighbors of  $y$  are alive, or  
     if  $y$  is dead and three neighbors of  $y$  are alive;  
      $x$  is dead otherwise.

(LN) is the mereotopological equivalent of the updating rule we introduced in *Chapter 1*: we did not make any substantive assumption about the metaphysics of laws, but a comparison with some philosophical accounts may be interesting.

According to a popular account, a law is the holding of a particular relation between universals:

Suppose it to be a law that  $F$ s are  $G$ s.  $F$ -ness and  $G$ -ness are taken to be universals. A certain relation, a relation of non-logical or contingent necessitation, holds between  $F$ -ness and  $G$ -ness. This state of affairs may be symbolized as ' $N(F,G)$ '.<sup>69</sup>

Whatever the merit of the proposal, the theory can hardly explain why (LN) is a law unless we are willing to introduce in the framework a whole new set of dark creatures – universals and some obscure relation of “contingent necessitation”. A more promising approach is the so-called “Best System Theory”<sup>70</sup> (henceforth BST): according to BST, laws in a world  $w$  are all the true generalizations contained within the *best* deductive system systematizing the particular facts in  $w$ . The rationale behind the idea is that we can summarize all the facts about a world using several deductive systems: some will be very powerful (e.g. the system including as axiom each particular fact will have full deductive power), others very simple (e.g. the simple including ' $2 + 2 = 4$ ' will be very simple). These two virtues compete: among all the deductive systems, the best is the one achieving the best combination of simplicity and power; all the true generalizations of this system will be the laws of nature. Of course, when we think about our world is really hard to judge if BST is extensionally adequate (i.e., if it classifies as “laws” all and only the generalizations that we regard as laws): but what happens in a digital world such as *Life*? Let us start with the obvious candidate for a deductive system (call it  $S_0$ ): a sentence describing the state of the world at  $t_0$  and a conditional like (LN) specifying the updating rule; by definition, ( $S_0$ ) is powerful enough to let us deduce all the “atomic facts” (i.e. facts like cell  $x$  being alive at  $t_x$ ); moreover, ( $S_0$ ) is also very simple – just two axioms. *If* this is the best system describing *Life*, (LN) will be counted as a law of nature, as desired. Now, let us consider a universe where at  $t_0$  all cells are dead: the evolution of the system, for any  $t$ , is obvious: the lattice will remain unchanged no

---

<sup>69</sup> Armstrong (1983) p. 85.

<sup>70</sup> See Ramsey (1978), Lewis (1973a), Lewis (1983b), Lewis (1994).

matter what. For this universe we may think of an alternative to  $(S_0)$ ,  $S_1$ : the first and only axiom of this system is ‘For any  $t$ , any cell is dead at  $t$ ’.  $(S_1)$  is very powerful, since every atomic fact can be deduced;  $(S_1)$  is also very simple, being composed by just one universal generalization. Given that both  $(S_0)$  and  $(S_1)$  are maximally powerful, the best system will be the simpler, that is  $(S_1)$ : however, according to  $(S_1)$  (LN) is *not* a law of nature. This example vividly illustrates a fundamental problem of BST: what laws there are in a digital universe depends on the initial conditions<sup>71</sup>. It does not matter that the computational level underlying a dead universe is based on (LN); it doesn’t matter that we programmed that world ourselves: contrary to our own source code, BST is telling us that (LN) is *not* a law of this universe. One option, as always, is just biting the bullet: maybe we are so ignorant about laws in digital universes that when we think we are coding laws we are actually doing something else. As a corollary of this view, the definition of worlds nomologically accessible from a given *Life* universe becomes unnatural: intuitively, given a world  $w$  of size  $s$ , we would consider all the worlds with size  $s$  as nomologically possible relatively to  $w$  (i.e. as worlds with the same laws of nature); however, if we follow BST, this simple definition is not available, since the dead universe and a randomly generated universe will *not* have the same laws of nature. The consequence is that establishing what is nomologically possible relatively to a digital world  $w$  becomes a very hard question even when there is no modal ignorance at all (i.e. we can know what is going on in all the possible alternatives).

Of course, there is another option, withdrawing BST and looking for something else outside the mainstream of philosophy; for example, a tentative solution would be to exploit the *locality* of digital universe and define the laws of nature starting from some sort of dispositional properties of cells<sup>72</sup>. Developing such a theory in detail goes beyond the scope of *this* work: it is however important to point out that “local” solutions may work for digital universes such as *Life* but are not suited for worlds (such as, probably, ours) with non-local fundamental dynamics.

---

<sup>71</sup> For a seemingly similar point in the context of a more general discussion on BST, see Woodward (2007), pp. 290-292.

<sup>72</sup> This view will be perfectly isomorphic to the computational simulation of *Life* developed in *Chapter 3*. However, it will still be undefined what is “dispositional” about the computational properties thus coded.

## 2.2 Layers

Whether there is such a thing as Reality,  
of which the various levels are only partial aspects,  
or whether there are only levels,  
is something that literature cannot decide.  
Literature recognizes rather the reality of the levels.  
*Italo Calvino*

As we have seen at the end of *Chapter 1*, the “ontological resolution” we use to capture *Life* phenomenology may greatly vary depending on the scope of our description – sometimes we want to talk about cells and aggregates, sometimes we want to talk *directly* about gliders, so we need somehow to include in the framework a handy way to meet this requirement. Many philosophers would happily subscribe to the idea that our world is basically the sum of different *layers* of objects, where higher-order entities are formed through mereological composition from lower-level ones: atoms make molecules, molecules make cells, cells make tissues, tissues make organs, organs make bodies. In the canonical formulation by Kim, the “fundamentalist picture” of reality is thus the following:

The Cartesian model of a bifurcated world has been replaced by that of a layered world, a hierarchically stratified structure of “levels” or “orders” of entities and their characteristic properties. It is generally thought that there is a bottom level, one consisting of whatever microphysics is going to tell us are the most basic physical particles out of which all matter is composed (electrons, neutrons, quarks, or whatever).<sup>73</sup>

While some have questioned the correctness of this picture for *our* universe<sup>74</sup>, it is indubitably a nice way of summarizing *Life*: cells compose basic patterns (say, gliders) which in turn compose more complicated patterns and so on. We can make more precise this intuition if we think that each “zooming out” is basically a restriction of the domain of discourse through mereological composition; in other words, the mereological

---

<sup>73</sup> Kim (1993) p. 337.

<sup>74</sup> See for example Shaffer (2003) or Ladyman, Ross (2007).

operation of summation is a mapping from a bigger (lower) domain to a smaller (higher) one, so that, when asserting ‘everything is white’ in front of a white lattice, we can take the quantifier to be ranging over atoms only; when asserting ‘there is an object’ in front of a glider moving in an empty universe, we can take the quantifier to be ranging over a domain made just by the glider.

Let us use the concept of *layer* informally to describe the “zooming levels” of *Life*: the claims we make when focusing on different layers *presuppose* the domain of the layer in question. More precisely, we will say that a *chain* is an ordered  $n$ -tuple of  $n$  layers (sets of elements) defined as follow:

$L_0$ ) The domain of atomic cells in *Life* universe (the cellular layer).

$L_n$ ) The domain obtained by mapping entities in the domain of  $L_{n-1}$  into a smaller non-empty domain. The mapping is such that each entity in  $L_{n-1}$  is mapped into one and one only entity in  $L_n$ .

Given the nature of *Life* universe, the chain will always be finite. In particular, we can single out a special layer, the “Eleatic layer”:

$L_u$ ) The domain obtained by mapping entities from any domain into the universal domain, i.e. the one with just one individual, the universe.

We say that an *admissible chain* is any chain of  $n$  layers starting with the cellular layer and ending with the Eleatic layer. The constraint on the function assures that each part of the lower layer is somehow assigned a new object in the higher domain, and no part of the lower layer is assigned to different higher level objects: apart from this, the mapping is completely unconstrained, which is something expected given that the mapping is conceptually a *mereological* operation and our *mereology* does not restrict composition in any way. In other words, the same apparatus may be adopted in different settings by adjusting the mapping so that it reflects different principles of summation<sup>75</sup>. Now that we have made some progress in defining *layers*, it is crucial to remember that layers do not actually *generate* objects, if not in a metaphorical way: given the atoms in

---

<sup>75</sup> Incidentally, we note here that each layer satisfies the conditions of the Minimalist program as stated in *Chapter 1*.

our universe and our mereology, we *already* have a domain with  $2^n - 1$  objects, the biggest one being the universe itself<sup>76</sup>. What layers do is to partition this huge set of objects in smaller and smaller domains that can be used to effectively *carve* reality<sup>77</sup>.

Before going further in connecting layers with the basic theory, we briefly address for the first time the “problem of emergence”, which has been mainly discussed in connection with properties (emergent *objects* are presumably such because they exhibit emergent properties). At a first approximation, we may say that *emergent properties* are non-obvious *global* properties arising from the local behavior of smaller entities – paradigmatic cases include the functioning of the brain, the wealth of a nation, the life of ant colonies. Of course, as we have already seen over and over in previous sections, CA show the phenomenon of emergence in its purest form: in *Life* all sorts of global pattern are in fact generated as a deterministic result of low-level local interactions. When speaking about emergence, it is very easy to conflate an epistemic version of emergence (what David Chalmers calls “weak emergence”<sup>78</sup>) with a metaphysical version of emergence (“strong emergence”). The crucial difference between the two is *supervenience*: if high-level properties *supervene* on low-level properties, it is only a weak case of emergence. Let us take a familiar example for illustrative purposes: biological properties are the high-level result of chemical properties; in particular, the chemistry of carbon-based chemical compounds is the basic brick of life on Earth. Are biological properties weakly or strongly emergent? To answer this question we should find out if biological properties *supervene* on chemical properties, which in turn requires us to answer the following question: two possible worlds, exactly identical in the arrangement of chemical properties, can possibly differ for the arrangement of biological properties? Or, in Kripke’s words<sup>79</sup>, once God has created all the chemical properties, is there any additional work to get the biological properties or do they come, so to speak, for free? The intuition here is pretty clear: once the chemical layer is in place, biological properties are no longer free to vary, since the upper layer is entirely constrained by the lower one. As a result, this is a case of *weak* emergence: while, in

---

<sup>76</sup> The theory of *layers* overlaps to some extent the theory of granular partitions in formal ontology, as put forward, for example, by Smith and Brogaard (2002) and Bittner and Smith (2003).

<sup>77</sup> We shall come back to the varieties of realism supported by the theory in *Chapter 3*.

<sup>78</sup> See Chalmers (2002).

<sup>79</sup> See the famous argument in Kripke (1980), p. 153.

some sense, biological properties are “unexpected”, “surprising” features of reality, they are indeed completely ontologically dependent upon chemical properties<sup>80</sup>.

Considering digital universes, it is clear that any object emerging from the cell-by-cell dynamics is only *weakly* emergent: in fact, there can be no difference in the emergent objects without a difference in the atomic level. Moreover, it looks trivial, given an admissible chain of layers and a non-atomic layer  $l$ , to define the objects in the  $l$ -domain as emergent with respect to layer  $l-1$ ; however, different chains would specify different emergent objects from the same emergent base, with no non-arbitrary way (ontologically speaking) to favor one chain over another. This choice can be made on epistemic bases, thus recovering the importance of some emergent objects in our understanding of *Life* dynamics: for example, one may wish to prefer layers recognizing objects such as *gliders* (that give us some predictive power over the world dynamics) instead of random regions of space-time. Is there some way to make this intuition more robust? We shall tackle this difficult question in the next section.

As stated above, *layers* are not part of the ontology of *Life*, but are “semantic shortcuts” allowing us to readily specify a group of objects and treat them as one. As such, we do not wish to include them in the specification of the formal theory regarding *Life*; a theory of *layers* pertains much more to a theory of semantical evaluation, as a theory of contextually implicit restrictions on the domain of quantification. In *Chapter 3* we shall sketch such a theory in the context of developing a working computational semantics for digital universes.

### 2.3 Emergence Reconsidered

We can only see a short distance ahead,  
but we can see plenty there that needs to be done.  
*Alan Turing (not necessarily about this work)*

As we emphasized in the previous section, from a purely metaphysical perspective the concept of emergence within our mereological framework is void of any substantive import. However, we can still address the problem of weak emergence from an epistemological perspective: is there a way to capture the fact that a layer of gliders is

---

<sup>80</sup> It is indeed disputed that there are real cases of strong emergence; most bets go to qualitative properties of mental states (qualia) – see for example Chalmers (1996).

epistemically different from a layer of randomly selected space-time region? Dan Dennett was the first philosopher to address this problem in the context of *Life* (in his language, the “reality of patterns”). Before explaining Dennett’s view, we briefly comment on two other discussions of emergence in digital universes put forward in recent years.

Bedau (1997) and Hovda (2008) develop a technical notion of emergence with the explicit goal of capturing the interesting case of emergence that scientists study every day: given the reducibility in principle of high-level properties to low-level ones, is there any way to make precise the “unpredictability” intuition (recall the quote from *Chapter 0*, ‘We get macro-surprises despite complete micro-knowledge’<sup>81</sup>)? Not surprisingly, they both take *Life* as their favorite example of a complex system exhibiting emerging features; while differing in details, the intuition behind both accounts is the following: a fact *F* about *Life* is emergent if and only if *F* can be deduced only by explicitly simulating the system dynamics. The reader, now familiar with CA, would indeed recognize the appeal of the proposal: given the system complexity (and its equivalence with a Universal Turing Machine) many facts about, say, gliders can only be established by simulating the universe step-by-step, cell-by-cell. However, while it may well be a philosophically adequate theory, it falls short of achieving its stated goal: this definition makes emergent facts scientifically *uninteresting*. What is scientifically interesting and practically useful (whenever the feature can be exploited) about complex systems is that emerging facts are somewhat stable, regular and fairly predictable: you can predict the choices of human beings without bothering in neuron-by-neuron computations, you can predict the properties of a city without bothering in head-by-head calculations and you can predict economic output of a nation without bothering in city-by-city simulation. In the context of *Life*, you can predict a glider path without computing the cell-by-cell updating rule: in other words, emergent features are not interesting because they cannot *always* be predicted *perfectly* without low-level calculation, but because they can *usually* be predicted with *good accuracy*. Bedau (1997) and Hovda (2008) cannot distinguish between gerrymandered emergent facts about *Life* and facts about gliders: in most cases, all these facts are just *emergent*, since they can be deduced by simulation only. However, facts about gliders are interesting for

---

<sup>81</sup> Epstein (1999), p. 48.

prediction and they can be exploited with some epistemical gain, other emergent facts are totally uninteresting for the scientific description of *Life*.

The pioneering Dennett (1991) made this point abundantly clear: if we describe the universe with *gliders*, we adopt a “design stance”, so that we can forget about the low-level dynamics and use generalizations to predict the evolution of the lattice – what is lost in precision (inferences are not deterministic), it is gained in speed: computing the path of a glider is much easier than compute the updating rule for all the cells in the universe<sup>82</sup>. In Dennett’s view, layers supporting efficient prediction of *Life* dynamics (by exploiting some important emergent features) are epistemically superior to layers which leave us with nothing better than the cell-by-cell calculation: notwithstanding the problem of defining “efficient prediction”, Dennett’s account is a very good refinement of the scientific notion of emergence<sup>83</sup>.

As a final note, Dennett (1991) got the scientific issue right, but the metaphysics completely wrong: in particular, the “predictability” criterion is used by Dennett as his own restriction to mereological composition (in his language, the answer to the question ‘When does a pattern exist?’), i.e. a pattern exists in some data if there is a description of the data that is more efficient than the bit map (whether or not one can concoct it)<sup>84</sup>. Given our favorite mereology, no restriction to the principle of composition is possible: by overlooking foundational, mereological considerations, Dennett<sup>85</sup> drew unwarranted metaphysical conclusions from epistemic principles.

## 2.4 Possible Digital Universes

But if there is a sense of reality (...)  
 then there must also be something  
 we can call a sense of possibility (...).  
 If he is told that something is the way it is, he will think:  
 well, it could probably just as well be otherwise.  
*Robert Musil*

---

<sup>82</sup> See Dennett (1991), pp. 39-41.

<sup>83</sup> The problem of predicting a glider position after  $n$  instants, in fact, seems to exhibit polynomial scaling both with a “design stance” algorithm and with a cell-by-cell algorithm: therefore, the usual polynomial vs. exponential distinction cannot be invoked here. However, in the context of one-dimensional CA, the framework of computational mechanics (as developed for example in Crutchfield (1994)) looks like a promising line of research (to our knowledge, nothing similar has been yet applied to *Life* – see Miller and Page (2007), pp. 233-234). *Appendix III* contains a brief description of the approach.

<sup>84</sup> See Dennett (1991), p. 34.

<sup>85</sup> Along with many commentators – see for example the discussion of “Rainforest Realism” in Ross (2000).

Ordinary objects have modal properties: I could have been a continental philosopher, Mitt Romney could have won the 2012 presidential election, my laptop can survive the replacement of its memory. While our everyday life is literally *full* of modal talk, providing a systematic account of modal properties (in both semantics and ontology) has been proved to be a daunting task. Almost all accounts on the market use the notion of “possible world” to analyze necessity and possibility: *de dicto* sentences like ‘There could have been unicorns’ are analyzed as ‘There is a possible world  $w$  such that at  $w$  there are unicorns’; *de re* sentences like ‘Jacopo could have been a continental philosopher’ are likewise analyzed with possible individuals, i.e. ‘There is a possible world  $w$  such that Jacopo exist at  $w$  and Jacopo is a continental philosopher at  $w$ ’<sup>86</sup>. While the dimension of time in *Life* is pretty obvious (and time itself has been represented elsewhere in works of formal ontology<sup>87</sup>), the space of possibilities is much less so. Intuitively, what we want is some sort of “principle of plenitude”, such that every possible way the lattice can be, our ontology can represent the lattice that way. A natural strategy to achieve the desired result is therefore to build possible worlds in a combinatorial fashion:

PW) For any lattice configuration  $c$ , there is a world where the lattice configuration is  $c$ .

(PW) guarantees the truth of ‘There could have been three gliders in the universe’, since, thanks to the ricombination, there is a possible world  $w$  such that at  $w$  there are three glides. However,  $c$  may be interpreted in several ways (from the most to the less restrictive):

i)  $c$  is interpreted as a whole three-dimensional universe of a given size, extending in space and time: possible worlds are generated by switching alive/dead cells to dead/alive cells in any possible combination throughout the universe history. Since every possible world will have the same number of atoms per instant and the same duration, this version of (PW) makes ‘The universe could have been one instant longer’ false. Also, for any given actual universe, the class of possible worlds will be finite.

---

<sup>86</sup> Divers (2002) is an excellent, in-depth discussion of possible worlds in contemporary philosophy.

<sup>87</sup> See for example the extensive treatment in Galton (2000), pp. 205-251.

ii)  $c$  is interpreted as the two-dimensional lattice configuration at  $t_0$  of a universe of a given size: possible worlds are generated combinatorially by varying the initial conditions. Every possible world will have the same number of atoms per instant, but different worlds may have different duration; ‘The universe could have been one instant longer’ will be true in this case, but ‘The universe could have had two more atoms at  $t$ ’ will be false. The class of possible worlds thus generated will be infinite (for any number of time instants  $n$ , there is a *finite* world whose duration is  $n$ ).

iii)  $c$  is interpreted as the two-dimensional lattice configuration at  $t_0$  of an arbitrary universe: possible worlds are generated combinatorially by varying the initial conditions *and* the lattice dimension; ‘The universe could have been one instant longer’ and ‘The universe could have had two more atoms at  $t$ ’ will turn out both true. Obviously, the class of finite possible worlds thus generated will be infinite.

A thorough investigation of the proper modal structure of digital universes goes beyond the scope of *this* work. However, two interesting facts deserve some comments: first, even universes obeying the Finite Nature hypothesis may not follow the “Finite Modal Nature” hypothesis, according to which the modal quantities associate with a given finite universe are themselves finite – interpretation (ii) and (iii) above make this abundantly clear. Second, the availability of a modal dimension (along the line of (i)-(iii)) may deceptively suggest that modal knowledge can be “empirically” gained in digital universe – but this is a mistake. On the one hand, evaluating necessity claims (e.g. ‘In every possible universe it is the case that  $P$ ’) with infinite universes may be impossible in a finite time; on the other, even possibility claims (e.g. ‘There is a possible universe  $w$  such that  $P$  is the case in  $w$ ’) may be hard to evaluate. One may be tempted to think that, given a lattice with  $n$  atoms, all the permutations of states for the  $n$  atoms should be *possible* state of *Life*: however, this is *not* true in general. Due to the existence of “Garden of Eden” configurations, i.e. configurations that may appear only as initial conditions for the universe but that cannot have predecessors, any statement of the form ‘There is a possible universe where configuration  $X$  is the lattice configuration at time  $t_1$ ’ is painstakingly hard to verify. Moreover, it is important to remember that *Life* is a

Universal Turing Machine: as such<sup>88</sup>, there is no way to know with certainty the lattice configuration after  $n$  time instants without explicitly simulating the universe evolution, from the Big Bang up to  $t_n$ .

The upshot of these preliminary considerations is that, even in a finite world where (as we shall see in the next chapter) the evaluation of complex sentences can always be effectively computed, modal knowledge is not trivial nor is it easily attainable: the problem – which surely deserves more attention than what has received here – is not the lack of the famous “telescope for possible worlds”, but the fact that, given the telescope *and* the possible worlds, it is computationally expensive to observe in detail what happens there.

## 2.5 Questions and Answers

Q<sub>1</sub>) If objects are just three-dimensional mereological sums, does this mean they cannot gain/lose parts?

A<sub>1</sub>) It depends on what ‘gain and lose parts’ means: given that cells exist at one instant of time, they are, in that instant, part of a process (or they aren’t) – there is no actual *change* involved. Of course, a process *can* gain/lose parts by having different group of cells at different time instants. To appreciate this point let us look at the following one-dimensional example (the vertical dimension represents time as usual):



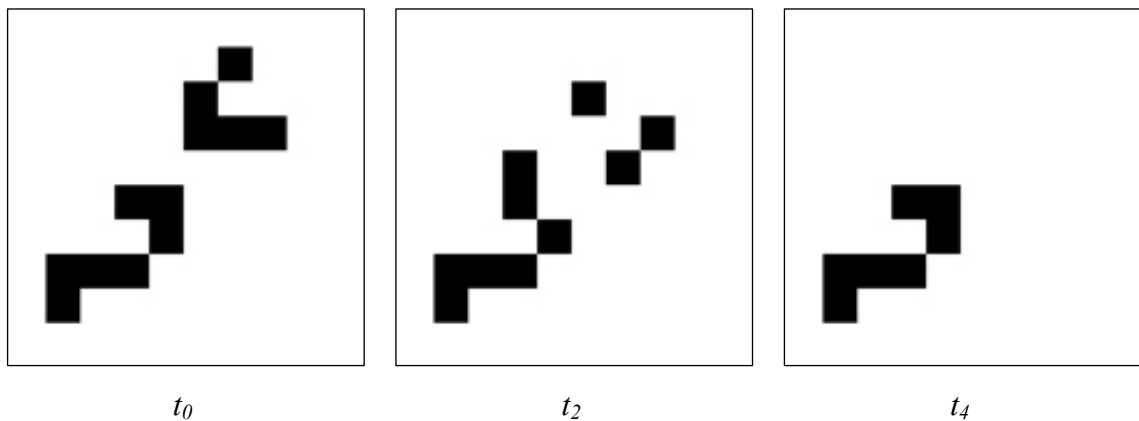
Intuitively, the process composed by the black cells *changes* from  $t_0$  (the first row) to  $t_1$  (the second row) by having at  $t_1$  cells that are *not* immediate successors of the cells at  $t_0$ . In *this* sense, three-dimensional mereological sums can gain/lose parts as many times as possible during their life span. As far as the modal profile of this process is concerned, the theory developed so far does not rule out any possibility: while it is certainly possible to develop a “digital modality” that would make processes modally frozen, others, more flexible accounts are perfectly compatible with the basic mereological framework.

---

<sup>88</sup> For a classical reference, see Turing (1936).

Q<sub>2</sub>) When you discussed laws of nature, you didn't mention *causality*, which is somehow a related, important notion: how does causation enter the picture?

A<sub>2</sub>) The link between laws of nature and causality is disputed in contemporary metaphysics: for example, while Donald Davidson requires that any instance of causation is to be subsumed under strict laws<sup>89</sup>, James Woodward takes a much more liberal stance by connecting causation to *invariance*<sup>90</sup>. The first “anomaly” in *Life* is that (LN) directly applies *only* to atoms: if (LN) is the unique law of the universe, should this imply that causation applies only to atoms? Let us look again at a “causal” episode, an *eater* devouring a *glider*:



There aren't strict laws relating gliders and eaters in *Life*, but gliders and eaters are just sums of atoms. The ontological structure of the problem reminds Davidson anomalous monism: mental events are physical events, but there are no strict laws for mental events: in fact, *Appendix III* is entirely devoted to employ CA as a model for Davidson's metaphysics in the context of the contemporary philosophy of psychology.

As a final consideration, we may note that the popular counterfactual approach to causation<sup>91</sup> is not easily implemented in *Life*. Roughly speaking, C causes E iff in the “closest” possible world where C does not happen, E does not happen (where the “closest” means the most similar to the actual world). As many commentators pointed out in other contexts<sup>92</sup>, the definition of the relevant notion of similarity is a highly non-trivial task: as simple as *Life* is, we can indeed see where the problem is in our case. Let us call  $w$  the lattice at  $t_0$ : which is the closest world where C (*the eater being in such and such a position*) does not happen? Apparently, the closest world  $w_1$  is any world which

<sup>89</sup> See for example Davidson (1970).

<sup>90</sup> See Woodward (2007).

<sup>91</sup> See Lewis (1973b).

<sup>92</sup> See for example the early “Nixon counterexample” by Kit Fine in Fine (1975).

differ from  $w$  by just one bit – just make dead one of the cells composing the eater;  $w_1$  is thus a world where  $C$  does not happen, but – as it turns out simulating the system –  $E$  (*the glider dying*) happens nonetheless, so  $C$  does *not* cause  $E$ . Of course, one may well say that the closest world is not  $w_1$ , but  $w_2$ , i.e. the world where all the cells composing the eater are dead: in  $w_2$   $C$  does not happen and  $E$  does not happen either, as desired.

To settle definitely the matter we would need an independent theory of trans-world identification for eaters; but even without such a theory, there are strong *prima facie* reasons to choose  $w_1$ , namely the fact that just by replacing one bit completely changes the eater's behavior. Of course, we are far from a conclusive account of causality in digital universes: hopefully, this first sketch and the arguments in *Appendix III* are a first step in the right direction.

### 3. Calculemus

#### 3.0 *The Aufbau Digitally Remastered*

What I cannot build, I do not understand.  
*Richard Feynman*

“Computational philosophy” should not sound completely unfamiliar to the sophisticated reader: on the one hand, the use of solvers to discover new proofs of interesting metaphysical claims gave in recent years surprisingly good results<sup>93</sup>; on the other, the constant grow – from niche to mainstream – of philosophy of information brought in the philosophical arena a set of relatively new conceptual tools from computer science<sup>94</sup>. However, the computational philosophy of *this* work is not directly related to any of these research paradigms.

Computational philosophy is, indeed, a particular view of philosophy, its methodology and its place in the world – a view with its roots well before computers and modern information technologies. Gottfried Leibniz wrote:

The alphabet of human thoughts is a catalog of primitive concepts, that is, of those things that we cannot reduce to any clearer definitions.<sup>95</sup>

Rudolf Carnap claimed that all concepts can be defined starting from a single concept and logic alone; under his account, a single world-sentence (which is a conjunction of all the truths about most basic objects of reality) entails all truths about the world. Lewis’ *Humean Supervenience* is an ontological version of this idea, where properties and supervenience take the places of concepts and definability: in particular, all truths about a world  $w$  supervene on the arrangement of the basic properties and relations of  $w$ <sup>96</sup>. Taken together, all these views suggest a specific idea of the role – and the potential impact – of philosophy: it is the business of philosophy to identify these primitive

---

<sup>93</sup> See for example Oppenheimer and Zalta (1991) on Anselm’s Proof.

<sup>94</sup> See for example Floridi (2011).

<sup>95</sup> Quoted in Chalmers (2012), p. 1.

<sup>96</sup> Strictly speaking, Lewis’ theory is a contingent claim about our world and worlds similar to ours. See for example Lewis (1994).

concepts and show how to derive non-primitive concepts – in other words, the final product of an ontological theory is a *generative* process, i.e. a procedure that may be used to derive, from the world-sentence of  $w$ , a virtually infinite number of arbitrary complex statements about  $w$ . If we require that this procedure should be effectively computable (why not?), the upshot is that successful philosophy is computable knowledge; it is therefore no surprise that Leibniz dream was a world where philosophers would settle their disagreements just by *calculating*.

Of course, the Leibniz/Carnap project is widely held to be a failure, for technical and philosophical reasons: *contra* Leibniz, it has been established that it is not possible to derive all arithmetical truths from simple axioms<sup>97</sup>; *contra* Carnap, it has been argued that many of the verificationist principles built into the *Aufbau* are simply conceptually misguided<sup>98</sup>; moreover, no one has the slightest idea of how a world-sentence of *this* world would look like, or what are the properties to be included in the Humean mosaic. That is exactly where digital universes fit naturally in the picture, since in a CA it is clear what the supervenience base is and how the world-sentence looks like; digital universe are also *finite* by definition, so that any problem of effective computability is resolved.

Leibniz and Carnap did not have computers to test this idea in a meaningful way. Can we take advantage of today's technology to better understand and assess their proposals? As we have seen, the approach combines philosophy, logic and computer science<sup>99</sup>: among the many possible ways to make something out of this intuition, we chose to write a computer simulation of *Life* and a computational version of the ontology we developed – we shall call the program *AufByte* in Rudolf Carnap's honor. At a first approximation, *AufByte* should be able to:

- i) simulate *Life* dynamics in an arbitrary lattice and with customizable initial conditions;

---

<sup>97</sup> Limitative theorems such as Gödel's fundamentally constrain the *unrestricted* version of Leibniz project, where any truth whatsoever can be deduced by a set of axioms (see Gödel (1931)). Of course, there is still room for a *characteristica* of the material world and the objects therein.

<sup>98</sup> See for example Friedman (1987).

<sup>99</sup> If the business of philosophy is to make knowledge computable, philosophical theories may become incredibly useful to the growing field of computational knowledge engines (see for example <http://www.venexia.eu> and <http://www.wolframalpha.com>). We shall come back to potential A.I. applications of the *AufByte* idea later in this chapter.

ii) evaluate the truth value of *any* sentence expressible with the predicates of our formal ontology in a given *Life* simulation (e.g. ‘the cell 3 is alive’, ‘there is a glider’, ‘there is something which is part of region z’, etc.).

In other words, given a digital universe *AufByte* will be an omniscient oracle, answering ‘yes’ or ‘no’ (true/false) to any question using its knowledge of the ontological structure of the world. Before looking into the details of the implementation, we first need to introduce the technology we adopted for the task.

### 3.1 Java Coding 101

Talk is cheap. Show me the code.  
*Linus Torvalds*

It is finally time to meet our new best friend: *Processing*, the language (and framework) used to develop the computational part of *this* research project.

*Processing* (henceforth often abbreviated in p5) is an *open source*, *Java based*, *object-oriented* programming language. While hopefully any reader will be familiar with the concept of programming language, we quickly review the other features of p5<sup>100</sup> (even if nothing fundamental really hinges on these details, it is a good practice to know the core properties of the tools we use to better understand their strength and limitations).

*Open source*: p5 was originally developed by Ben Fry and Casey Reas at the MIT Media Lab; since the very beginning the project was designed to be “*open*” – one of p5 stated aim is to be a gentle introduction to programming practices for inexperienced programmers (in particular, visual artists and designers), thanks to the easy syntax and the immediate, rewarding visual feedback. *Open source* means that the source code is freely available to anyone: if you would like, you can take a look at what is really going on behind the scene of p5. Moreover, successful open source projects like *Processing* generally attract talented programmers and create a huge community of followers, fans,

---

<sup>100</sup> There are many good introductory books on *Processing*. Shiffman (2008) and Reas, Fry (2008) are good places to learn the basics; Shiffman (2012) is an excellent intermediate text, with a chapter devoted to cellular automata.

developers; in practice, this means that the constantly growing interest around p5 makes solving coding problems pretty easy: you can learn most of what you need just by reading the official forum, since it is likely that someone around the globe already had *your* problem and shared the solution.

*Java based:* p5 is based on a famous, older language, *Java*. Java is one of the most popular language around – it can be found on cars, music players, mobile phones, international space stations, just to name a few<sup>101</sup>. One of the most known and advertised feature of Java is the ‘Write Once, Run Everywhere’ property: after you wrote a Java program you can just run the *same* code on another hardware and another operating system and it will still work as expected. How is that possible? Java is a partially interpreted, partially compiled language. Compiled languages<sup>102</sup> use other programs (compilers) that translate the code directly into machine executable code – different machines will therefore require different compiled versions (e.g. a program compiled for Windows should be recompiled to run under Linux). Interpreted language are executed step-by-step by an interpreter, without a prior conversion of all the code into low-level machine code: different machines can run the same program given that they both have installed the interpreter. Java makes a bit of both: the original program is compiled into an intermediate language (called bytecode) and then a virtual machine<sup>103</sup> (the Java Virtual Machine) acts as the interpreter between the bytecode and the machine. So, given that your Mac and my PC both have installed a JVM (of course, each system has a devoted JVM), any Java program (and, *a fortiori*, any p5 program) can be executed on the two devices without any change. Finally, it is worth noting that p5 implements easy references to many, but not all Java capabilities (as well as inheriting Java typical data structure and object-oriented nature, see below). However, since it runs on a JVM, you can always add pieces of Java to do more advanced stuff.

---

<sup>101</sup> The educational literature on Java is pretty much endless. Schildt (2011) is a “beginner’s guide” containing a useful historical introduction and comparison with other well-known languages.

<sup>102</sup> This is a simplification. In theory one can have an interpreted version of usually compiled languages and *vice versa*. However, for historical and practical reasons, some languages are almost always compiled/interpreted.

<sup>103</sup> Virtual machines are exactly what the name says, i.e. software programs emulating a physical machine; given Alan Turing original contributions to computability theory (the Universal Turing Machine and the uniformity between input data and program instructions) virtual machines should not be very surprising.

*Object-oriented*: object-oriented programming (also called OOP) is a *programming paradigm*, i.e. it is a way to develop software programs; languages are said to be object-oriented if they support this particular style (of course, a given language may support more than one paradigm<sup>104</sup>). The reader familiar with philosophy may think of OOP as Plato’s programming paradigm (while, one may suggest, functional programming is Church’s paradigm): the building blocks are objects called classes, which are abstractly defined as full-fledged objects with properties (variables) and capabilities (functions). Any time the program needs to introduce an object, a concrete instance of the class is built according to the abstract specifications of that class and inherits its properties and capabilities: at no time (call it a Leibnitian twist to the basic Plato’s idea) other parts of the program can interact with its properties unless specified by its capabilities. For example, a software program manipulating circles will create a Circle class with appropriate properties, like *radius* and *circumference*, and capabilities, like *calculateCircumference*. Any time we need a new circle in our program, we build an instance of our Circle class, say “circle1”, which, of course, possesses all the properties and functions typical of the “idea of Circle” as specified in the abstract class definition. Moreover, in a typical OOP fashion, we would specify that, to create a concrete circle, you need to specify an essential property, i.e. its radius (it does not make sense to speak of a circle without a radius). In our case, we require that any concrete instance, like “circle1”, is created with a given radius, that should be specified upon creation (e.g. ‘circle1 = new Circle(5)’, where 5 is the radius); we also require that the newly created object should calculate its *circumference* using the *calculateCircumference* function (and, of course, the radius). In this way we are sure that there is no way, from *outside*, to change the value of *radius* or *circumference* – everything that can happen to a circle is embedded in the class. In other words, it does not matter what else the program does: as far as circles are concerned, no variable will be changed except by the very functions of the Circle class. Encapsulating parts of the code in (almost) self-sufficient objects has many design advantages: re-usability (you can copy the Circle class to another program and instantly all the properties and functions relative to circles become available), maintenance (if, *per impossible*, the formula for the circumference changes tomorrow, the code will be changed just once, inside that particular function, no matter how many

---

<sup>104</sup> Again, this is a simplification: almost any modern language, despite its historical roots, presents or emulates the design patterns typical of the most important paradigms (imperative programming, OOP, functional programming).

times circles are used in the program – i.e. there is no need to look at thousands of lines and correct them one by one), inheritance (you can define a “child” class of the Circle class, and it will inherit all the Circle features), polymorphism (you can use the same, human readable label for functions that have similar purposes but different underlying mechanisms – e.g. you can define a drawing method *print* for circles and for squares and invoke them in the same way, ‘circle1.print()’ and ‘square1.print()’). Finally, it is worth stressing that OOP matches our ontological intuitions pretty well, since, philosophically speaking, we are already accustomed to think about reality as a collection of token from a set of abstract types, interacting in various ways<sup>105</sup>.

Concluding this brief overview of p5, it is important to note that *Processing* comes with a friendly *integrated development environment* (IDE). An IDE is a software that helps writing programs in a given language, by providing syntax checking, online help functions, compilation of the source code etc.. While in principle more complex IDEs can be used with this language (such as the famous Java IDE *Eclipse*<sup>106</sup>), the simple, intuitive original p5 framework will be perfect for our computational needs.

### 3.2 The Computational Cost of Ontology

W: Beauty often seduces us on the road to truth.

H.: This doesn't bother you?

W.: That you were wrong? I try to work through the pain.

H.: I was not wrong. Everything I said was true. It fit. It was elegant.

W.: So, reality was wrong.

H.: Reality is almost always wrong

*House M.D.*

To talk *about* digital universes, *AufByte* should be able to do two, quite separate, things: first, modeling digital universes, their properties, their dynamics; second, implementing a computational semantics, which, consistently with the ontology, computes the truth value of any statement that can be produced with the language of the formal theory. As we shall see, the first is very easy, the second is very hard.

---

<sup>105</sup> Most manuals for object-oriented languages (Java, C#, etc.) contain an introductory section on OOP in general – see for example Purdum (2012).

<sup>106</sup> The IDE is freely available from the *Eclipse* project website: <http://www.eclipse.org>.

Leaving to the interested reader a thorough inspection of the source code, a *Life* simulation requires two basic classes: *Universe* and *Atom*. *Universe* is a “container” object for the cells in our simulation, *Atom* is the class modeling each cell in the automaton. Running *AufByte* will create an instance of *Universe*, which in turn will create a collection of  $n$  objects of type *Atom*. The CA main logic is then implemented in a classic OOP fashion: at each time step in the simulation, each cell in the universe calls its own *update()* method – *update()* obviously contains the local transition function as specified by the rules of *Life*. Finally, at the end of the updating session, the current state of the automaton is visualized on the screen (again, in a classic OOP fashion, using the *draw()* method) – while, technically speaking, the update of the cells occurs sequentially, the result of each cycle is the same as a synchronous update<sup>107</sup>.

On top of the digital universe, we need to model the computational semantics: as stated in the introduction, our goal is an algorithm that evaluates the truth value of *any* sentence of the ontology in a given *Life* simulation. A first design proposal is to add two software layers to *AufByte*: one layer will manage the computational representation of the semantics, another layer will manage syntactical and semantical aspects of the queries we ask *AufByte* to evaluate (i.e. check for well-formed formulas, simplify predicates through definitions, implement the recursion algorithm for satisfaction of arbitrary formulas). We shall deal with each of the layer in turn.

Given the availability of a thoroughly understood, general mathematical framework for the semantics of first order theories, it seems natural to use computational counterparts of model-theoretic objects to implement the semantics of *AufByte*. We use the *Structure* class as a container class for our model-theoretic entities – a domain of quantification (for which we use a *Domain* class) and the interpretation of non-logical predicates (a function from  $n$ -places predicates to  $n$ -tuples in the domain). Of course, it is exactly at this point that the code should precisely mirror our ontology: in particular, the domain is built, starting from the atomic cells in the CA, according to the specification of mereology (i.e. the resulting domain is closed under the mereological operation of summation); the interpretation of non-logical predicates (like ‘is Alive’) is constrained by the axioms of our theory (i.e. each cell is alive or dead, not both) and by

---

<sup>107</sup> The trick is computationally trivial, but it introduces an inefficiency that is not present in the abstract specification of the CA (namely, the fact that each cell has *two separate properties* storing its state, the current state and the previous state).

the universe under examination (cell  $c$  is in the extension of ‘is Alive’ iff  $c$ , in the universe being simulated, has the property of *being alive*).

The *Formula* class contains all the properties and the methods concerning syntactic validity and satisfaction; in particular, once we ask *AufByte* to evaluate a first-order sentence, a new object of type *Formula* is created in memory. The formula is then tested for syntactic validity (through the methods exposed by the *Language* class, containing the characters for predicates, variables, constants of the theory) and for free variables (variables not bounded by quantifiers). If the formula is closed, *AufByte* may compute its truth value through a standard Tarski-style recursion – as in standard first-order model theory, we use the notion of satisfaction by an assignment (modeled by a special *Assignment* class) to recursively evaluate sentences<sup>108</sup>.

To better understand the general architecture, let us go through an example with a very simple universe:

|     |     |
|-----|-----|
| 0_0 | 0_1 |
| 1_0 | 1_1 |

Once the universe is in place, an instance of *Structure* creates an appropriate domain and interpretation<sup>109</sup>: as in the picture, we shall use the code ‘row#\_column#’ to represent the objects in the domain (the code for non-atomic objects is just the concatenation with ‘&’ of the code of their atoms):

Domain = { 1\_1, 1\_0, 1\_0&1\_1, 0\_1, 0\_1&1\_1, 0\_1&1\_0, 0\_1&1\_0&1\_1, 0\_0, 0\_0&1\_1, 0\_0&1\_0, 0\_0&1\_0&1\_1, 0\_0&0\_1, 0\_0&0\_1&1\_1, 0\_0&0\_1&1\_0, 0\_0&0\_1&1\_0&1\_1 }

Interpretation for ‘=’ = { <1\_1, 1\_1>, <1\_0, 1\_0>, <1\_0&1\_1, 1\_0&1\_1>, <0\_1, 0\_1>, <0\_1&1\_1, 0\_1&1\_1>, <0\_1&1\_0, 0\_1&1\_0>, <0\_1&1\_0&1\_1, 0\_1&1\_0&1\_1>, <0\_0, 0\_0>, <0\_0&1\_1, 0\_0&1\_1>, <0\_0&1\_0, 0\_0&1\_0>, <0\_0&1\_0&1\_1, 0\_0&1\_0&1\_1>, <0\_0&0\_1, 0\_0&0\_1>, <0\_0&0\_1&1\_1, 0\_0&0\_1&1\_1>, <0\_0&0\_1&1\_0, 0\_0&0\_1&1\_0> }

<sup>108</sup> For computational reasons partial functions are used for assignments as suggested by Barwise, Etchemendy (2002) pp. 500-506.

<sup>109</sup> For simplicity we list here just the extension of two predicates, *identity* and *parthood*. It is worth stressing that the actual symbols used for predicates, quantifiers etc. are parameters that can be settled directly in the appropriate *Language* class of the program.

$\langle 0\_1 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 0\_0, 0\_0 \rangle$ ,  $\langle 0\_0 \& 1\_1, 0\_0 \& 1\_1 \rangle$ ,  $\langle 0\_0 \& 1\_0, 0\_0 \& 1\_0 \rangle$ ,  
 $\langle 0\_0 \& 1\_0 \& 1\_1, 0\_0 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 0\_0 \& 0\_1, 0\_0 \& 0\_1 \rangle$ ,  $\langle 0\_0 \& 0\_1 \& 1\_1, 0\_0 \& 0\_1 \& 1\_1 \rangle$ ,  
 $\langle 0\_0 \& 0\_1 \& 1\_0, 0\_0 \& 0\_1 \& 1\_0 \rangle$ ,  $\langle 0\_0 \& 0\_1 \& 1\_0 \& 1\_1, 0\_0 \& 0\_1 \& 1\_0 \& 1\_1 \rangle$  }

Interpretation for 'P' = {  $\langle 1\_1, 1\_1 \rangle$ ,  $\langle 1\_0, 1\_0 \rangle$ ,  $\langle 1\_0, 1\_0 \& 1\_1 \rangle$ ,  $\langle 1\_1, 1\_0 \& 1\_1 \rangle$ ,  
 $\langle 0\_1, 0\_1 \rangle$ ,  $\langle 0\_1, 0\_1 \& 1\_1 \rangle$ ,  $\langle 1\_1, 0\_1 \& 1\_1 \rangle$ ,  $\langle 0\_1, 0\_1 \& 1\_0 \rangle$ ,  $\langle 1\_0, 0\_1 \& 1\_0 \rangle$ ,  
 $\langle 0\_1, 0\_1 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 1\_0, 0\_1 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 1\_1, 0\_1 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 0\_0, 0\_0 \rangle$ ,  
 $\langle 0\_0, 0\_0 \& 1\_1 \rangle$ ,  $\langle 1\_1, 0\_0 \& 1\_1 \rangle$ ,  $\langle 0\_0, 0\_0 \& 1\_0 \rangle$ ,  $\langle 1\_0, 0\_0 \& 1\_0 \rangle$ ,  
 $\langle 0\_0, 0\_0 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 1\_0, 0\_0 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 1\_1, 0\_0 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 0\_0, 0\_0 \& 0\_1 \rangle$ ,  
 $\langle 0\_1, 0\_0 \& 0\_1 \rangle$ ,  $\langle 0\_0, 0\_0 \& 0\_1 \& 1\_1 \rangle$ ,  $\langle 0\_1, 0\_0 \& 0\_1 \& 1\_1 \rangle$ ,  $\langle 1\_1, 0\_0 \& 0\_1 \& 1\_1 \rangle$ ,  
 $\langle 0\_0, 0\_0 \& 0\_1 \& 1\_0 \rangle$ ,  $\langle 0\_1, 0\_0 \& 0\_1 \& 1\_0 \rangle$ ,  $\langle 1\_0, 0\_0 \& 0\_1 \& 1\_0 \rangle$ ,  
 $\langle 0\_0, 0\_0 \& 0\_1 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 0\_1, 0\_0 \& 0\_1 \& 1\_0 \& 1\_1 \rangle$ ,  $\langle 1\_0, 0\_0 \& 0\_1 \& 1\_0 \& 1\_1 \rangle$ ,  
 $\langle 1\_1, 0\_0 \& 0\_1 \& 1\_0 \& 1\_1 \rangle$  }

Going through the interpretation of  $P$  one may find puzzling that, for example, the extension thus defined does not include  $\langle 0\_0 \& 0\_1, 0\_0 \& 0\_1 \& 1\_1 \rangle$ ; in fact, if we count how many objects make 'P(x,  $0\_0 \& 0\_1 \& 1\_1$ )' satisfied, we get three as an answer, which should be surprising given that the ontology accepts unrestricted composition. As it turns out, it is computationally easier to deal with an interpretation of *parthood* that assigns to every object just the atomic cells composing it. How can we respect the axiom of unrestricted composition then? The extensional and atomistic nature of the ontology can be exploited to solve the semantics conundrum: for any pair of objects  $a$  and  $b$ , ' $a$  is part of  $b$ ' is logically equivalent to 'every atom that is part of  $a$  is part of  $b$ '. This trick guarantees that, by properly rearranging formulas before the evaluation, we can get the truth values we want while maintaining a smoother implementation of the semantics.

With this structure safely stored in the memory of the computer, *AufByte* can easily verify that ' $1\_1$  is part of something'<sup>110</sup> is true in the universe, since:

---

<sup>110</sup> For convenience, we use here the label representing the object in the computational domain as its name in the language.

- i) ‘1\_1&1\_0 is part of something’ is logically equivalent to ‘every object  $x$  that is an atomic part of 1\_1&1\_0 is part of a object  $y$ ’.
- ii) there are many  $y$  in the domain having 1\_1&1\_0’s atomic parts (1\_1 and 1\_0) as parts: 0\_0&1\_0&1\_1, 0\_0&0\_1&1\_0&1\_1, etc.

This software architecture has many attractive features: it sharply divides the world (*Life* simulation) from our theory *about* the world (ontology) from our language *for* the theory (semantics), allowing for a very transparent code and, eventually, an easy comparison between competing ontologies<sup>111</sup>; it employs a straightforward computational semantics, where each special class is the natural counterpart of some model-theoretic tool already well understood by logicians, philosophers, computer scientists<sup>112</sup>; it is fully general: although the semantics is naturally constrained by the digital universe, the methods can be easily readapted to any computational domain of objects and properties.

Unfortunately, however, this architecture is troublesome in a crucial aspect: its applicability. If we simulate a slightly bigger world than our previous example – say, a 6x6 lattice – it turns out that *Processing* cannot properly generate the domain of the semantics due to memory constraints. Technically speaking, the problem is due to the size of the array that stores in the working memory the explicit representation of elements in the domain: once you reach a value close to the maximum allowed for integers (2147483647 items)<sup>113</sup>, the array cannot “grow” anymore: literally, there is no place for other objects in the domain of quantification. Incidentally, we may note that the computational impasse clarifies the debate about the so-called “innocence” of mereology. In particular, it becomes obvious that the Minimalist view introduced in *Chapter 1* does *not* reduce the ontological costs of classic mereology, an ontological cost that *AufByte* can *straightforwardly* and *unambiguously* quantify in the bytes

---

<sup>111</sup> For example, it is easy to add custom ontologies imposing restrictions on composition and then evaluate the same sentences to spot differences in truth values between ontologies.

<sup>112</sup> It is very easy to evaluate complex formulas thanks to the algorithm that transform each occurrence of a complex predicate into its definition with simple concepts – e.g. ‘ $a$  overlaps  $b$ ’ gets first translated to ‘There is something that is both part of  $a$  and  $b$ ’ and then evaluated. In this way, the rules for semantic evaluation are kept to a minimum (generally speaking, one rule for each undefined, primitive concept).

<sup>113</sup> Why is this particular value so crucial? Like most modern programming languages, p5 uses integers to index arrays, so that you can use commands like `myArray[128]` to get the content of a given cell from the array. However, since p5 represents a standard integer as a 32 bit signed number in the computer memory, an array with more than  $2^{32}$  / 2 objects cannot have a proper index (in practice, however, the computer will usually run out of memory well before reaching this theoretical threshold).

required by  $2^n - 1$  objects. If a Quinean-style slogan is needed, for computational philosophers ‘to be is to occupy space in memory’: the fact that sometime is epistemically convenient to treat some objects as one – as we did ourselves introducing the concept of *layer* – does not change in any way how one should count objects in metaphysical seriousness<sup>114</sup>.

One may be tempted to blame p5 for this failure, but, as it turns out, most general purposes programming languages have similar limitations for arrays and lists<sup>115</sup>. On second thought, we may be tempted to blame the whole issue of semantic evaluation: couldn’t we build an oracle without explicitly representing model-theoretic structures? In fact, given that *AufByte* is in the spirit of the Leibniz/Carnap proposal, it is natural to expect some sort of isomorphism between *semantic* evaluation of complex sentences through recursion on simpler ones, and *syntactic* deduction of complex facts starting from simple premises about the state of the universe. To test this alternative strategy we set up a program that produces, given as input the description of a *Life* universe, the list of “atomic facts” of that universe in the format of Prolog clauses<sup>116</sup>. Prolog is a programming language “inspired” by first-order logic (the name comes from the French ‘PROgrammation en LOGique’), widely used in academia and Artificial Intelligence. A typical Prolog program consists of a series of sentence, describing a situation, and some rules, defining the relations of the predicates used in the sentences. The program runs when the user asks queries: Prolog will use the facts and the rules it knows about the situation to logically *deduce* the answer to the user’s question<sup>117</sup>. As an example, here is how you may check whether Aristotle’s famous syllogism is indeed valid:

```
man(socrates).  
mortal(X) :- man(X).
```

---

<sup>114</sup> For a more articulate discussion of the Minimalist View see Berto, Carrara (2009). It is maybe time to accept that general extensional mereology is *indeed* very counterintuitive, but we do not need to make excuses – why did we suppose in the first place that the structure of the world would reflect our naïve intuitions?

<sup>115</sup> For example, Microsoft latest – and very popular – language, C#, has pretty much the same limitations.

<sup>116</sup> Both the C# “interpreter” and the Prolog program are available under request.

<sup>117</sup> Unfortunately, there is no space here to further discuss Prolog spectacular features, especially as an educational programming language (for wannabe programmers and philosophers as well). The reader may wish to read Blackburn, Bos, Striegnitz (2006) for a nice introduction and Bratko (2011) for advanced Prolog topics in Artificial Intelligence.

The first line is a fact: intuitively enough, it states that Socrates is a man. The second line is a rule ( $\text{:-}$  is the *if* symbol), stating that if something is a man, then it is mortal. If we now query Prolog with the following sentence:

```
mortal(socrates).
```

Prolog will output ‘yes’ as an answer, meaning that it was able to deduce the target sentence from the list of facts and rules. Luckily, Prolog clauses do not need to be inserted by hand every time the program runs, but they can be saved in *pl* files and loaded when needed through a specific procedure. In our case, this means that we can generate all the basic facts of *Life* universe:

```
object('1_0').  
partOf('1_0', '1_0&1_1').
```

and then we can ask Prolog to evaluate the following:

```
partOf('1_0', X).
```

Clever as it may be this second, syntactical approach, it fails for pretty much the same reason as the first one: even for small universes (30 atoms) the *pl* file contains too many facts – and Prolog cannot manage all that information for architectural constraint.

Finally, one may be tempted to dismiss the problem as ‘just a *practical* problem’ (yes, philosophers do that a lot when their ideas are plainly inadequate for the real world). However, it is crucial to understand that it’s not *just* a practical problem, something that NASA, or Google, could solve easily. The point is that the computational complexity of the task of generating a semantics for an atomic universe and unrestricted mereological composition is *exponential*: in fact, the resources needed scales roughly as  $2^n$  as the number of atoms grows<sup>118</sup>. So, while it may be true that Google and NASA could easily simulate the  $2^{32}$  universe, they still would not go very far: the domain of the semantics grows so rapidly that soon we would have more

---

<sup>118</sup> For more precise definitions and an excellent introduction to the field see Moore, Mertens (2011).

elements than atoms in the observable universe (which are about  $2^{80}$ ). Generally speaking, computational complexity considers polynomial time algorithms as “feasible” computations: things growing faster (like our domain) make for practically intractable problems<sup>119</sup>; in other words, simulating the semantics for a *Life* universe with just 1000 atoms (involving a domain with  $2^{1000}$  objects) looks like a sci-fi fantasy<sup>120</sup>.

The upshot of the discussion is that our failure to generalize *AufByte* to larger universe is not due to any conceptually obvious reason: the Leibniz/Carnap project of computing the truth value of arbitrary sentences does *not* fail for philosophical arguments, nor for problems with computing deductions in first-order theories<sup>121</sup>. The failure of the project is basically due to another source of *a priori* arguments – a source philosophers are not very familiar with: the *computational complexity* of a philosophical theory. Given that we are not aware of any research project resembling *this* work and the construction of *AufByte*, it is not surprising the issues did not come out before in the philosophical literature. While philosophers generally know very little about computational complexity, why did the A.I. and engineering community working with computational ontologies not recognized the problem? The answer is obvious once the latest works in qualitative spatial reasoning (and applied ontology/conceptual modeling) are examined: axiomatic theories of *parthood* never include the axiom of unrestricted composition<sup>122</sup>; indeed, in some cases even transitivity and extensionality are dropped<sup>123</sup>, resulting in mereological theories very different from the one developed for *Life*<sup>124</sup>.

---

<sup>119</sup> While it may look somewhat arbitrary to people outside the field, there is a general consensus among practitioners that the polynomial/exponential divide captures pretty well the feasible/unfeasible distinction. For some philosophical thoughts about the issue, see the philosophers’ friendly Aaronson (*forthcoming*).

<sup>120</sup> In the field of computational complexity, a satisfaction-related problem is the 3-SAT problem. 3-SAT has been heavily discussed in the context of the P vs. NP conjecture; in particular, 3-SAT is a so-called NP-complete problem (for an introduction to P vs. NP and NP-complete problems, see the first chapter of Moore, Mertens (2011)). However, our problem is *not* directly related to this: 3-SAT is about finding a model satisfying a given formula, but we are interested in finding the truth value of a formula given a model (the ontology of *Life*).

<sup>121</sup> As well known, meta-logical properties of first-order logic are less than ideal from a computational perspective – in particular, first-order logic is only semi-decidable (see Boolos, Burgess, Jeffrey (2002), Chapter 17).

<sup>122</sup> See for example the excellent survey in Cohn, Renz (2008).

<sup>123</sup> See for example the arguments in Guizzardi(2005), pp. 151-155.

<sup>124</sup> However, see Pontow, Dazinger, Schubert (2007) for a Prolog implementation of general extensional mereology in the context of heart anatomy. Unfortunately, their model treats a very limited number of atomic components, and the conclusion of the article mentions the problems in generalizing the approach: ‘the search for a mathematical model of structural relations that represents a compromise between computational complexity, algebraic strength and ontological validity will continue to be a subject of future investigation’, Pontow, Dazinger, Schubert (2007), p. 325.

The development of a mereological semantics under a straightforward set-theoretic model is thus a practical failure: in the forthcoming section we shall explore alternative ways to make *AufByte* “practically” useful.

### 3.3 Computational Ontology on the Cheap

Debugging is twice as hard as writing the code in the first place.  
Therefore, if you write the code as cleverly as possible,  
you are, by definition, not smart enough to debug it.  
*Brian W. Kernighan*

There is an easy way out to our complexity problem: drop the cursed axiom and replace it with something that has reasonable computational properties. Unfortunately, we have conceptual reasons not to do so. As we have seen in *Chapter 1*, David Lewis’ argument is a strong case against restrictions to the principle of composition. Second, implementing a computational ontology makes a not so obvious point about the *status* of mereology as a theory. In particular, building the interpretation function for *parthood* with unrestricted composition is a totally *a priori* task, since no actual properties of the simulated universe enter into the algorithm. While interpreting predicates like ‘isAlive’ requires some if-conditions to check whether a given cell has actually the property of *being alive*, no such if-clauses are needed to build the extension of the predicate ‘PartOf’<sup>125</sup>. In other words, once you have the atoms, the unrestricted composition principle does all the job for you, so that there is no mereological difference between the two worlds below:



No property in the universe can make any difference to the basic ontological structure of the world: it is never an empirical question whether something is part of something else, since everything is settled, once and for all, with the domain of atoms – what becomes an empirical, psychological matter is to single out a particular object in the

---

<sup>125</sup> The reader is strongly encouraged at this point to check in the *AufByte* code the relevant algorithm that unambiguously settles this issue.

mereological domain and call it ‘Jacopo’ – however, *parthood* relations with what *now* we call Jacopo are already there (so to speak) since the beginning of the universe. The metaphysical upshot is a realist version of conventionalism: reality is mind (and concept) independent; moreover, reality is, in some sense, *fully packed*: any possible composite object exists, no matter how strange or disconnected; what we do with our concepts is *select*, out of the arrays of pre-existing objects, some portions of reality that are interesting for various reasons<sup>126</sup> – a picture which is in stark contrast with “anti-realist” (cookie-cutting) conventionalism<sup>127</sup>, i.e. the idea that reality does not come with ready-made objects but it is instead the human mind that carves some “neutral stuff” into material objects with identity and persistence conditions<sup>128</sup>. According to this account, the mereological predicate ‘PartOf’ is analogous to the identity predicate ‘=’ and different from ‘isAlive’, ‘isDead’ etc.: exactly as in the case of identity, the extension of the relevant concept can be settled without looking at the arrangement of properties in the universe. It is therefore not surprising that identity can be *defined* within extensional mereology (i.e. identity is a limiting case of sharing parts), since both notions naturally belong to the very “algebra of Being”.

Let us now suppose we change the composition axiom adding some restrictions, e.g. by stipulating that only connected atoms properly compose new objects. In this case the extension of *parthood* will no longer be totally *a priori*, making the two worlds below distinguishable by some purely mereological formula (for example, ‘*a* and *d* are part of something’ is true only in the squared world):



In particular, the domain of atoms, together with a mereological axiom, is no longer sufficient to entirely settle the extension of the ‘PartOf’ predicate: the interpretation

<sup>126</sup> For a presentation and a critical discussion of this metaphysical picture, see for example Sidelle (1992), Sosa (1999) and Eklund (2007).

<sup>127</sup> For a defense of stuff ontology see Jubien (1993), Sidelle (1989). For a critical discussion of both types of conventionalism see Rossi, Tagliabue (2009).

<sup>128</sup> Stuff ontologists cannot accept mereology as a primitive theory (since reality has no objects *per se*), but they can give a straightforward operationalist interpretation to the concept of mereological summation (*any* portion of space-time can be carved and considered, by convention, as a full-fledged object).

function now requires that some “empirical” conditions on atoms have to be checked before deciding.

This revision of the original theory may be unpalatable for two reasons: first, it may be argued that mereology (like identity) should be totally *a priori*, as an ontological theory as well as a semantical one: just as the logic of identity entirely constrained the extension of the predicate in model theory, the logic of *parthood* should do the same. Second, since the nice analogy with identity (whose extension is computed *a priori*) will no longer hold, but identity, via extensionality, will still be definable in the theory, the restriction to composition may force us to drop also the axiom of extensionality for the sake of conceptual consistency (an axiom we may wish to retain for independent reasons).

If we are not willing to give up the principle of unrestricted composition, what are the remaining options? If we drop the requirement of a strict isomorphism between *AufByte* computational semantics and standard, model-theoretic constructions, a way out is provided by a more “dynamical” approach to semantic evaluation. The intuition here is that the evaluation of formulas does not require the entire semantics to be present *at once* in the memory of the computer; instead, we only need to represent explicitly the object that is currently being evaluated as having / not having the property expressed by the predicate (given a compositional semantics, any evaluation boils down to evaluate atomic formulas). Given  $n$  atoms, we use a  $n$ -bit number to represent all the possible objects in the domain, such that the mereological sum of objects  $i$  and  $j$  is represented by a string of  $n$  bits where  $i$  and  $j$  are 1 and all others are 0<sup>129</sup>. So, when evaluating an arbitrary formula in which the assignment of  $x$  is the  $m^{\text{th}}$  object in the domain, we get the atoms of the  $m^{\text{th}}$  object by converting  $m$  to a binary number of  $n$  digits and then counting the 1s in the string: given the extensional, atomistic nature of the underlying theory, the atoms composing a given object are all that matters to evaluate the satisfaction of base formulas<sup>130</sup>.

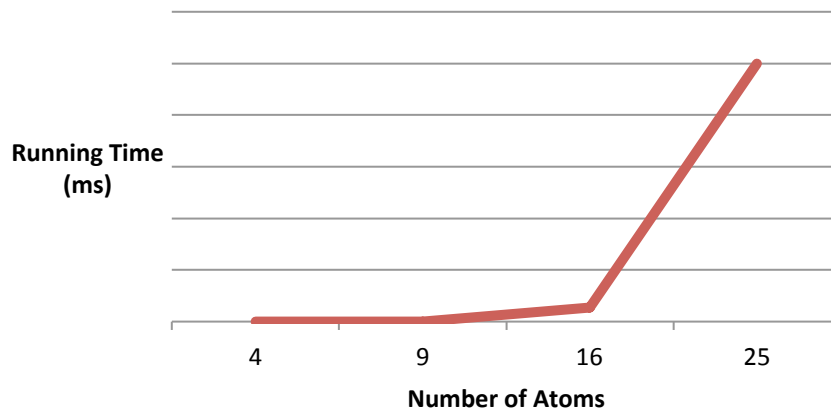
The new approach has the advantage of managing the working memory of the computer much more efficiently; however, this advantage comes at the cost of departing from traditional set-theoretic constructions for semantics, which are wonderfully elegant

---

<sup>129</sup> A  $n$ -bit number can represent  $2^n$  objects. Given that the string with all 0s is not used (since there is no “null individual” in the theory), the total number is  $2^n - 1$  objects, as required by mereology (e.g. the universe is the string with all 1s).

<sup>130</sup> Once we have the atoms, it is of course trivial to get their properties by checking the simulation of the universe.

but, as we have seen, have not been devised with computational efficiency in mind. In practice, *AufByte* code becomes a little less transparent, since the *Structure* class now allows for *two* different mechanisms of satisfaction: a classic, model-theoretic approach, and a new, dynamic approach<sup>131</sup>. It is also very important to understand that the new semantics does not *solve* the problem, it merely avoids the crash; in particular, the computational cost of ontology is exactly the same, but it is now paid in a more abundant resource, time, instead of memory. The awful scaling properties of general extensional mereology are still there, even if the failure is not so spectacular. While now simple statements (*‘a is part of something’*) can be easily computed also for big universes, the following graph illustrates what happens to the algorithm running time with more complex cases (like the nested quantifiers in the definition of the universe, i.e. *‘there is an  $x$  such that everything is part of  $x$ ’*)<sup>132</sup>:



Finally, it is important to stress that the point we made above on the *a priori* status of mereology is still valid in the new semantics: in particular, it is only because of the axiom of unrestricted composition that we know “in advance” all the objects in the domain and, therefore, we can use the *n*-digit number to navigate through the domain.

<sup>131</sup> The sophisticated reader may here appreciate the benefits of OOP: changing *Structure* required only minor changes to *Formula*: in particular, all the code for the Tarski-style recursive evaluation remained untouched.

<sup>132</sup> As expected, the statistical coefficient for the correlation between running times and number of objects in the ontology is practically 1.0 (0.99996).

### 3.4 The Reality of Time

Time is an illusion. Lunchtime doubly so.

*Douglas Adams*

As we have seen in *Chapter 2*, introducing *Life* temporal dimension in the ontology requires a non-trivial extension of the basic mereotopological machinery we set up for *Life* snapshots. The first dilemma is *endurance* vs. *perdurance*: do cells exist at different times or not? As anticipated in the previous chapter, the computational side of the issue is pretty unproblematic. Given that we wish to represent change, our semantics should be able to talk about past *and* present objects, but the goal can be achieved either by having each cell stores its previous states (endurance), or by having a three-dimensional array of cells, each existing exactly in one instant (perdurance) – for the philosophical reasons previously given, *AufByte* adopts a three-dimensional account of change.

A more interesting issue is the ontological status of past events. In a basic *Life* simulation, only *present* things can be referenced in the semantics, since there is no memory of past states and they cannot be recovered starting from the present state. If we want to quantify over past events, there must be *something* we can quantify over – but what? The philosophical options on the market are basically two, *presentism* vs. *eternalism*<sup>133</sup>: on a presentist account of time, only present things *exist*, while eternalists are happy to countenance past (and future) entities in their catalogue of the world. In many regards, the eternalist view is the natural outcome of taking the analogies between time and space seriously: just as people in New York exist even if they are far from you in space, dinosaurs exist even if they are far from you in time – in both cases, the logical structure of the existential quantification (‘There are New Yorkers’, ‘There were dinosaurs’) is pretty much the same. As in many debates where a “realist” side faces a “fictionalist” side, the fictionalists want to assert the same sentences the realists do without committing themselves to the same ontological costs. In this case, the presentist may wish to say that ‘There were dinosaurs’ means something like ‘In the past, there were dinosaurs’<sup>134</sup>. The strategy solves the ontological problem, just like asserting ‘According to a story (by J.K. Rawlings), there are wizards’ does not imply any

---

<sup>133</sup> For an overview of the debate, see Markosian (2010).

<sup>134</sup> Technically speaking, this is equivalent to say that the existential quantifier is in the scope of an intensional operator, so the ontological commitment is avoided.

unwanted commitment to wizards. What is the computational counterpart of this debate? An eternalist philosopher will say that when we quantify over past states of *Life* evolution, we are actually quantifying over existing cells and properties that are stored in the working memory just as the present state is; a presentist philosopher cannot use *past* cells stored in the working memory, so a *representation* of past cells should be used for the semantics (just as we use stories to represent Harry Potter's properties). However, how can we represent *within Life* past states of the world without being committed to their existence (remember the credo: 'to be is to take up some space in the memory')?

For sure, at  $t$  we cannot have *Atom* objects instantiated except those that are *present*: so, to make sense of sentences about the world at  $t-1$ , we should use something *in* the cells to represent the tensed predicates 'is alive', 'was dead' and so on. The upshot of this strategy is that a presentist version of *Life* will occupy pretty much the same space in memory as an eternalist version: if we judge ontological commitment by bytes, it is clear that presentism is *not* cheaper than eternalism. But even if we do not judge ontological commitment in this "rough" way, the point surely deserves some attention: it is conceptually impossible to have a presentist semantics without tampering with the basic properties and objects of *Life*, resulting in a strange theory that can no longer claim to be more commonsensical than the alternative.

Of course we could represent the past *outside Life* by saving a representation of past world states: the universe itself remains the old same *Life* we know and love, but we collect a growing "story" about the world as the universe evolves and we use it to give us the "fake" truth-makers for sentence about the past (pretty much as stories by J.K. Rawlings give us fake truth-makers for sentences about wizards). Unfortunately, using this strategy will commit us to a different *category* of objects in the ontology of *Life*, i.e. *representations* of *Life* past states: we started with just our space-time with cells and states, but now we must recognize the existence of something *outside* space-time, i.e. stories about past states. We can appreciate here with clarity and precision a familiar move in many philosophical debates: we start with a world of "concrete" objects and properties and then we introduce "abstract" objects and properties to account for certain behavior of "concrete" objects. In its most general features, the issue is that modeling entities to avoid unwanted commitments requires either some form of representation

within the world, or some form of representation outside the world – so, either you change the ontology or you introduce something *outside* the world: the concrete/abstract ontological divide is thus understood, in its full generality, as a debate between in-world/out-world objects<sup>135</sup>. If this is true, presentism requires one of these unpalatable consequences, a deep metaphysical revision of *Life* structure or a commitment to “abstract” objects<sup>136</sup>. Moreover, we can drive another point home by noting that in a digital universe *both* “concrete” and “abstract” objects are, fundamentally, *virtual* objects (i.e. just bytes): the commitment (measured in the quantity of information needed by the ontology) of the non-realist account is much closer to the realist account than what is usually acknowledged in the general debate (presumably because, say, sets and propositions are imagined as “thin and rarified” entities, while dinosaurs are “heavy and full-bloated” objects)<sup>137</sup>.

For all these reasons, we shall adopt an eternalist conception of time and let ourselves directly store in memory an arbitrary number of past states. Before moving on, it may be interesting to answer one last question about time and CA: are there conditions that would allow for a presentist world without changes in ontology (either direct changes or the introduction of *abstracta*)? Suppose that all we can store about a universe is the present state of the lattice, with no “hidden variables” inside cells recording past events; suppose further that we know the CA updating rule: why cannot we just *compute* the truth value of sentences about the past? The answer in *Life* is pretty obvious: given the present lattice, there is no way to know with certainty how the lattice was at the previous instant. However, there is a class of CA that violates this constraint: if the updating rule is *reversible*<sup>138</sup>, information is always conserved during an update, so it is possible at any time to recover the previous state of the lattice. Ontologically speaking, these universes are incredibly cheap: *talking* about those worlds is very complicated, but their ontology is transparent and light (no strange properties around, no *abstracta*, no

---

<sup>135</sup> In the particular case of presentism, a further point needs some clarifications: if only *present* entities exist, what is the status of the abstract representations of the past? Does it make sense, from a presentist perspective, to say that stories about the past exist *simpliciter*?

<sup>136</sup> Of course, someone may argue that abstract objects are not troublesome entities at all. The point, however, is slightly different, since it is a metaphysical theory about time (not mathematics) that *forces* us to include *abstracta* in our ontology. While this does not seem strange in our world (since we already have independent uses for sets and numbers), in *Life* this looks like an unnecessary burden.

<sup>137</sup> This may not seem a crucial point - until you seriously entertain the following hypothesis: what if *our* world is a digital universe? We shall address this question directly in *Appendix IV*.

<sup>138</sup> See for example the CA presented in Berto, Rossi, Tagliabue (2010), *Chapter I*, and the discussion of reversibility therein. For a classic reference, see Fredkin, Toffoli (1982).

past states in memory): we get an improvement in what Quine calls ontology, paid for in the coin of ideology<sup>139</sup>.

### 3.5 Layers in Computational Semantics

The limits of my language  
mean the limit of my world.  
*Ludwig Wittgenstein*

We introduced *layers* in *Chapter 2* as semantical devices to single out (epistemically relevant) parts of the universe and treat them as non-decomposable objects. Semantically speaking, *layers* are just restrictions of the domain of quantifications: when we see a glider floating alone in the universe, the sentence ‘there is one object alive in the universe’ strikes us as a *true* sentence, even if mereology implies that, whenever there is *one* glider, there are *many* objects alive in the lattice. To examine a more familiar situation, when your best friend opens the fridge and says ‘there is no beer’, he is not saying that Reality does *not* contain beer (after all, this may not be the best of all possible worlds, but it is certainly not the worst): he is saying that *in the fridge*, there is no beer. In our non-philosophical moments, we are so good at picking up contextual clues that restricting quantification to limited portions of the world is an automatic, effortless process: can we teach our program to do the same semantic magic? There are two main strategies that may be adopted: on the one hand, one may insist that this kind of pragmatic sensibility is only one aspect of a more general ability – let us call it “common sense reasoning” – so that we cannot tackle the problem without considering the huge, unsolved, problem of formalizing common sense, the true nemesis of any A.I. researcher; on the other, one may adopt a more modest attitude and just ask the following question: is there anything *within* the formal structure we have already built that can help us during the evaluation of contextually restricted formulas? While we made abundantly clear with *this* project that we love ambitious and never-ending challenges, in this case we settle for the modest alternative and try to sketch some preliminary ideas.

---

<sup>139</sup> See Lewis (1986b), p. 4, for the original, opposite slogan.

Our starting point for some intuitions is the Relevance Theory (RT), proposed in Sperber and Wilson (1986) as a cognitively plausible development of the theory of conversation put forward in Grice (1989)<sup>140</sup>. RT holds that human communication tends to maximize *relevance*: i.e. ‘the greater the positive cognitive effects with the smaller mental effort to get them, the greater the relevance of the input for the individual’<sup>141</sup>. In particular, every utterance communicates a *presumption of its own optimal relevance*, where “optimal relevance” is spelled out as *i)* the utterance is relevant enough to be worth processing; *ii)* it is the most relevant one compatible with the communicator’s abilities and preferences. In the case of ‘there is no beer’, the utterance is indeed optimally relevant, since it is worth processing and it is the most relevant, given my preferences and psychological attitudes<sup>142</sup>; for the same reasons, in the case of the glider floating alone in the universe, the utterance ‘there is only one object alive in the universe’ is indeed optimally relevant. In the context of computational semantics, a first translation of the intuition behind RT may be the following: if a sentence is *false* taken at its face value, perhaps there is a more relevant interpretation of the quantification that makes the sentence *true*; if such an interpretation can be found, it is to be preferred to the literal reading (*ceteris paribus*). Let us try to see step by step how the algorithmic procedure may work for ‘there is only one object alive in the universe’:

- 0) The formula ‘There is one object alive’ is true just in case:
  - a. there is one atomic object with the property of *being alive* or there is a non-atomic object such that each of its atomic parts has the property of *being alive*; call the object  $x$ ;
  - b. for any object  $y$  that is alive,  $x = y$ .
- 1) ‘There is one object alive’ is false because (0.b) does not hold in the unrestricted domain. Check if there is a contextual restriction on the domain of quantification.
- 2) Let  $L$  be an *admissible chain of layers*, in the sense of *Chapter 2*, composed by three layers: the atomic layer, the glider layer (i.e. the layer with two objects, the glider and its dead complement) and the Eleatic layer.

---

<sup>140</sup> For an overview in the context of contemporary pragmatics, see Korta and Perry (2012).

<sup>141</sup> Korta and Perry (2012), Section 3.2.1.

<sup>142</sup> For example, ‘there is no yellow, slightly alcoholic, German liquid’ would be an utterance that is worth processing, but less relevant given my overall cognitive abilities and mental states.

- 3) ‘There is one object alive’ is false in the atomic layer because (0.b) does not hold in that layer. Go to the next layer.
- 4) ‘There is one object alive’ is true in the glider layer.
- 5) Since the glider layer is “more relevant” than the atomic layer, there is a contextually relevant restriction: so, metaphysically speaking, the sentence is false (as it should be); pragmatically speaking, the sentence is true (as it should be), given the existence of a suitable layer.

Steps 0-5 are relatively straightforward computational steps. However, we did not clarify the most important part of the derivation: why should the *glider layer* be “more relevant” than the *atomic layer*? Since the only kind of pragmatic trick we are considering involves the size of the domain, the answer is that the *glider layer* is smaller than the *atomic layer* – and since it is smaller, it takes less time to be processed (as required by RT original intuition). It should be obvious that these sketchy suggestions are still very far from a full-fledged theory of quantification in a communicative setting: first, we need to understand the process of selecting an admissible chain of layers that is *adequate* for the sentence we are evaluating (given the combinatorial explosion of our mereology, it could take years to find a layer that makes the formula true<sup>143</sup>). Moreover, the theory as stated has the following unwelcome consequence: any layer composed by two objects, one cell that is alive and its not-alive complement, would make ‘There is one object alive’ true<sup>144</sup>.

Notwithstanding all the work that still needs to be done, digital universes are a natural benchmark to test and improve theories of language and cognition: in this context, the notion of *layer* may be a useful bridge between the mereological nature of the mind-independent world and the concepts epistemically relevant for the human mind.

---

<sup>143</sup> A natural tactics would be to try first with layers containing instantiations of known concepts: if you have the mereotopological definition of a *glider* as a process with such-and-such properties, you could look for layers containing exactly these properties.

<sup>144</sup> A natural suggestion (see also the previous footnote) would be to impose the following metrics on layers with the same size: the greater is the number of objects recognized by the layer that are instantiations of known concepts (e.g. gliders), the greater is the relevance of that layer.

### 3.6 Lessons from AufByte

In theory, there is no difference between theory and practice.

In practice, there is.

*Yogi Berra*

We started the chapter by linking *AufByte* to some historical precedents, Leibniz's *Mathesis Universalis* and Carnap's Aufbau. A formal theory encompassing everything, in which complex concepts are explicitly built from simple ones and every truth about the world can be deduced within the theory, is an amazing image for a philosopher's mind. Unfortunately, the history of the two precedents shows that the goal is not easily attainable: some truths just cannot be deduced and the whole idea of "primitive concepts" is much messier than what we thought.

*AufByte* tackles the same basic challenge, taking advantages of the peculiar features of digital universes: they are finite, they are ontologically neat and, moreover, they can be easily implemented in a computer, so that, in the end, philosophers would just sit down in front of Processing and say 'Calcuemus!'. Our (not so) little computational experiment with a foundational ontology made the original Leibniz/Carnap project come to life, even if in a very limited domain; nonetheless, it shows that the ambitious project of formalizing the logical structure of (portions of) reality is even more appealing today, when we know much better the relevant theoretical constraints and when technology provides us with a rewarding, exciting feedback on our theorizing.

Among the several issues raised in this chapter, three topics strike us as particularly important:

*The interplay between mereology and realism:* computational ontology helped us frame the problem of ontological commitment in a non-standard, yet very precise and effective way. In particular, when it comes to mereology we have seen there is no clear notion of "innocence" to be invoked (considerations about *layers* excluded); moreover, we have seen that the version of realism supported by our ontology is a "super-realism", a world *fully-packed* with objects. Given the chosen mereology, this does not look very surprising; however, it is important to stress that accepting this ontology means *to reject* others; in particular, our super-realism is not compatible with "conventionalism" (or any

kind of stuff-ontology), nor is it compatible with the commonsensical distinction between *fiat* and real objects. To better appreciate the point, let us use the terminology of Varzi (2005) and distinguish between *de re* and *de dicto* boundaries: *de re* boundaries ‘carve nature at its joints’, as Plato’s butcher would say; *de dicto* boundaries mark entities whose existence is conventional. So, my dog has *de re* boundaries, since it is a *bona fide* object, a living thing, while France has *de dicto* boundaries, since the extension of political entities are completely arbitrary. Generalizing a bit, we may say that *de re* objects have persistent and identity conditions completely independent from our interests and psychological attitude, while *de dicto* objects persist and exist as *fiat* entities. While it is a matter of *stipulation* whether the Lakers, moving from Minnesota to California, are one and the same team persisting through time, or two teams with the same name, it is a *metaphysical* problem whether our friend John, who lost his memories in a car accident, is still the same person as before, or someone new “born” in that body after the crash – it may be hard to answer the question, but surely it is not up to *us* whether John survived or not. With this distinction at hand, we can classify realist theories on a continuum: on one extreme, conventionalism, maintaining that, after all, there are just *de dicto* boundaries; what looks like a *real* object is actually no different from a basketball team or a nation (in a slogan, minds *create* objects). On the other extreme, super-realism, maintaining that, after all, there are just *de re* objects; what looks like an arbitrary stipulation is no less real and mind-independent than John and my dog<sup>145</sup> (in a slogan, minds *select* already existing objects). In the middle there is the commonsensical, moderate position we started with: John is a *bona fide* object, France is a *fiat* entity. Different accounts require different mereological approaches: while super-realism and common-sense agree on mereology as a general “algebra of Being”, but disagree on the principle of composition, conventionalism cannot even accept mereology as a purely *formal* theory. Theories require *objects* and conventionalism doesn’t have them: there is no domain of quantification available *before* some boundary is drawn and some convention is settled. But if mereology can just quantify over conventional objects it does not seem to be the general, *a priori*, all-encompassing

---

<sup>145</sup> For example, let’s agree that the identity of the Lakers supervenes on spatio-temporal material objects. According to this picture we have two disconnected, gerrymandered composite objects, overlapping for the first of part of Lakers history; however, only the second object has the Californian part. The survival of the team does *not* amount to an ontological stipulation, but to a semantical (arbitrary) decision: should ‘Lakers’ be used as a name of the first or the second of these two objects?

theory we thought it was<sup>146</sup>: there is an appropriate mereology for any “conceptual scheme” and there is no way to say that one version is right and the others are wrong<sup>147</sup>. Of course, these considerations could have been made even *without* the help of the *AufByte*: however, it is worth stressing that the ontological import of a theory shows itself more dramatically in the context of a computational framework. It is also important to note that what emerged with *AufByte* is a somewhat non-standard view on *parsimony*: it is a recurring theme in ontology that to multiply the *number* of entities is not troublesome, since Ockham’s razor only bans the multiplication of *kinds* of entities – so you can go from 1 to 1000 *As* as long as you do not need to introduce *Bs* in the theory. This is certainly a sensible claim, one that intuitively captures the idea that the primitive notions should be kept to a minimum. According to a computational perspective, things look different: if a new theory implies a change in the complexity class (from  $n$  to  $2^n$  objects in the semantics), it does not look so innocent after all, and it is certainly *more* complex. Philosophers arguing against classical mereology have been missing the point for decades: it is not about counting or about numbers, it is about scaling properties.

*Fundamental vs Not-So-Fundamental ontological structures*: *AufByte* required a robust understanding of the computational aspects of the semantics of a mereological theory. In particular, *AufByte* highlights the crucial importance of the composition axiom for the computational semantics: even if the term “mereotopology” suggests otherwise, mereology is more “fundamental” than topology. As we have discussed at length in this chapter, the proposed mereology constrains *a priori* the domain of quantification, so it must effectively be in place before doing any evaluation whatsoever. In other words, the computational perspective highlighted an interesting fact which is not that evident from the axiomatic perspective.

*The cost of computing truths*: the digital version of the Leibniz/Carnap project does not suffer from the problems that plagued the original attempts: we are doing better than Leibniz, since, in principle, every truth is computable in a finite time; we are also doing

---

<sup>146</sup> A similar point is made in Berto, Tagliabue (*forthcoming*) and it is briefly discussed in *Appendix II*.

<sup>147</sup> It is not clear if some philosophers defend conventionalism *and* general extensional mereology. However, Varzi (2005b) and Casati, Varzi (1999) suggest that Varzi is shifting back and forth between the two perspectives.

better than Carnap, since, in principle, our primitive notions (basically, *Life* states and the proto-geometry) are not obviously flawed by positivistic prejudices and the like. As we have remarked many times, a striking truth emerging from *AufByte* is that our digital Aufbau suffers from computational complexity due to the principle of unrestricted composition. We just want to add here two final considerations on this topic. First, any practical application using the principle is severely constrained in generality<sup>148</sup> (or, in any case, requires a carefully designed architecture to limit the downsides of combinatorial explosion); from the opposite perspective, if it turns out that a successful “knowledge engine” can be built on a weaker mereology, how could we still claim that the cursed axiom represent reality<sup>149</sup>? Second, it has been suggested that computational tools may be applied to ethical and legal reasoning<sup>150</sup>: if computing ethical principles and legal rules cannot be, in some way, decoupled from the mereological structure of the universe, the project of a computationally administered law becomes hopeless. When discussing about ethics and combinatorial scenarios, another famous image from Leibniz metaphysics comes to mind, i.e. God observing all possible worlds and then choosing the one with the *best* balance of good and bad. In the spirit of negative theology (the view that God’s nature can be somewhat captured by saying what God is *not*) *AufByte* makes us sure that the God of any digital reality is *not* using a 32-bit computer.

### 3.7 Questions and Answers

Q<sub>1</sub>) What kind of answers the *AufByte* oracle can compute?

A<sub>1</sub>) In theory, all the ontology we developed can be easily incorporated into the *AufByte*; in practice, at the moment of writing the end of *this* work (February 2013) not every concept is computationally available. However, *AufByte* is already more than just a proof of concept: for example, almost any mereotopological query can be evaluated (with much less than one hundred lines of code). Due to the recursive nature of the computational semantics, there is no limit (except time and computer memory) to the

---

<sup>148</sup> A problem we ourselves overlooked in a previous work: see the semantics presented in Berto, Rossi, Tagliabue (2010), pp. 93-95.

<sup>149</sup> In theory, an easy way out is to claim that true mereology accepts the principle of unrestricted composition, but that “regional” sub-domains may admit some contextual restrictions. However, at that point the challenge becomes the development of a new formal framework where “true” and “fake” mereological principles can coexist without problems.

<sup>150</sup> See for example Rossi, Rossi, Sommaggio (2010).

complexity of the queries: just like humans understand ‘the father of the father of John’ or ‘the population of the capital of France’, *AufByte* has no problem in understanding nested referential expressions like ‘the closure of the neighborhood of  $x$ ’. Finally, it is worth noting that no optimization has been done to computationally exploit particular facts in the semantics of digital universes: for example, if ‘Neighborhood of  $x$ ’ is contained twice in a sentence, the semantics will compute the referent of the expression twice instead of saving the value after the first time. At this point, since we are more interested in the conceptual import of the experiment, we chose to make the code as general as possible without bothering too much about low-level performance optimization.

Q<sub>2</sub>) Does *AufByte* understand the theory?

A<sub>2</sub>) We let ourselves speak figuratively of *AufByte* as the computational counterpart of an oracle; moreover, many times we highlighted the strong connection between formal ontology and knowledge representation for Artificial Intelligence. Is therefore natural to ask: by teaching ontology to our laptop, are we making it (a little bit) “intelligent”?

There are many ways to argue for/against the attribution of intelligence to artificial systems – and we can’t survey all the possibilities here. However, acknowledging that the following will hardly settle the matter, we list three different perspectives suggesting (not proving!) that *AufByte* has something like a true understanding of this very limited universe. The first perspective cannot be other than Turing’s, who famously changed the ‘can this machine think?’ question into the ‘can this machine pass my conversational test?’ question<sup>151</sup>; in our setting, the proposal amounts to ask: can *AufByte* be as accurate as a human in evaluating queries about *Life*? The answer is *yes*. The second perspective is the “inner model” idea, as presented for example in Berto, Rossi, Tagliabue (2010): intelligent artificial systems should possess an internal model of the external world they perceive/live in<sup>152</sup>. The condition seems to be satisfied, since the very purpose of teaching our ontology to the computer is to formally model the world of *Life*. Finally, the third perspective comes from the field of lexical competence, which studied extensively the problem of attributing knowledge of a concept to an intentional system.

---

<sup>151</sup> Turing (1950) is a landmark in the history of A.I. and philosophy of cognitive sciences.

<sup>152</sup> See Berto, Rossi, Tagliabue (2010), Chapter I.

Following Marconi (1996) and Marconi (1997), we may say (as a very rough approximation) that a system  $S$  understands the word  $W$  in language  $L$  if and only if:

- i)  $S$  knows the semantical (i.e. truth-conditional) contribution of  $W$  to any complex expression of  $L$  containing  $W$ .
- ii)  $S$  knows the inferential patterns licensed by  $W$  in any complex expression of  $L$  containing  $W$ .<sup>153</sup>

For example, I understand the meaning of ‘pera’ [‘pear’] in Italian since I know what is the reference of ‘pera’ in ‘la pera sta sul tavolo’ [‘the pear is on the table’] (so, I know how the word contributes to the truth value of the sentence) and I also know that from ‘la pera sta sul tavolo’ I can infer that ‘un frutto sta sul tavolo’ [‘A fruit is on the table’]. In the *AufByte* case, the relevant conditions seem satisfied for the predicates of the theory<sup>154</sup>, so that we could say, for example, that *AufByte* understand the meaning of ‘part of’.

None of these three arguments can be considered a final proof of “intelligence”: however, taken together they make a *prima facie* good case that the ontology built inside the code provides some (very primitive and limited) form of artificial understanding.

Q<sub>3</sub>) Can *AufByte* be applied *outside* digital universes?

A<sub>3</sub>) *AufByte* is a philosophical experiment and was designed and implemented as such. However, it is easy to see that even this small Processing project can be easily extended to solve more ambitious problems: on the one hand, the computational semantics can be readapted to other ontologies built in the same *Aufbau*, “bottom-up” approach; on the other, the user interaction system could be changed to accommodate natural language queries. In particular, once a natural language grammar for digital universe has been prepared, a module mapping natural language to formal language sentences will do the trick: users ask ‘ $b$  is part of the neighborhood of  $a$ ’, the query get translated in the first-order equivalent and then evaluated as usual by the *AufByte* oracle. It is also easy to see

---

<sup>153</sup> The conditions as stated are not Marconi’s, but they seem to capture pretty well the core of his proposal in this context.

<sup>154</sup> While no inferential algorithm is already implemented in *AufByte*, it will be trivial to expand upon the existing code and add the new feature (see also the last question of the section).

how the semantics could be developed to answer *wh*-questions instead of simple yes/no queries: at a first approximation, you can translate ‘what is the neighbor of *a* that is also part of *c*’ with something like ‘return the object(s) *x* in the domain satisfying the condition  $Nxa \ \& \ Pxc$ ’.

Whatever the pros and cons of the proposed approach, it is worth noting the currently growing effort of the IT community in the manufacturing of automated Q&A systems: IBM Watson, Apple Siri, Wolfram Alpha, Venexia, are just some among the most prominent attempts in applied Artificial Intelligence to bridge the gap between humans and machines in the understanding of natural language and management of procedural and encyclopedic knowledge. This research paradigm stands or falls depending on two main variables: the precision and consistency of the formal representation of reality and the effectiveness of the computational semantics behind the scenes. Call it “computational ontology” or any other way, it looks like something philosophy may fruitfully contribute to.

## 4. Conclusion

### 4.0 Looking Back, Going Forward

It always takes longer than you expect,  
even when you take into account Hofstadter's Law.  
*Hofstadter's Law*

At the end of *this* small investigation, a recap of the main themes may be useful. The following (likely containing some overlapping, repetitions and omissions) is a list of the core issues raised throughout the dissertation:

*CA as experiments in silico: Life* allows for a smooth manipulation of all the interesting features of reality while exhibiting an astonishingly variety of emergent behavior. Moreover, *Life* is a never-ending source of new intuitions: as Dan Dennett put it, *Life* is ‘a prodigious versatile generator of philosophically important examples and thought experiments of admirable clarity and vividness’<sup>155</sup>, such that ‘every philosophy student should be held responsible for an intimate acquaintance’ with it<sup>156</sup>. If you believe that reduction of higher-level to lower-level entities entails trivial behavior, if you believe that strictly determinism produces predictable results, if you believe that space-time regions and material objects are obviously different things – well, you should give *Life* a chance of challenging your convictions.

*Digital worlds as formal frameworks*: the logical structure of digital universes is indeed very rich, allowing for a neat and interesting formal modeling. The basic structures involved are clear and undisputed, and many standard arguments can be rephrased to be evaluated in the context of CA. The interplay between the naive world, our ontological theory and our favorite semantics can be studied with precision in the framework. Finally, given that digital universes are computer friendly, we have as a

---

<sup>155</sup> Dennett (1991), p. 37.

<sup>156</sup> Dennett (1991). p. 37.

bonus the chance of implementing the formal theory, getting additional insights and learning non-trivial things (see below).

*Computing ontology as “sanity check”*: by computing our ontology we learned, basically, one thing over and over again: *details matter*. It was only through the application of formal concepts to the “practical” task of building an oracle for *Life* that we encountered many challenges that were not visible when sitting in the arm-chair of philosophy. Modeling philosophical notions in a computational context helped us see problems from a different perspective and gave fresh insights on many topics (Minimalism and the innocence of mereology, the notion of computational parsimony, the ontological import of a theory etc.). The computational philosopher’s credo is therefore the following: if a theory cannot be modeled by some computational structure, it is likely too vague to be interesting or fully grasped.

Looking forward, it is easy to see that there is still much work to be done within the basic framework; again with some overlapping, repetitions and omissions, the following is a list of topics that surely deserve a better look:

*Advanced computational semantics*: the latest *AufByte* “stable release” at the time of writing this chapter (February 2013) does not contain a temporal dimension. Moreover, implementing *layers* and *modality* is a natural (but challenging) extension to the framework. The good news is that the “production cycle” of the program does not need to follow the consuming protocol of academic publishing: as soon as new features are developed and bugs are fixed, the source code can just be released online.

*Quantifying emergence*: weak emergence is a very important concept in the scientific practice and CA are a natural environment for testing our intuitions. As we have seen, Dennett suggested that emergent features are the ones that can be exploited for prediction. However it is not clear how to *detect* those features: do we have algorithmic procedures to discover such regularities? Computational mechanics – as introduced in Crutchfield (1994) – seems a promising approach, but more work is needed also on the

philosophical side (e.g. what is the link between emergent features and the laws discovered by special sciences?).

*The metaphysics of properties:* most of the time, we talked about *objects*, not properties. We suggested that states in a CA can be thought as perfectly natural properties in Lewis' sense, but we did not go much farther than that. Can *Life* be used to give an alternative (or complementary) account of naturalness? Can we exploit the computational nature of CA to *measure* how much natural a property is?

*Comparison with Chalmers' latest work:* David Chalmers' latest book, *Constructing the world*, is an explicitly attempt to revive Carnap's project as applied to the real world, not a toy universe. Both *this* work and Chalmers' are, in some sense, different ways of making some philosophical progress by "updating" the Leibniz/Carnap dream: a comparison between the two approaches could indeed be very interesting<sup>157</sup>.

I should have liked to write an essay in piecemeal, unsystematic metaphysics, offering a new, independent proposal on some specific topic. It was not to be. All in all, *this* work is more general than most: it is not about a single argument or a particular problem; instead, I tried to convey the overall picture of digital philosophy as it may apply across several parts of philosophy. It is no wonder that there is much work to do to fill in the details, but without a general recognition of the territory, it would not even be clear what details would still be needed.

#### *4.1 AufByte and the Future of Philosophy*

Progress isn't made by early risers.  
It's made by lazy men trying  
to find easier ways to do something.  
*Robert Heinlein*

In *Chapter 3* we emphasized how abstract works in formal ontology, when coupled with a computational semantics, can indeed be a fertile field of interdisciplinary collaboration

---

<sup>157</sup> For example, digital universes looks like a finite, computational version of Chalmers' Cosmoscope (see Chalmers (2012), pp. 114-118).

with computer science and Artificial Intelligence. As concluding remarks, we just lay out here two “sociological” thoughts on the practice of philosophy inspired by the *AufByte* project.

With the explosion of neuroscience as a mainstream enterprise<sup>158</sup> (and the resulting marketing hype in the general and scientific press), it is becoming extremely hard to find philosophers that do *not* know what the *nucleus accumbens* is; on the other hand, it is still almost impossible to find a philosopher that knows what a *compiler* is. Of course, it is good news that practitioners are developing some kind of scientific sensibility; moreover, it is *very* good news that theorizing is informed by the latest empirical findings. However, one question comes naturally to mind to the old-fashioned philosopher: if we do not fully understand how gliders emerge in *Life*, what level of *precision* and *clarity* can be expected from claims relating psychological attitudes to neuronal activity?

The second note is directly related to the first. If we want to understand brains, isn’t computer science a very natural fit for the philosopher’s toolkit? Digital universes are abstract, precise models supported by a mature and formal theory: if we want to develop a theory of representations, for example, *Life* looks like a natural benchmark. Moreover, it is arguably the computational nature of the brain that makes it interesting: biological *explanations* can only go that far without invoking computational concepts. A side-effect of mixing (theoretical and practical) computer science with philosophy is to make philosophy *relevant* once again. At the beginning of the last century, prominent philosophers were standing at the very frontier of human knowledge, in contact with (and respected by) the most important scientists of the time. If – as many argue<sup>159</sup> – the next big thing is the advent of artificial minds, philosophers will have all the skills needed to regain a central place in the quest for solutions to the most pressing and exciting challenges for humanity.

As a final note, I would like to quote what a wise philosopher wrote at the beginning of his Ph.D. dissertation:

---

<sup>158</sup> A picture is worth a thousand words: a quick look at Google Ngram is much more revealing than any bibliographic reference.

<sup>159</sup> See for example Kurzweil (2005).

It is the profession of philosophers to question platitudes that others accept without thinking twice. A dangerous profession, since philosophers are more easily discredited than platitudes, but a useful one. For when a good philosopher challenges a platitude, it usually turns out that the platitude was essentially right; but the philosopher has noticed trouble that one who did not think twice could not have met. In the end the challenge is answered and the platitude survives, more often than not. But the philosopher has done the adherents of the platitude a service: he has made them think twice.<sup>160</sup>

In the end, what may be valuable in the unorthodox methodology of *this* work is a meta-challenge for standard philosophers. It is the profession of computational philosophers to question platitudes philosophers accept without thinking twice, as we noticed many troubles that one who did not try to compute his ontology could not have likely met. In the end most philosophical theories we discussed survived intact. But we hopefully made the standard philosophers a service: we have made them think twice.

---

<sup>160</sup> Lewis (1969), p. 1.

## Appendix I: The Formal Theory

### I.1 Introduction

Arrest this man  
He talks in math  
*Radiohead*

For the geeky reader, we now restate all our axioms using the precise language of logic. Our basic language will be classical first-order logic with identity and the usual set of Boolean connectives ( $\neg$ ,  $\&$ ,  $\vee$ ,  $\rightarrow$ ,  $\leftrightarrow$ ), whose syntax and semantics can be found in any textbook<sup>161</sup>. The order of the axioms and definitions follows (almost) exactly the discussion in the main text: new notions are introduced when needed and some alternative axiomatizations are discussed.

### I.2 Axioms from Chapter 1

We start by listing the lexical axioms of mereology:

PL.1) Everything is part of itself.

$Pxx$

PL.2) Two distinct things cannot be part of each other.

$Pxy \& Pyx \rightarrow x = y$

PL.3) Any part of any part of a thing is itself part of that thing.

$Pxy \& Pyz \rightarrow Pxz$

For convenience, we introduce some new predicates:

$O_{\text{def}}$ ) Two objects *overlap* iff there is an object that is part of both.

$Oxy =_{\text{def}} \exists z (Pzx \& Pzy)$

$U_{\text{def}}$ ) Two objects *underlap* iff there is an object of which they are both parts.

$Uxy =_{\text{def}} \exists z (Pxz \& Pyz)$

$PP_{\text{def}}$ ) Any part of an object is a *proper part* iff it is not identical with that object

---

<sup>161</sup> For an excellent introduction, see Barwise, Etchemendy (2002).

$$PP_{xy} =_{\text{def}} P_{xy} \ \& \ \neg P_{yx}$$

Adding to (PL.1)-(PL.3) the supplementation principle (PS) allows us to derive the perfect extensionality of the domain (PE):

PS) If an object is not part of another, some part of the former does not overlap the latter.

$$\neg P_{yx} \rightarrow \exists z (P_{zy} \ \& \ \neg O_{zx})$$

PE) Two objects are identical iff they have the same parts.

$$(\exists z PP_{zx} \vee \exists z PP_{zy}) \rightarrow (x=y \leftrightarrow \forall z (PP_{zx} \leftrightarrow PP_{zy}))$$

If we introduce an axiom for a mereological upper bound, i.e. an object every thing is part of, we can state the principle of unrestricted summation:

U) There is a maximal element of which everything is part.

$$\exists z \forall x P_{xz}$$

PC.1) For any two objects, there is a smallest thing of which they are parts.

$$\exists z \forall w (O_{wz} \leftrightarrow (O_{wx} \vee O_{wy}))$$

Since the theory is extensional, there is a unique individual that is the sum of any two objects: if  $\iota$  is a description operator, we can define a sum operator (+) as follows:

$$x + y =_{\text{def}} \iota z \forall w (O_{wz} \leftrightarrow (O_{wx} \vee O_{wy}))$$

Since there is no null element, the dual of (PC.1) – the product – is stated in conditional form:

PC.2) If two things overlap, there is a largest thing that is a part of both.

$$O_{xy} \rightarrow \exists z \forall w (P_{wz} \leftrightarrow (P_{wx} \ \& \ P_{wy}))$$

Finally, we need to add the characteristic Democritean flavor:

A<sub>def</sub>) Any object is *atomic* iff it has no proper parts.

$$Ax =_{\text{def}} \neg \exists y PPyx$$

AT) Everything is ultimately composed by atomic objects.

$$\exists y (Pyx \ \& \ Ay)^{162}$$

Given the mereological basis, the following topological axioms and definitions should be straightforward:

TL.1) Everything is connected to itself.

$$Cxx$$

TL.2) If one thing is connected to another, then also the latter is connected to the first.

$$Cxy \rightarrow Cyx$$

E<sub>def</sub>) One thing is *enclosed* in another iff everything connected to the first is also connected to the second.

$$Exy =_{\text{def}} \forall z (Czx \rightarrow Czy)$$

TL.3) If one thing is a part of another, everything connected to the first is connected to the second.

$$Pxz \rightarrow Exy$$

IPP<sub>def</sub>) One thing is an *internal proper part of* another iff the first is a proper part of the second and everything connected to the first overlaps the second.

$$IPPxy =_{\text{def}} PPyx \ \& \ \forall z (Czx \rightarrow Ozy)$$

TPP<sub>def</sub>) One thing is an *tangential proper part of* another iff the first is a proper part of the second and something connected to the first does not overlap the second.

$$TPPxy =_{\text{def}} PPyx \ \& \ \exists z (Czx \ \& \ \neg Ozy)$$

SC<sub>def</sub>) One thing is *self-connected* iff any two parts that make up the whole of it are connected to each other.

$$SCx =_{\text{def}} \forall y \forall z (\forall w (Owx \leftrightarrow Owz) \vee Owy \rightarrow Cyz)$$

We are now in a position to list the principles characterizing *Life* (for readability, let us

---

<sup>162</sup> Atoms allow for an alternative, simpler axiomatization: (AT) and (PS) can be replaced by:  $\neg Pxy \rightarrow \exists z (Az \ \& \ Pzx \ \& \ \neg Pzy)$ , which implies an atomistic version of (PE):  $x=y \leftrightarrow \forall z (Az \rightarrow (Pzx \leftrightarrow Pzy))$ .

use  $\exists^n x \phi x$  as an abbreviation for ‘there are  $n$  distinct objects  $\phi$ -ing’):

LL.1) *Being alive* and *Being dead* exclusively and exhaustively define each cell's state.

$$Ax \rightarrow (Wx \vee Bx)$$

$N_{\text{def}}$ ) One cell is another's neighbor iff they are connected.

$$Nxy =_{\text{def}} Ax \ \& \ Ay \ \& \ Cxy$$

$NH_{\text{def}}$ ) A cell's *neighborhood* is the mereological sum of its neighbors.

$$NH(x) =_{\text{def}} \iota z \forall w (Pwz \leftrightarrow Nwx)$$

NA) Each cell has exactly nine neighbors.

$$\exists^9 y Nxy$$

Alternatively, we could use lexical axioms for *neighbor* and define *connection* accordingly:

NL.1) Only two atoms can be each other neighbors.

$$Nxy \rightarrow Ax \ \& \ Ay$$

NL.2) Every atom is neighbor of itself.

$$Nxx$$

NL.3) If one atom is neighbor of another, then also the latter is neighbor of the first.

$$Nxy \rightarrow Nyx$$

$C_{\text{def}}$ )  $x$  and  $y$  are *connected* iff there is one  $x$ -atom which is the *neighbor* of one  $y$ -atom

$$Cxy =_{\text{def}} \exists z (Az \ \& \ Pzx) \ \& \ \exists w (Aw \ \& \ Pwy) \ \& \ Nzw$$

Finally, we add a “Finite Nature” axiom to the effect that the number of atoms in the domain is #:

FN) There are # atoms.

$$\exists^{\#} x Ax$$

### I.3 Axioms from Chapter 2

Adding a temporal dimension to *Life* requires a modification to the basic language of mereotopology: in what follows we shall use the symbol “ $t$ ” as a variable ranging only over entities in the domain which can be intuitively regarded as time instants. The following axioms therefore systematize the concept of time in *Life*:

TIL.1) No time instant precedes itself.

$$\neg >tt$$

TIL.2) If  $t$  precedes  $t'$  and  $t'$  precedes  $t''$ , then  $t$  precedes  $t''$ .

$$>tt' \ \& \ >t't'' \rightarrow >tt''$$

TIL.3) If  $t$  and  $t'$  are distinct, either  $t$  precedes  $t'$  or  $t'$  precedes  $t$ .

$$\neg t = t' \rightarrow >tt' \vee >t't$$

IS<sub>def</sub>)  $t'$  is the *immediate successor* of  $t$  iff  $t$  precedes  $t'$  and  $t$  does not precede any other instant preceding  $t'$ .

$$IS t't =_{\text{def}} >tt' \ \& \ \forall t''' (>t'''t' \rightarrow \neg >tt''')$$

IP<sub>def</sub>)  $t$  is the *immediate predecessor* of  $t'$  iff  $t'$  is the immediate successor of  $t$ .

$$IP tt' =_{\text{def}} IS t't$$

TIF) There is an instant with an immediate successor but no immediate predecessor (i.e. the first instant) and an there is an instant with an immediate predecessor but no immediate successor (i.e. the last instant); any other instant has both an immediate predecessor and an immediate successor.

$$BBt =_{\text{def}} \neg \exists t' IP t't$$

$$BCt =_{\text{def}} \neg \exists t' IS t't$$

$$TIF) (\neg BBt \ \& \ \neg BCt) \rightarrow (\exists t' IS t't \ \& \ \exists t'' IP t''t)$$

Now that we have variables ranging over well-behaving time instants, we need to connect time with the atomic objects in our universe and update the “Finite Nature” axiom accordingly. We introduce the primitive predicate “ $Ixt$ ” ranging over atoms and times:

CT) Any atomic cell exist at one instant of time.

$$Ax \rightarrow \iota Ixt$$

FN) For each  $t$ , there are  $\#$  atomic cells at  $t$ .

$$\exists^{\#}x (Ax \ \& \ Ixt)$$

The final part of the axiomatic theory should spell out *Life* characteristic dynamics: the state of a given cell in the lattice depends only on a specific subset of the lattice. We can use the expressive resource of mereotopology to precisely model the phenomenon – first, we need to introduce the notion of *temporal connection*:

TC<sub>def</sub>)  $x$  existing at  $t_x$  and  $y$  existing at  $t_y$  are *temporally connected* iff an  $x$ -atom is connected to an  $y$ -atom and one of  $\langle t_x, t_y \rangle$  is the immediate successor of the other.

$$TC_{xy} =_{\text{def}} \exists z \exists t (Izt \ \& \ Pzx) \ \& \ \exists v \exists t' (Ivt' \ \& \ Pvy) \ \& \ (IS_{tt'} \vee IS_{t't})$$

Two cells are now *neighbors* if they are connected and exist in the same instant:

N<sub>def</sub>) Cell  $x$  existing at  $t_x$  is *neighbor* of cell  $y$  existing at  $t_y$  iff  $x$  is connected to  $y$  and  $t_x = t_y$ .

$$N_{xy} =_{\text{def}} \exists t Ixt \ \& \ \exists t' Iyt' \ \& \ C_{xy} \ \& \ t = t'.$$

We are now in a position to define the notion of successor/predecessor for a given cell and state *Life* dynamics:

ICS<sub>def</sub>)  $x$  existing at  $t_n$  is *the immediate successor* of  $y$  existing at  $t_m$  iff  $t_n$  is the immediate successor of  $t_m$  and every neighbor of  $x$  is temporally connected to  $y$ .

$$ICS_{xy} =_{\text{def}} \exists t' Ixt' \ \& \ \exists t Iyt \ \& \ IS_{t't} \ \& \ \forall z (N_{zx} \rightarrow TC_{zy})$$

ICP<sub>def</sub>)  $x$  existing at  $t_n$  is *the immediate predecessor* of  $y$  existing at  $t_m$  iff  $y$  is the immediate successor of  $x$ .

$$ICP_{xy} =_{\text{def}} ICS_{yx}$$

TCS) Any atomic cell has one *immediate predecessor/successor* (except for cell living at the last/first instant of time).

$$Ax \rightarrow \iota y ICS_{yx} \ \& \ \iota z ICP_{zx}$$

LN) If  $x$  is the immediate successor of  $y$ ,  $x$  is alive if  $y$  is alive and two or three neighbors of  $y$  are alive, or if  $y$  is dead and three neighbors of  $y$  are alive;  $x$  is dead

otherwise.

$$\text{ICS}_{xy} \rightarrow \{ [(Wy \ \& \ (\exists^2 z \ (Nzx \ \& \ Wz) \vee \exists^3 z \ (Nzx \ \& \ Wz))) \vee (By \ \& \ \exists^3 z \ (Nzx \ \& \ Wz))] \\ \rightarrow Wx \}$$

#### *I.4 Concluding Remarks*

The above list of axioms is not intended as a full-fledged formal specification of the ontology, since many technical features of the language were not defined in details (in the end, since our main concern is computational, the formal details are less important than the choices made during the implementation). However, we believe they provide an appealing starting point for the reader to get the feeling of the logical structure of the digital ontology. We hope to turn very soon to formal philosophy to fully develop these sketchy remarks.

## Bibliography

If you copy from one author, it's plagiarism.

If you copy from two, it's research.

*Wilson Mizner*

Aaronson, S., *forthcoming*, 'Why Philosophers Should Care About Computational Complexity', in B. J. Copeland, C. J. Posy, O. Shagrir (eds.), *Computability: Gödel, Turing, Church, and Beyond*, MIT Press.

Armstrong, D., 1983, *What Is a Law of Nature?*, Cambridge: Cambridge University Press.

Barwise, J., Etchemendy J., 2002, *Language, Proof and Logic*, Stanford: CSLI Publications.

Batty, M., 2005, *Cities and Complexity, Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals*, Cambridge: MIT Press.

Beane, S. R., Davoudi Z., Savage M., J., 2012, 'Constraints on the Universe as a Numerical Simulation', <http://arxiv.org/abs/1210.1847>.

Bedau, M., 1997, 'Weak Emergence', in J. Tomberlin (ed.), *Philosophical Perspectives, 11: Mind, Causation, and World*, Blackwell Publishers, pp. 375-399.

Berto, F., Carrara M., 2009, 'To Exist and to Count: a Note on the Minimalist View', *Dialectica*, 63 (3): 343-356.

Berto, F., Rossi G., Tagliabue J., 2010, *The Mathematics of the Models of Reference*, London: College Publications.

Berto, F., Tagliabue J., 2012, 'Cellular Automata', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.

Berto, F., Tagliabue J., *forthcoming*, 'Either the World is Digital or Not'.

Berlekamp, E., Conway J., Guy R., 1982, *Winning Ways for Your Mathematical Plays (Vol. 2)*, London: Academic Press.

- Bermúdez, J. L., 2005, *Philosophy of Psychology: A Contemporary Introduction*. London: Routledge.
- Bittner, T., Smith B., 2003, ‘A Theory of Granular Partitions’, in M. Duckham, M. F. Goodchild, M. F. Worboys (eds.), Taylor & Francis Books, pp. 117-151.
- Blackburn, P., Bos J., Striegnit K., 2006, *Learn Prolog Now!*, London: College Publications.
- Boolos, G., Burgess J. P., Jeffrey R. C., 2002, *Computability and Logic*, 4<sup>th</sup> Edition, Cambridge: Cambridge University Press.
- Borgo, S., Lesmo L., 2008, eds., *Formal Ontologies Meet Industry*, Amsterdam: IOS Press.
- Bratko, I., 2011, *Prolog Programming for Artificial Intelligence*, 4<sup>th</sup> Edition, Dorchester: Addison-Wesley.
- Cabello, A., 2013, ‘The Contextual Computer’, in Zenil (2013), pp. 595-604.
- Cappelen, H., 2012, *Philosophy without Intuitions*, New York: Oxford University Press.
- Casati, R., Varzi A., 1999, *Parts and Places*, Cambridge: MIT Press.
- Chalmers, D., 1996, *The Conscious Mind*, Oxford: Oxford University Press.
- Chalmers, D., 2002, ‘Strong and Weak Emergence’, in P. Clayton, P. Davies (eds.), *The Re-Emergence of Emergence*, Oxford University Press, pp. 244-255.
- Chalmers, D., 2012, *Constructing the World*, Oxford: Oxford University Press.
- Chen, H., Chen S., Doolen G., Lee Y.C., 1983, ‘Simple Lattice Gas Models for Waves’, *Complex Systems*, 2: 259–267.
- Clarke, B. L., 1981, ‘A Calculus of Individuals Based on “Connection”’, *Notre Dame Journal of Formal Logic*, 22: 204-218.
- Cohn, A. G., Renz J., 2008, ‘Qualitative Spatial Representation and Reasoning’, in F. van Hermelen, V. Lifschitz, B. Porter (eds.), *Handbook of Knowledge Representation*, Elsevier, pp. 551-596.
- Creutz, M., 1986, ‘Deterministic Ising dynamics’, *Annals of Physics*, 167: 62-76.

- Crutchfield, J. P., 1994, 'The Calculi of Emergence: Computation, Dynamics, and Induction', *Physica D*, 75: 11-54.
- Crutchfield, J. P., Mitchell M., 1995, 'The Evolution of Emergent Computation', *Proceedings of the National Academy of Sciences*, 92(23): 10742.
- Davidson, D., 1970, 'Mental Events', in L. Foster, J. W. Swanson (eds.), *Experience and Theory*, Duckworth, 79-102.
- De Laguna, T., 1922, 'The Nature of Space-I', *Journal of Philosophy*, 19: 393-407.
- De Roever, W.P., Engelhardt K., 1998, *Data refinement: Model-oriented proof methods and their comparison*, Cambridge: Cambridge University Press.
- Dennett, D., 1991, 'Real Patterns', *Journal of Philosophy*, 88: 27-51.
- Divers J., 2002, *Possible Worlds*, London: Routledge.
- Eklund, M., 2007, *The Picture of Reality as an Amorphous Lump*, in J. Hawthorne, T. Sider, D. Zimmerman (eds.), *Contemporary Debates in Metaphysics*, Blackwell, pp. 382-396.
- Einstein, A., 1905, 'Zur Elektrodynamik bewegter Körper', *Annalen der Physik*, 17: 891.
- Einstein, A., 1915, 'Die Feldgleichungen der Gravitation', *Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin*: 844-847.
- Epstein, J.M., 1999, 'Agent-based Computational Models and Generative Social Science', *Complexity*, 4(5):41-60.
- Evans, G., 1978, 'Can There Be Vague Objects?', *Analysis*, 38 (4): 208.
- Ferrario, R., Oltramari A., 2009, eds., *Formal Ontologies Meet Industry*, Amsterdam: IOS Press.
- Feynman, R. P., 1965, *The Character of Physical Law*, Cambridge: MIT Press.
- Fine, K., 1975, 'Review of Lewis (1973a)', *Mind*, 84: 451-458.
- Fine, K., 1999, 'Things and their Parts', *Midwest Studies in Philosophy*, 23 (1): 61-74.

- Flegg, G. H., 2001, *From Geometry to Topology*, New York: Dover Publications.
- Floridi, L., 2009, 'Against Digital Ontology', *Synthese*, 168 (1): 151-178.
- Floridi, L., 2011, *The Philosophy of Information*, Oxford: Oxford University Press.
- Franceschetti, D. R., Campbell B. W., Hamneken J. W., 1992, 'Hamming Nets, Ising Sets, Cellular Automata, Neural Nets and Random Walks', *American Journal of Physics*, 61: 50-53.
- Fredkin, E., Toffoli T., 1982, 'Conservative logic', *International Journal of Theoretical Physics*, 21: 219-253.
- Fredkin, E., 1993, 'A New Cosmogony', in *PhysComp '92: Proceedings of the Workshop on Physics and Computation*, IEEE Computer Society Press, pp. 116-121.
- Friedman, M., 1987, 'Carnap's Aufbau Reconsidered', *Noûs*, 21 (4): 521-545.
- Galton, A., 2000, *Qualitative Spatial Change*, Oxford: Oxford University Press.
- Gödel, K., 1931, 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme', *Monatshefte für Mathematik und Physik*, 38: 173-98.
- Grim, P., Mark G. R., Denis P. St., 1998, *The Philosophical Computer: Exploratory Essays in Philosophical Computer Modeling*, Cambridge: MIT Press.
- Grice, H.P., 1989, *Studies in the Way of Words*, Cambridge: Harvard University Press..
- Guizzardi, G., 2005, *Ontological Foundations for Structural Conceptual Models*, Twente: Telematica Instituut Fundamental Research Series No. 15.
- Hanson, J. E., Crutchfield J. P., 1992, 'The Attractor-Basin Portrait of a Cellular Automaton', *Journal of Statistical Physics*, 66: 1415-1462.
- Hintikka, J., 1999, 'The Emperor's New Intuitions', *Journal of Philosophy*, 96 (3): 127-147.
- Hoare C. A. R., He J., 1998, *Unifying Theories of Programming*, London: Prentice Hall.
- Hodges, W., Lewis D. K., 1968, 'Finitude and Infinitude in the Atomic Calculus of Individuals', *Noûs*, 2: 405-410.

- Hordijk, W., Crutchfield J. P., Mitchell M., 1996, 'Embedded Particle Computation in Evolved Cellular Automata', in *Proceedings of the Conference on Physics and Computation PhysComp96*.
- Hovda, P., 2008, 'Quantifying Weak Emergence', *Minds and Machine*, 18 (4): 461-473.
- Hovda, P., *forthcoming*, 'Tensed Mereology'.
- Ilachinski, A., 2001, *Cellular Automata*, Singapore: World Scientific Publishing.
- Ilachinski, A., 2004, *Artificial War*, Singapore: World Scientific Publishing.
- Jubien, M., 1993, *Ontology, Modality, and the Fallacy of Reference*, Cambridge: Cambridge University Press.
- Keefe, R., 2000, *Theories of Vagueness*, Cambridge: Cambridge University Press.
- Kim, J., 1993, *Supervenience and Mind*, Cambridge: Cambridge University Press.
- Korta, K., Perry J., 2012, 'Pragmatics', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.
- Koslicki, K., 2008, *The Structure of Objects*, Oxford: Oxford University Press.
- Kripke, Saul. 1980, *Naming and Necessity*, Cambridge: Harvard University Press.
- Ladyman, J., Ross D., 2007, *Every Thing Must Go*, Oxford: Oxford University Press.
- Kurzweil, R., 2005, *The Singularity is Near*, New York: Viking Penguin.
- Lewis, D. K., 1969, *Convention*, Cambridge: Harvard University Press.
- Lewis, D. K., 1973a, *Counterfactuals*, Oxford: Blackwell Publishers.
- Lewis, D. K., 1973b, 'Causation', *Journal of Philosophy*, 70: 556-567.
- Lewis, D. K., 1983a, *Philosophical Papers, Volume I*, Oxford: Oxford University Press.
- Lewis, D. K., 1983b, 'New Work for a Theory of Universals', *Australasian Journal of Philosophy*, 61: 343-377.
- Lewis, D. K., 1986a, *Philosophical Papers, Volume II*, Oxford: Oxford University Press.

- Lewis, D. K., 1986b, *On the Plurality of Worlds*, Oxford: Blackwell Publishers.
- Lewis, D. K., 1988, 'Vague Identity: Evans Misunderstood', *Analysis*, 48 (3): 128-130.
- Lewis D. K., 1991, *Parts of Classes*, Oxford: Blackwell.
- Lewis, D. K., 1994, 'Humean Supervenience Debugged', *Mind*, 103: 473-490.
- Lewis, D. K., 1999, *Papers in Metaphysics and Epistemology*, Cambridge: Cambridge University Press.
- Lloyd, S., 1997, 'Universe as Quantum Computer', *Complexity*, 3 (1): 32-35.
- Lloyd, S., 2007, *Programming the universe : a quantum computer scientist takes on the cosmos*, New York: Knopf, Borzoi Books.
- Lloyd, S., 2013, 'The Universe as Quantum Computer', in Zenil (2013), pp. 567-582.
- Mainzer, K., Chua L., 2012, *The Universe as Automaton. From Simplicity and Symmetry to Complexity*, Heidelberg: Springer.
- Marconi, D., 1996, 'On the referential competence of some machines', *Artificial Intelligence Review*, 10 (1-2): 21-35.
- Marconi, D., 1997, *Lexical Competence*, Cambridge: MIT Press.
- Marr, D., 1982, *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*, San Francisco: W.H. Freeman and Company.
- Markosian, N., 2010, 'Time', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.
- Miller, J., Page S., 2007, *Complex Adaptive System*, Princeton: Princeton University Press.
- Moore, C., Mertens S., 2011, *The Nature of Computation*, Oxford: Oxford University Press.
- Oppenheimer, P. E., Zalta E. N., 1991, 'On the Logic of the Ontological Argument', in James Tomberlin (ed.), *Philosophical Perspectives 5: The Philosophy of Religion*, pp. 509-529.

- Piccinini G., 2010, 'Computation in Physical Systems', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.
- Pontow, C., Dazinger F., Schubert R., 2007, 'A Framework in Prolog for Computing Structural Relationships', *Data & Knowledge Engineering*, 62 (2): 308-326.
- Purdum, J., 2012, *Beginning Object-Oriented Programming with C#*, Indianapolis: Wrox.
- Putnam, H., 1982, 'Why There Isn't a Ready-Made World', *Synthese*, 51: 141-67.
- Ramsey, F., 1978, *Foundations*, London: Routledge and Kegan Paul.
- Reas, C., Fry, B., 2008, *Getting Started with Processing*, Sebastopol: O'Reilly.
- Ross, D., 2000, 'Rainforest Realism: A Dennettian Theory of Existence', in D. Ross, A. Brook, D. Thompson (eds.), *Dennetts Philosophy A Comprehensive Assessment*, MIT Press, pp. 147-168.
- Rossi, M., Rossi G., Sommaggio P., 2010, *La Legge nella Società della Semi-Immortalità*, Milano: Lampi di Stampa.
- Rossi, M., Tagliabue J., 2009, 'Le Tre Maschere del Convenzionalista', *Rivista di Estetica*, 41: 109-124.
- Schildt, H., 2011, *Java, A Beginner's Guide, 5th Edition*, New York: McGraw-Hill Osborne Media.
- Shaffer, J., 2003, 'Is There a Fundamental Level?', *Noûs*, 37 (3): 498-517.
- Shiffman, D., 2008, *Learning Processing: A Beginner's Guide to Programming Images, Animation, and Interaction*, Burlington: Morgan Kaufmann.
- Shiffman, D., 2012, *The Nature of Code: Simulating Natural Systems with Processing*, Version 1.0.
- Sidelle, A., 1989, *Necessity, Essence, Individuation*, Ithaca: Cornell University Press.
- Sidelle, A., 1992, 'Identity and the Identity-Like', *Philosophical Topics*, 20: 269-192.
- Sider, T., 2001, *Four-Dimensionalism: An Ontology of Persistence and Time*, Oxford: Oxford University Press.

- Simon, H. A., 1962, 'The Architecture of Complexity', *Proceedings of the American Philosophical Society*, 106 (6): 467-482.
- Smith, B., Brogaard B., 2002, 'Quantum mereotopology', *Annals of Mathematics and Artificial Intelligence*, 36: 153-175.
- Sosa, E., 1999, 'Existential Relativity', *Midwest Studies in Philosophy*, XXIII: 132-143.
- Sperber, D., Wilson D., 1986, *Relevance: Communication and Cognition*, Oxford: Blackwell.
- Tagliabue, J., 2013, 'A New Kind of Philosophy. Manifesto for a Digital Ontology', in H. Zenil (ed.), *Irreducibility and Computational Equivalence*, Springer, pp. 325-339.
- Turing, A. M., 1936, 'On Computable Numbers with an Application to the Entscheidungsproblem', *Proceeding of the London Mathematical Society*, 42: 230-265.
- Turing, A. M., 1950, 'Computing machinery and intelligence', *Minds and Machines*, 59: 433-460.
- Van Inwagen, P., 1990, *Material Beings*, Ithaca: Cornell University Press.
- Van Inwagen, P., 1994, 'Composition as Identity', in *Philosophical Perspectives, Vol. 8, Logic and Language*, pp. 207-220.
- Varzi, A. C., 2000, 'Mereological Commitments', *Dialectica*, 54 (4): 283-305.
- Varzi, A. C., 2005a, 'Change, Temporal Parts, and the Argument from Vagueness', *Dialectica*, 59 (4): 485-498.
- Varzi, A. C., 2005b, 'Teoria e pratica dei confini', *Sistemi intelligenti*, 17 (3): 399-418.
- Varzi, A. C., 2009, 'Mereology', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.
- Wheeler, J.A., 1990, 'Information, Physics, Quantum: The Search for Links', in W. Zurek (ed.), *Complexity, Entropy, and the Physics of Information*, Addison-Wesley, pp. 309-336.
- Williamson, T., 1996 *Vagueness*, London: Routledge.
- Williamson, T., 2008, *The Philosophy of Philosophy*, Oxford: Blackwell

Wolfram, S., 2002, *A New Kind of Science*, Champaign: Wolfram Media.

Yalowitz, S., 2012, 'Anomalous Monism', in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*.

Woodward, J., 2007, *Making Things Happen*, Oxford: Oxford University Press.

Zenil, H., 2013, ed., *A Computable Universe*, Singapore: World Scientific Publishing.

Zuse, K., 1982, 'The Computing Universe', *International Journal of Theoretical Physics*, 21: 589-600.