INVERTED THEORY NETWORKS

By
Nico Christodoulides

SUBMITTED IN FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN MATHEMATICS
AT
UNIVERSITY OF CAPE TOWN
CAPE TOWN, SOUTH AFRICA
DECEMBER 2005

UNIVERSITY OF CAPE TOWN
DEPARTMENT OF
MATHEMATICS AND APPLIED MATHEMATICS

The undersigned hereby certify that they have read and recommend to the Faculty of Science for acceptance a thesis entitled "**Inverted Theory Networks**" by **Nico Christodoulides** in fulfillment of the requirements for the degree of **Master of Science in Mathematics**.

Dated: December 2005

Supervisor: _____
Ingrid Rewitzky

Readers: _____
...

_____
...

UNIVERSITY OF CAPE TOWN

Date: **December 2005**

Author: **Nico Christodoulides**

Title: **Inverted Theory Networks**

Department: **Mathematics and Applied Mathematics**

Degree: **M.Sc.** Convocation: **December** Year: **2005**

Signature of Author

To Nikki

Kiana

and Mikael

my Earth

Sun

and Moon

Table of Contents

Abstract

The logicatom is defined, and it is argued that this represents the quantum of knowledge.
Theory networks encapsulating a set of logicatoms and the dynamic relations between them, are defined. It is shown that these structures can emulate cellular automaton systems and in particular, simulate universal Turing machines.
The regulating principle of natural selection is formalised together with its necessary and sufficient conditions. It is proven that there exists inverted theory networks (an analogous construct to theory networks) that satisfy all the requirements specified for natural selection to regulate their dynamics. The applicability of inverted theory networks to modelling thought is analysed. Further, inverted theory networks are proposed as a candidate for the pregeometry hypothesised by Wheeler.

Acknowledgements

This thesis represents a long and arduous journey in my life. Along this journey, many mentors, friends and loved ones carried me when I had given up and thrown in the towel. I would thus like to take this time to briefly describe this journey and portray the heroines in my story.

When I was fourteen years old, I distinctly remember sitting at my desk, bouncing a tennis ball on a maths book and promising myself that I will one day understand why the ball falls. This promise led me to learn and love mathematics and physics, resulting in me studying towards a masters degree in cosmology at the University of Cape Town in 1994. My thesis topic at that time entailed the analysis of Berry's phase, a phenomenon in quantum physics that is solely explained using geometric concepts. The idea was to see if all of quantum theory could be rewritten using geometrical concepts, with the final objective being to 'geometrise quantum theory' (as opposed to quantising general relativity). Needless to say, after two years, I had achieved nothing except frustration and confusion. My supervisor then stipulated that I write a review of all research done in this field in order to complete my degree. From my perspective, I had failed in my original promise and thus threw in the towel. My dream of becoming an academic who works for the sole purpose of discovering truth collapsed and I entered the financial markets, selling my mathematical skills to asset management companies solely for the accumulation of capital. I had given up searching for the truth and tried to replace it with capital, a clearly inferior and, in my opinion, diametrically opposed objective.

It was at this point that I met my wife, a zoologist who introduced me to the concepts of evolutionary biology. She gave me Richard Dawkins book, "The Selfish Gene". On reading it, that little child that needed to understand why the tennis ball fell, awoke within me once again. I thought he was dead - he was actually only sleeping. Dawkins concept of a meme fascinated me and I decided to attempt to mathematically formalise them. I restarted a masters degree in pure mathematics focusing on logic. On attempting to mathematically formalise memes, I started recognising characteristics in my structures that reminded me of the research I had done ten years previously. Here the writings of John Wheeler guided me in seeing the similarities between my structure and his hypothesised pregeometry. In effect, the last ten years have led me full circle and I believe that I have come some way in solving that original thesis topic.

Every story has its villains, heroes and heroines. They are all fundamental in bringing the saga to completion. As I worked on my thesis I remember telling various professors in mathematics,

physics and biology what I believed: That the universe was one big brain - that the same regulating principle that constructed unalterable beliefs in my brain created the unalterable physical laws that we observe - that the mathematical platform used to describe dynamical knowledge in my head and that used to describe physical law was one in the same - that physical law itself evolved by natural selection - that the reason the ball fell was because general relativity ended up being one of the fittest laws that survived. These statements evoked both scorn and laughter and it was at this point that my mentors and loved ones came to the fore. Firstly to my wife Nikki, who never stopped believing in me, even though I often did. Thank you. Without you, none of this would have been possible. Secondly, to my supervisor Ingrid who allowed me the academic freedom to express these ideas and helped me sculpt this work into the robust mathematical proofs that constitute this thesis. Thank you. Without your input and guidance, $(TEA)^{-1}$ theory would have remained a hand waving exercise.

Cape Town, South Africa
December 29, 2005

Nico Christodoulides

Chapter 1

Introduction

The three postulates that form the foundation of this thesis are introduced. These motivate the problem statement and thesis objective. I conclude with an overview of the remainder of the thesis.

1.1 Background

Bertrand Russell, Richard Dawkins and John Archibald Wheeler provide me with the postulates from whence this thesis arises.

1.1.1 Russell's Logical Atomism

The primary hypothesis adopted in this thesis is that of **logical atomism.** Ludwig Wittgenstein and Bertrand Russell were the prime exponents of this philosophy. The logical atomism view of reality assumes that all knowledge must begin with sensory experience. Genuine information about the world must be acquired by *a posteriori* means, so that nothing can be thought without first being sensed. From this beginning, Russell argued that everything else follows by logical analysis. Simple facts like 'It is raining' are the atomic facts or 'logical atoms' upon which all human knowledge is grounded. In particular, Russell claims in the fifth chapter of The Problems of Philosophy (1912): **"Every proposition which we can understand must be composed wholly of constituents with which we are acquainted."** This statement forms the founding argument in the formal definition of a 'logical atom'.

1.1.2 Dawkins' Meme

Richard Dawkins hypothesised the existence of a 'meme' [15, 16]. In order to define a meme, he referred to the analogous construct, the gene. Biological organisms are defined by their genotype and

their phenotype [26]. The genotype (nucleic acids) represents the underlying genetic coding while the phenotype (proteins) is the expression of the genotype within an environment. Dawkins defined the meme as "a unit of information residing in the human brain" [15]. Just as the phenotype of a particular gene complex in a species determines a particular trait e.g. blue eyes in human beings, the phenotype of a meme complex represents a concept that can be understood, learnt or sensed. This can be represented as a collection of words, music or visual images. One can view Dawkins meme as equivalent to Russell's logical atom. However, Dawkins' genius came in observing the regulating principle of these entities. Dawkins hypothesised that the dynamic behaviour of memes is governed by **natural selection**. From an intuitive perspective, consider the following example. This thesis represents the phenotype of a meme complex existing in the author's brain. By reading it, the reader has allowed the meme complex to make a copy of itself in the reader's brain. Thus memes have the property of **reproduction**. Now the reader will understand this thesis in a different way to the author (or any other reader for that matter) due to the incoming knowledge interacting with the existing knowledge in the reader's head. Thus the meme complex can be said to have **mutated** as a copy was made. Finally, depending on whether the reader thinks this thesis is of any value to the scientific community or not, he/she may recommend others to read it, or he/she might forget entirely about it. Thus the meme complex exhibits the property of **differential fitness** i.e. its spread and survival depends upon its makeup - in this particular instance, its acceptance within the scientific community. This fitness may be quantified by the number of citations in future scientific work. The three properties stated in bold are exactly the necessary requirements of natural selection [17, 47]. The second hypothesis is encapsulated in the statement: **'Natural selection acts on memes and regulates their survival, resulting in the fittest meme surviving.'**

1.1.3 Wheeler's Pregeometry

Einstein's theory of general relativity [22] elevated the importance of the underlying spacetime structure in physics. Prior to the theory, the spacetime continuum was regarded as the arena in which the laws of physics act. Einstein's field equations dictated that energy curved spacetime and spacetime in turn prescribed the dynamics of classical energy. In Wheeler's words [55], general relativity "dethroned spacetime from a post of preordained perfection high above the battles of matter and energy, and marked it as a new dynamic entity participating actively in this combat." What was previously perceived as a gravitational force field is now known to be the effects of curved spacetime. Further, physical laws such as the conservation of energy and momentum ended up being a mathematical consequence of the geometry. Misner and Wheeler took these beautiful concepts to

the next logical step by asking the question: 'Is the spacetime continuum all there is to physics?' In other words, can curved spacetime *solely* represent all the laws of physics. To answer this question, the theory of geometrodynamics was born. Geometrodynamics is the study of the geometry of curved *empty* space and the relative dynamics of subspaces therein, as prescribed by the Einstein field equations. Misner and Wheeler [78] went some way to show that classical physics embodying gravitation, electromagnetism, non-quantised charge and non-quantised mass can be represented as purely geometrical phenomena. This theory reached its explanatory limit when attempting to discuss quantum phenomena. The limitation in the theory was identified in that it was constrained to operate in a differentiable manifold. There was no natural way of modelling the dynamics in the underlying topology. To overcome this barrier, Wheeler hypothesised the existence of a **'pregeometry'**. Wheeler argued that spacetime itself must be understood in terms of the more fundamental structure. The underlying principle of such a structure was to be found in its simplicity. In particular, Wheeler stated [79]: "All of physics, in my view, will be seen someday to follow the pattern of thermodynamics and statistical mechanics, of regularity based on chaos, of 'law without law'. Specifically, I believe that everything is built higgledy-piggledy on the unpredictable outcomes of billions and billions of elementary quantum phenomena, and that the laws and initial conditions of physics arise out of this chaos by the action of a regulating principle, the discovery and proper formulation of which is the number one task of the coming third era of physics."

1.2 Problem statement and thesis objective

These disparate topics are linked in the following way: Russell's logical atom is analogous to Dawkins' meme. Russell's postulate describes properties regarding the quanta of thought; Dawkins postulates what regulates these quanta. The question as to what this has to do with physics and Wheeler's pregeometry, comprising the hypothesised fundamental building blocks of physical law was elegantly answered by G.F.R. Ellis in [23]: "Human thoughts can cause real physical effects." If I have the intention of picking up a stone and throwing it, the result would be the physical effect of a stone hurtling through the air. "At present there is no way to express this interaction in the language of physics, even though our causal schemes are manifestly incomplete if this is not taken into account. The minimum requirement to do so is to include the relevant variables in the space of variables considered. That then makes these variables and their effects a part of physics - or perhaps of fundamental physics". Thus Wheeler's pregeometry must comprise the 'variables' that model intent i.e. thought. My hypothesis is that the structure that models the quanta of thought is Wheeler's

hypothesised pregeometry. Further, the regulating principle sought after by Wheeler is none other than Darwin's law of natural selection, originally suggested in 1859 as the principle mechanism of evolutionary change.

The above paragraph guided my research program resulting in the question **"Can a formal paradigm be created in which to model these postulates?"**. I split this problem statement up into 2 objectives: The construction objective encompasses defining a formal mathematical space comprised of entities that represent Dawkins' memes or Russell's logical atoms. Further, this objective encapsulates showing that the dynamics of the space is regulated by the principle of natural selection. The application analysis objective encompasses investigating if this space can be applied to analysing the dynamical properties of knowledge and whether it serves as a candidate for modelling pregeometries in physics. Needing to define a space that comprises a set of elements representing knowledge, I naturally enter the formal arena of knowledge representation - description logics. Research within this broad mathematical arena is guided by Russell's claim that "every proposition which we can understand must be composed wholly of constituents with which we are acquainted". This is interpreted as saying that 'new' knowledge is made up of 'existing' knowledge i.e. all knowledge is comprised of knowledge. I use this to define the basic entity of my space - the *logicatom*. I then proceed to formally construct platforms comprising dynamic sets of these entities i.e. theory networks and inverted theory networks. In order to prove that these structures are regulated by natural selection, I derive the necessary and sufficient requirements for it to be said that natural selection regulates the dynamics of a space. The construction objective is met through the construction of a particular inverted theory network, whereupon I prove that it is regulated by natural selection. The application analysis objective is met since it completely guided the construction of the space under consideration. Various case studies are given throughout the thesis that show the various applications of these structures to multiple fields of study.

1.3 Layout

Chapter 2 begins by providing an overview of knowledge representation using modal logic and proceeds to motivate the definition of the logicatom. Theory networks, a space comprising these entities, are constructed and formally defined using basic maps in modal logic. I then proceed to show how knowledge is actually modelled within theory networks and conclude by showing how these structures can simulate the dynamics of cellular automata systems. All work done in this chapter is the author's original work.

In Chapter 3, the theory of natural selection is formalised resulting in the proofs of the necessary and sufficient conditions for it to regulate a space. The search for theory networks that are regulated by natural selection lead me to an analogous construct, an inverted theory network. The chapter is concluded with a proof showing the existence of an inverted theory networks that satisfies all the requirements of natural selection. All work done in this chapter is the author's original work.

In Chapter 4, I argue why inverted theory networks can be used to define a pregeometry, with specific emphasis on classical physical observables such as the dimension of space-time. Further, I argue how quantum theory could arise within this platform and discuss how this same formalism could model the dynamics of knowledge i.e. thought. Work done in this chapter encompasses a mathematical reformulation of existing research in physics.

Chapter 5 concludes the thesis by detailing exactly what has and has not been achieved. Further research objectives relevant to this work are stated in a single hypothesis.

Appendix A contains the mathematics pertaining to all algorithms used in simulation programs written to numerically verify the proven results.

Chapter 2

Constructing Theory Networks

This chapter introduces theory networks, the formal structure that underpins the foundation of all work in this thesis. Section 2.1 reviews basic concepts and notation used in the basic modal language. Section 2.2 informally motivates the definitions required in the construction of a theory network. Section 2.3 then proceeds to formalise these notions using the language of modal logic. This formalisation will result in a set of tools that will be used to reason about these spaces. Section 2.4 provides various case studies showing how knowledge is described within theory networks. Section 2.5 provides case studies showing how theory networks can simulate the dynamics of cellular automata systems. Section 2.6 concludes the chapter by summarising all results.

2.1 Knowledge representation using modal logic

"Logic is the glue that binds together methods of reasoning, in all domains."

D. Gries and F.B. Schneider [33].

I review propositional calculus and the basic modal language using definitions and examples from Blackburn, de Rijke and Venema's authoritive book on modal logic [7]. Refer to [7] for the historical background concerning modal logic as well as examples showing how modal logic can be used for knowledge representation.

Definition 2.1.1. The language of **classical propositional calculus** (PropCal) is built up using a countable set of propositional variables Φ usually denoted by $p, q, r, \ldots$. The set $Form(\Phi)$ of the well-formed formulas (wff) of PropCal are constructed using the rule:

$$\phi ::= p \mid \bot \mid \neg\psi \mid \psi \vee \eta \mid \psi \wedge \eta$$

where p ranges over all propositional variables in Φ. This means that a formula is either a propositional variable, the propositional contradiction, a negated formula, the disjunction of a formula or the conjunction of a formulae.

Abbreviations for implication $\phi \rightarrow \psi := \neg\phi \vee \psi$, bi-implication $\phi \leftrightarrow \psi := (\phi \leftrightarrow \psi) \wedge (\psi \leftrightarrow \phi)$ and the propositional tautology $\top := \neg\bot$ are used throughout the thesis. If the set of propositional variables $\Phi_n = \{p_1, p_2, \ldots, p_n\}$ is finite, I denote the language as PropCal_n. One introduces the semantics of PropCal using valuations. A valuation can be viewed as a mapping $w : \Phi \rightarrow \{0, 1\}$, assigning to each propositional variable in Φ the values TRUE$(= 1)$ or FALSE$(= 0)$. This is extended inductively to all propositions in $\mathit{Form}(\Phi)$. The set $\mathbf{W}$ of all possible valuations is referred to as the set of possible worlds. Any proposition true under all valuations is called a tautology; conversely, any proposition false under all valuations is called a contradiction. I denote the set of all valuations that map ψ to TRUE by $[\psi] = \{w \mid w(\phi) = 1\} \subseteq W$. $[\psi]$ is known as the *meaning* of ψ. Two propositions ϕ and ψ are said to be logically equivalent $\phi \equiv \psi$ iff $[\phi] = [\psi]$.

PropCal can be axiomatised as a logic using a particular *system*. Examples include Gentzen (natural deduction systems), Beth (Tableau systems) and Hilbert-style systems [64]. In a Hilbert-style system, the following inference rules are supplied to the system:

- *Modus ponens*: From ϕ and $\phi \rightarrow \psi$, infer ψ.
- *Uniform substitution*: From ϕ, infer θ where θ is obtained from ϕ by uniformly replacing propositional variables in ϕ by arbitrary formulae.

It is then proven that they preserve the valuations. Certain tautologies in the system are chosen as axioms e.g. $\neg\neg p \leftrightarrow p$. One then proceeds to show that these axioms are sound in that every theorem produced from these axioms and the inference rules is a tautology. Completeness is proven by showing that every tautology is in fact a theorem. PropCal will be used to describe the logicatom I wish to model. However, as will be shown in Section 2.2, a relational structure is, by construction, a necessity in the definition of the logicatom. Modal logic provides one with a simple, yet expressive language for talking about relational structures.

Definition 2.1.2. The *basic modal language* $\mathit{ML}(\Diamond, \Phi)$ is built up using a set of propositional variables Φ and a unary modal operator $\Diamond$ ('diamond'). The well-formed formulas $\phi \in \mathit{Form}(\Diamond, \Phi)$ of the basic modal language are given by the rule:

$$\phi ::= p \mid \bot \mid \neg\psi \mid \psi \vee \eta \mid \psi \wedge \eta \mid \Diamond\psi$$

where p ranges over all propositional variables Φ.

The difference with PropCal is that one can prefix a formula by a diamond. The dual operator $\Box$ ('box') for the diamond is defined by $\Box\phi := \neg\Diamond\neg\phi$. There are various readings for diamond and box which hints towards the power these structures have in modelling knowledge. Examples include

auto-epistemic logic [56] where the basic modal language is used to reason about knowledge itself. In this case $\Box\phi$ is interpreted as meaning 'the agent knows that ϕ'. In *provability logic*, $\Box\phi$ is read as 'it is provable in some arithmetic theory that ϕ'. More than one modality can be used to define a basic modal language as in the case of the Basic Temporal Language. Here the modalities $\langle F\rangle, \langle P\rangle$ are interpreted as follows: $\langle F\rangle\phi$ reads 'ϕ will be true at some point in the **F**uture', while $\langle P\rangle\phi$ reads 'ϕ was true at some point in the **P**ast'. The list continues and detailed application examples are shown in [7]. One interprets the basic modal language using relational structures. Clarifying the notation and terminology used with relational structures, a binary relation R from a set X to a set Y is defined as a subset of the cartesian product $X \times Y$. Binary relations are denoted by the symbols $R, S, T, \ldots$. The notation xRy means that $(x, y) \in R$. In all cases that follow, I consider binary relations for the case when $X = Y$. The image set of x under R is denoted by $R(x) = \{y \mid xRy\}$. In the interpretation of the basic modal language, the semantics is given by Kripke structures or frames.

Definition 2.1.3. A *frame* for the basic modal language is a pair $\mathcal{F} = (W, R)$ such that

(i) W is a non-empty set of possible worlds.

(ii) R is a binary relation over W.

This simple relational structure provides a setting in which to define the models of the basic modal language.

Definition 2.1.4. A *model* for the basic modal language $ML(\Diamond, \Phi)$ is a pair $\mathcal{M} = (\mathcal{F}, V) = (W, R, V)$ where V is a valuation mapping each proposition variable $p \in \Phi$ to a set of worlds $V(p) \subseteq W$. Formally V is a map from the propositional variables to the power set of W i.e. $V : \Phi \longrightarrow \mathcal{P}(W)$.

One uses these definitions to interpret the basic modal language in models, as specified by the following satisfaction definition.

Definition 2.1.5. Consider a model $\mathcal{M} = (W, R, V)$ with $w \in W$ a world in the model. One inductively defines the notion of a formula ϕ being true in $\mathcal{M}$ at world w as follows:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$ with $p \in \Phi$.
- It is not the case that $\mathcal{M}, w \models \bot$.
- $\mathcal{M}, w \models \neg\phi$ iff it is not the case that $\mathcal{M}, w \models \phi$.
- $\mathcal{M}, w \models \phi \vee \psi$ iff $\mathcal{M}, w \models \phi$ or $\mathcal{M}, w \models \psi$.
- $\mathcal{M}, w \models \Diamond\phi$ iff $\exists v \in W$ such that wRv and $\mathcal{M}, v \models \phi$.

The relational structure together with the fact that these notions are intrinsically local, in the sense that each formula is evaluated inside a particular world, provide the powerful description language required to achieve the construction objective. Global truth is naturally defined as truth in every world.

Definition 2.1.6. A formula ϕ is *globally true* in a model $\mathcal{M} = (W, R, V)$ if it is satisfied at every world $w \in W$ in the model $\mathcal{M}$

The following example [7] shows the intuition behind all the definitions thus far:

Example 2.1.1. Consider the frame $\mathcal{F} = (w_1, w_2, w_3, w_4, w_5, R)$ where $w_i R w_j$ iff $j = i + 1$. The model $(\mathcal{F}, V)$ over $\Phi_3 = \{p, q, r\}$ has valuation V defined by:

$V(p) = \{w_2, w_3\}$

$V(q) = W$

$V(r) = \emptyset$

The following is true:

(i) $\mathcal{M}, w_1 \models \Diamond\Box p$
One has $\mathcal{M}, w_3 \models p$ since $w_3 \in V(p)$. Thus $\mathcal{M}, w_2 \models \Box p$ since the image set of w_2 is $R(w_2) = \{w_3\}$. Finally, $\mathcal{M}, w_1 \models \Diamond\Box p$ since $w_1 R w_2$ and $\mathcal{M}, w_2 \models p$ as required.

(ii) $\mathcal{M} \models \Box q$
It is true that $\forall i \mathcal{M}, w_i \models q$ (q is true at all worlds) by definition of the valuation.

The above definitions and simple example shows two important aspects of models in the basic modal language that further reinforce my view that this is the correct mathematical toolkit in which to analyse the thesis objective. Firstly, the example shows that one has a structure in which global propositions (e.g. $\mathcal{M} \models \Box q$ in Example 2.1.1 can be viewed as a law that specifies the physics throughout the universe) and local propositions (e.g. $\mathcal{M}, w_1 \models \Diamond\Box p$ in Example 2.1.1 can be viewed as a belief that determines an individuals actions) could live in harmony with each other. This is exactly a property that I would require of any pregeometry, if I am to adopt Wheeler's hypothesis that pregeometries underly all observables in nature. The observation of someone picking up a stone and throwing it has 2 distinct features - the person's intent to pick up the stone and throw it and the gravitational law that determines the path the stone follows. Secondly, the relation R endows the space with a geometric structure, a requirement for any pregeometry. Further, this geometric structure is intricately linked to the 'local and global laws' of the model, the geometrodynamical view of physics. To motivate this point, I will proceed to show some examples of how globally true formulae in a model determine global geometrical aspects of the underlying relational structure. Towards this end, one defines the following system of axioms for reasoning about frames.

Definition 2.1.7. A *normal modal logic* Λ is a set of formulas that contains the following axiom schemata:

Taut : All propositional tautologies

Dual : $\Diamond p \leftrightarrow \neg\Box\neg p$

K : $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$

and is closed under the inference rules:

Modus ponens: From ϕ and $\phi \rightarrow \psi$, infer ψ

Uniform substitution: From ϕ, infer θ where θ is obtained from ϕ by uniformly replacing propositional variables in ϕ by arbitrary formulae.

Generalisation: From ϕ, infer $\Box\phi$

Throughout this thesis, I will work with the normal modal logic **K**, unless I specifically emphasise that I am working in standard propositional calculus. Inference within **K** will be denoted by $\vdash_K$ while $\vdash$ denotes inference in PropCal. Example 2.1.2 shows how the axiom schemata of models in a normal modal logic specify properties of the relation R of a frame $\mathcal{F} = (W, R)$ and thus from another perspective, the geometry. For the physicist reading this work, Example 2.1.2 shows the well known correspondence proofs [7] linking a property of a relation in a model, to a modal formula that is globally true in the model. Intuitively, the example show how the axioms in a model constrain the 'geometry' of the underlying space.

Example 2.1.2. Consider a model $\mathcal{M} = (W, R, V)$. Then $\forall w \in W, wRw$ (i.e. R is reflexive) iff $\forall\phi \in Form(\Diamond, \Phi)$, $\mathcal{M} \models \Box\phi \rightarrow \phi$
Choose an arbitrary $w \in W$. Now assume $\mathcal{M}, w \models \Box\phi$ for some $\phi \in Form(\Diamond, \Phi)$. Then $\forall v \in W$ such that wRv one has $\mathcal{M}, v \models \phi$. But wRw implying $\mathcal{M}, w \models \phi$. Thus $\mathcal{M}, w \models \Box\phi \rightarrow \phi$. Since this is true for every $w \in W$, one can deduce $\mathcal{M} \models \Box\phi \rightarrow \phi$ as required. Conversely, assume R is not reflexive. Thus $\exists w \in W$ such that not wRw. To falsify $\mathcal{M} \models \Box\phi \rightarrow \phi$ for some $\phi \in Form(\diamond, \Phi)$, consider a valuation V where $V(\phi) = W - \{w\}$. Then, by definition, $\mathcal{M}, w \not\models \phi$. On the other hand, consider v s.t. wRv. This implies $w \neq v$, which in turn implies $v \in V(\phi)$, allowing one to conclude $\mathcal{M}, v \models \phi$. Since v was arbitrary, one has $\mathcal{M}, w \models \Box\phi$ implying $\mathcal{M}, w \not\models \Box\phi \rightarrow \phi$ and thus finally $\mathcal{M} \not\models \Box\phi \rightarrow \phi$.

Armed with these tools, I will proceed to construct a logicatom and the space wherein it lives - a theory network.

2.2 Manufacturing consent

"Saddam Hussein is a threat to America. He's a threat to our friends. He's a man who said he wouldn't have weapons of mass destruction, yet he has them. He's a man that not only has weapons of mass destruction, he's used them. " - G.W. Bush[1]

This section informally guides one towards the formal definition of the logical atom and the structure of the space, theory networks, in which it lives. I will proceed by modelling one of my beliefs. This will in turn motivate the definitions and behavioural requirements of logicatoms. Suppose I believe Chomsky's thesis of 'manufacturing consent', an argument stating how the population of the USA live in an orwellian world order where capital (equated with power) dictates what the masses must know and believe [37]. In order to present this belief to any rational person, I am required to argue why I believe this. So here goes.

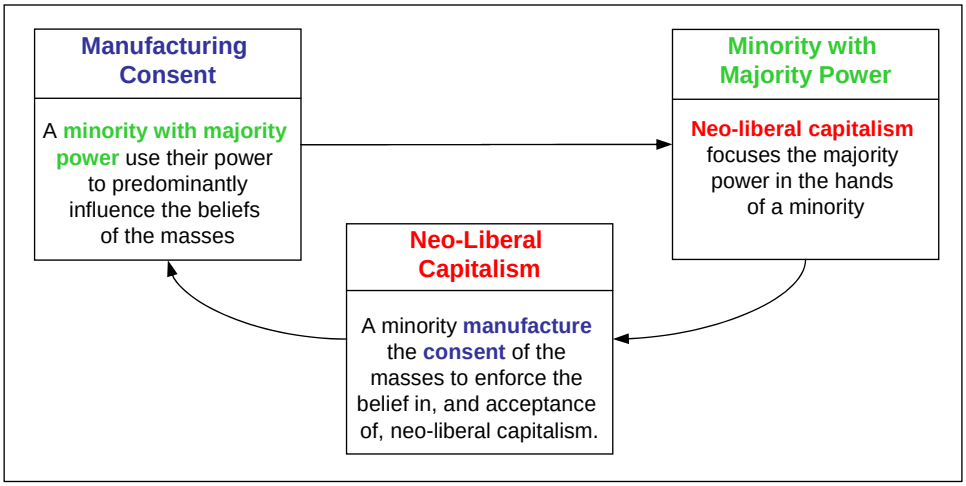

Figure 2.1: The thesis of manufacturing consent

First of all, I believe that the short term profit objective of neo-liberal capitalism, the dominant model adopted by the USA concentrates the majority capital into the hands of a minority (as confirmed by a reading of any fortune 500 magazine). I thus conclude that a minority have the majority power (where I use capital to quantify power). The system is stable in the sense that it is self-perpetuating, since the powerful minority can use their power to manufacture the consent of the masses to accepting neo-liberal capitalism. This allows them to further their own capital interests

[1] During a speech at the Illinois Police Academy in Springfield, Illinois.

and thus accumulate more capital (and thus power by definition). Any alternative structure in society that truly questions this model or the powerful minority interests will find themselves in conflict with those who have the power to regulate what the masses know and believe through media companies owned by them, and governed by the regulating principle encapsulated in the mantra: 'Our short term profit objective dictates that ....'. The most important aspect of the above argument is not to convince the reader to join anti-globalisation groups (although I personally don't think that's a bad idea), but that in trying to explain why I believe in manufacturing consent, I was required to introduce other beliefs of mine e.g. that neo-liberal capitalism implies the minority have the majority capital. My beliefs are built up using existing beliefs. Figure 2.1 shows how I construct my belief of manufacturing consent using my other beliefs. The header of each rectangle represents the names of the various beliefs. The body represents a description in terms of associated beliefs. The arrows point to any beliefs that are mentioned in the description. Note that each individual rectangle does not represent a belief by itself. However, the combination of the three rectangles together with their relationships do. To observe this, start in the manufacturing consent rectangle and move counter-clockwise (in the opposite direction of the arrows), reading each description in turn.

The argument that beliefs comprise other beliefs is analogous to Russell's statement: "Every proposition which we can understand must be composed wholly of constituents with which we are acquainted." The natural question arises as to whether there is a set of quantum beliefs that one can use to build up my belief in manufacturing consent? I adopt Minsky's [54] argument in doubting "the feasibility of representing ordinary knowledge effectively in the form of many small independently true propositions" - formally, the propositional variables of a description logic. In other words, attempting to define the quantum of belief (hereafter referred to as the logicatom) as having the formal representation of propositional variables in some logic will not serve my purposes. Even if one constrains the problem to only modelling knowledge ('fact' as opposed to a 'belief'), I argue that there can be no fundamental building blocks to 'knowledge' since these building blocks would require 'knowledge' to describe them. Consider the statement '*It is raining*'. This statement can only be constructed once the concept '*water falls from the sky*' exists. This in return can only be constructed once the concept of '*water*' and '*sky*' exists. For '*sky*' to exist, I need the concept of '*the view of space from earth*' ... It is my thesis that this process could continue ad infinitum. This argument implies that the logicatom, the building block of knowledge and beliefs is constructed using other logicatoms i.e. knowledge and beliefs. In other words, I require something analogous to a recursive definition. The diagram depicting the belief in manufacturing consent steers me in

defining the logicatom as an entity that has a particular name (e.g. manufacturing consent) and believes something about the other logicatoms in its universe (e.g. 'A minority have the power to influence the beliefs of the masses'). To formalise this, I proceed as follows:

Definition 2.2.1. A *logicatom* μ over PropCal_n is defined as an element of the cross product $\Phi_n \times \mathit{Form}(\Phi_n)$, represented as an ordered pair (p, ϕ) where $p \in \Phi_n$ is a propositional variable and $\phi \in \mathit{Form}(\Phi_n)$ is a formula in PropCal_n.

I refer to the logicatom μ with $\mu := (p, \phi)$ as having name p and believing ϕ. Formally, I define the two projection operators $N : \Phi \times \mathit{Form}(\Phi) \to \Phi$ and $B : \Phi \times \mathit{Form}(\Phi) \to \mathit{Form}(\Phi)$ with $N(\mu) = p$ representing the name and $B(\mu) = \phi$ representing the belief of the logicatom μ. In terms of notation, μ,ν will usually be used to represent logicatoms, $p, q, r, \ldots$ proposition variables and $\phi, \psi, \ldots$ the well formed formulae.

Now consider a set of these logicatoms. Intuitively, the belief of a logicatom is a belief about the logicatoms in the universe in which it resides. For example, assume logicatom μ defined by $\mu := (p_0, p_1 \wedge p_2)$ has name p_0 and believes say that both the logicatoms named p_1 and p_2 are true. This would only make sense if their was only one logicatom named p_1 (respectively p_2) in its respective space. I thus only consider sets $\mathcal{U}$ of logicatoms over PropCal_n that satisfy the condition $\forall \mu, \nu \in \mathcal{U}$, $\mu = \nu$ iff $N(\mu) = N(\nu)$. I have an entity that lives in a space and believes something about the space it lives in, its belief being made up of elements comprising the space.

Example 2.2.1. A pictorial example of the concepts defined using PropCal_3 is shown in Figure 2.2. The names of the logicatoms are shown in the boxes above the ellipses, and their respective beliefs in the underlying ellipses.

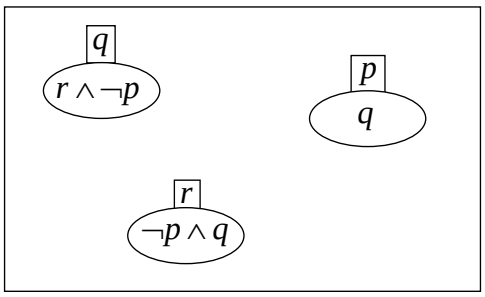

Figure 2.2: A logicatom universe over PropCal_3

A logician may at first sight find it strange that names are atomic beliefs. The previous example of manufacturing consent motivates this definition. If this does not suffice, I ask the reader to bear with

me, as the value thereof will become clear in later sections. I now need to consider the structures required to model the dynamical behaviour of logicatoms. Intuitively, this will be encapsulated in the belief revision of each logicatom. Now the natural mechanism (from a physics perspective) would be to allow logicatoms to interact and through their interaction, change their state. This will result in a new (and maybe different) logicatom universe. In order for a logicatom to interact, it needs to know of the existence of other logicatoms. This can be represented as a binary relation R on $\mathcal{U}$. Now guided by modern theories of physics that are predominantly local in their description of nature, I impose a constraint on the binary relation R, in that it must be local. I now constrain the definition of a logicatom universe to be a pair $(\mathcal{U}, R)$ where $\mathcal{U}$ comprises a set of uniquely named logicatoms over PropCal$_n$ and $R \subseteq \mathcal{U} \times \mathcal{U}$ is a local binary relation. Before I proceed to formalise the definition of the local relation, I present an example showing how such a relation can be defined.

Example 2.2.2. Given logicatoms $\mu, \nu \in \mathcal{U}$, define the local relation of $\mathcal{U}$ to be a binary relation $R \subseteq \mathcal{U} \times \mathcal{U}$ given by

$$\mu R \nu \text{ iff } \vdash B(\mu) \rightarrow N(\nu)$$

The inference $\vdash$ is the standard inference of PropCal. This definition says that two logicatoms μ, ν are related if and only if the belief of μ implies the name of ν in PropCal$_n$. This relation is clearly local since one can determine all the inferences of the propositional variables for any proposition ϕ in PropCal$_n$. This relation is shown for the example of logicatoms in Figure 2.3 below. It will be referred to as the *affirmed implication relation*.

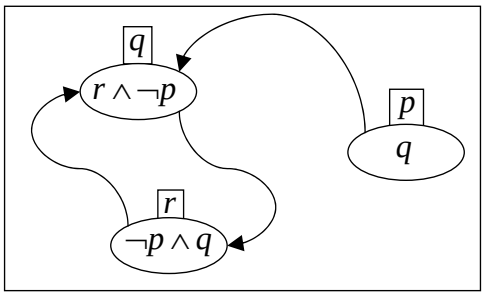

Figure 2.3: A universe comprising related logicatoms

One can formalise the definition of the local relation of a logicatom universe using the definition of a *relation generating function*.

Definition 2.2.2. A *relation generating function* over PropCal$_n$ is defined as a function $R_f : Form(\Phi_n) \rightarrow \mathcal{P}(\Phi_n)$ that maps formulae to sets of proposition variables.

Given logicatoms μ, ν over Φ_n and a relation generating function $R_f : \mathit{Form}(\Phi_n) \to \mathcal{P}(\Phi_n)$, one can define the binary relation by

$$\mu R \nu \text{ iff } N(\nu) \in R_f[B(\mu)]$$

Relation generating functions capture the essence of the *local* requirement, since what μ is related to is determined solely by its belief $B(\mu)$ using a well defined function R_f.

Definition 2.2.3. A *logicatom universe* over $\mathrm{PropCal}_n$ is defined as the pair $(\mathcal{U}, R_f)$ where $\mathcal{U}$ is a set of n distinct logicatoms satisfying the unique naming constraint i.e. for every logicatom $\mu, \nu \in \mathcal{U}$, $\mu = \nu$ iff $N(\mu) = N(\nu)$, and $R_f : \mathit{Form}(\Phi_n) \to \mathcal{P}(\Phi_n)$ is a relation generating function that generates the binary relation $R \subset \mathcal{U} \times \mathcal{U}$ capturing local relationships between logicatoms.

The final requirement for completion of this platform, is 'change'. This is clear since our objective will be to show that the way a logicatom universe changes is regulated by natural selection. Once again, motivated by physics, I require this change to be local in nature. Now given a logicatom $\mu \in \mathcal{U}$ in a logicatom universe $(\mathcal{U}, R_f)$, consider the image set of μ under R, the binary relation generated by the relation generating function R_f. One has

$$\begin{aligned} R(\mu) &= \{\nu \in \mathcal{U} \mid \mu R \nu\} \\ &= \{\nu \in \mathcal{U} \mid N(\nu) \in R_f[B(\mu)]\} \end{aligned}$$

The image set $R(\mu)$ specifies the 'local' set of logicatoms that could possibly influence how μ changes. Now the name or the belief of μ could possibly change. If one was to allow the name of μ to change, one could in no way guarantee that the new set of logicatoms generated will satisfy the unique naming constraint. For the required dynamic behaviour, I will thus only allow the beliefs of logicatoms to (possibly) change. This will be encapsulated in a transition rule that maps a logicatom universe $(\mathcal{U}, R_f)$ to a new logicatom universe $(\mathcal{U}', R_f)$ by changing the beliefs of the logicatoms in $\mathcal{U}$. The new belief of a logicatom will be determined solely by information local to the logicatom i.e. the set comprising the logicatom and its related logicatoms. (Note that the local criterion specified is exactly the same as for transition functions of cellular automata [2].) The relation generating map will generate a new binary relation $R' \subset U' \times U'$ from the new set of beliefs.

[2] See Definition 2.5.2

Example 2.2.3. An example of a transition rule $T : Form(\Phi) \rightarrow Form(\Phi)$ that updates the beliefs of logicatoms in a universe $\mathcal{U}$ is defined as follows: Assume $\mu R \nu_i$ for some set $\{\nu_1, \nu_2, \ldots, \nu_k\} \subseteq \mathcal{U}$. T specifies that the belief of μ gets updated by uniformly substituting every occurrence of $N(\nu_i)$ in $B(\mu)$ with $B(\nu_i)$ and evaluating the new proposition. This update rule is shown in Figure 2.4 where I have used the *affirmed implication* relation as specified in Example 2.2.2. Note that the transition rule updates the beliefs of the logicatoms and consequently, the relations between them, resulting in a new logicatom universe $\mathcal{U}'$.

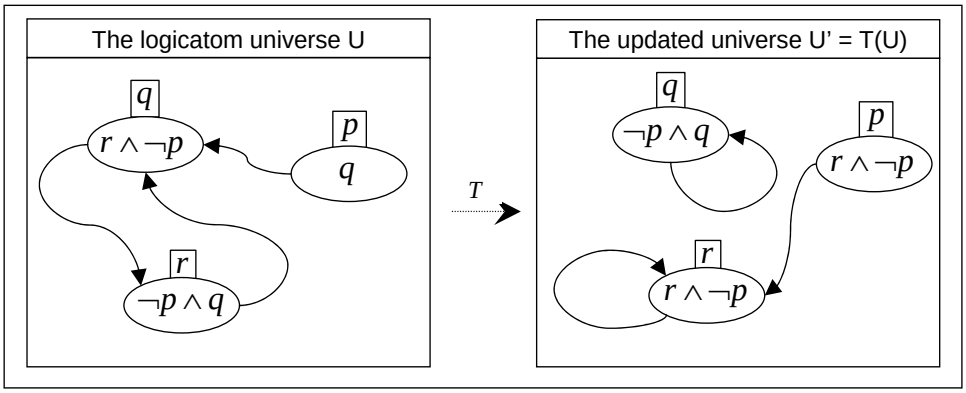

Figure 2.4: The dynamics of a logicatoms universe

A space comprising logicatoms has been defined. These logicatoms 'believe' something about the universe they live in. This belief determines who they 'see' in their universe. This in turn determines how their belief will change. I thus have all the tools to define a theory network over PropCal$_n$.

Definition 2.2.4. A theory network over PropCal$_n$ comprises a 3-tuple $(\mathcal{U}, R_f, T)$ where (U, R_f) is a logicatom universe and $T : Form(\Phi_n) \rightarrow Form(\Phi_n)$ is a local transition function that maps a logicatom's belief to a new belief. The updated belief of the logicatom is determined solely by its current belief and its related set of logicatoms' beliefs.

The 3-tuple allows me to generate a series of logicatom universes $((U_0, R_f), (U_1, R_f), (U_2, R_f), \ldots)$. A logicatom universe (U_i, R_f) is determined by the universe (U_{i-1}, R_f) and the transition rule T. The local binary relation is generated using the relation generating function R_f, and can change from one universe to the next.

I will conclude this section by intuitively arguing how this simple structure satisfies the behavioural characteristics I am after in terms of the construction objective. From a biological perspective, consider Dawkins' behavioural analogy of memes (or beliefs) with genes. Now the phenotype of the gene determines its morphology and physiological control i.e. its expression within the natural environment. In our terminology, the morphology of a logicatom is encapsulated in the local binary

relation R generated by the relation generating function. The morphology thus specifies its relation to other logicatoms. By the analogy with genes this 'phenotype' should be determined by the logicatom itself i.e. locally. Section 3.5.2 will expand on this analogy and incorporate all other aspects required in the living sciences i.e. procreation, parenthood, transmission and mutation of genes etc. From a physics perspective, Klinger and Cahill define the entities comprising their pregeometric space as "information denoting relationships" [10]. My definitions and requirements analogously stipulate that the information content of a logicatom specifies its relationships with other logicatoms, reinforcing the path I am following to create a structure that can serve to model a pregeometry in physics. Finally, from a knowledge modelling perspective, all one can conclude at the current time is that I have a structure built up using the language of logic, a platform for modelling knowledge. Towards this end, I will delve deeper into modal logic, with the objective of creating an alternative representation of theory networks and logicatom universes, that will be shown to be very useful for reasoning about the content and dynamics of this space.

2.3 Alternative representations of theory networks

I introduce this section by stating the standard definitions and theorems [7] that encapsulate the invariance results used for models in a basic modal language. These definitions and theorems will be used to arrive. at an alternative representation for logicatom universes and theory networks. This alternative representation will allow me to extend the current definitions, allowing me to show how knowledge is modelled within this platform.

Definition 2.3.1. Let $\mathcal{M}_1 = (W_1, R_1, V_1)$ and $\mathcal{M}_2 = (W_2, R_2, V_2)$ be two models of the basic modal language $ML(\Diamond, \Phi)$.
$\mathcal{M}_2$ is a *submodel* of $\mathcal{M}_1$ if $W_2 \subseteq W_1$, $R_2 = R_1 \cap (W_2 \times W_2)$ (i.e. R_2 is the restriction of R_1 to W_2) and $\forall p \in \Phi$, $V_2(p) = V_1(p) \cap W_2$.
$\mathcal{M}_2$ is a *generated submodel* of $\mathcal{M}_1$ (i.e. $\mathcal{M}_2 \rightarrowtail \mathcal{M}_1$) if $\mathcal{M}_2$ is a submodel of $\mathcal{M}_1$ and the following closure condition holds: If $w \in W_2$ and wR_1v then $v \in W_2$.

Theorem 2.3.1. *Let $\mathcal{M}_1 = (W_1, R_1, V_1)$ and $\mathcal{M}_2 = (W_2, R_2, V_2)$ be two models in the basic modal language such that $\mathcal{M}_2$ is a generated submodel of $\mathcal{M}_1$. Then for every modal formula $\phi \in Form(\Diamond, \Phi)$ and every world $w \in W_2$ one has*

$$\mathcal{M}_1, w \models \phi \text{ iff } \mathcal{M}_2, w \models \phi.$$

Definition 2.3.2. Let $\mathcal{M}_1 = (W_1, R_1, V_1)$ and $\mathcal{M}_2 = (W_2, R_2, V_2)$ be two modal models. A *strong homomorphism* f from $\mathcal{M}_1$ to $\mathcal{M}_2$ written $f : \mathcal{M}_1 \to \mathcal{M}_2$ is a function between W_1 and W_2 with the following properties:

(i) w and $f(w)$ satisfy the same propositional variables;

(ii) $\forall w, v \in W_1$, wR_1v iff $f(w)R_2f(v)$.

An *isomorphism* is a bijective strong homomorphism.

Definition 2.3.3. Let $\mathcal{M}_1 = (W_1, R_1, V_1)$ and $\mathcal{M}_2 = (W_2, R_2, V_2)$ be two modal models. A mapping f from $\mathcal{M}_1$ to $\mathcal{M}_2$ is a *bounded morphism* if it satisfies the following conditions:

(i) w and $f(w)$ satisfy the same propositional variables;

(ii) f is a homomorphism with respect to the relations i.e. if wR_1v then $f(w)R_2f(v)$;

(iii) If $f(w)R_2v_2$, then $\exists v_1$ such that wRv_1 and $f(v_1) = v_2$.

If there is a surjective bounded morphism from $\mathcal{M}_1$ to $\mathcal{M}_2$, then $\mathcal{M}_2$ is a *bounded morphic image* of $\mathcal{M}_1$.

Definition 2.3.4. Let $\mathcal{M}_1 = (W_1, R_1, V_1)$ and $\mathcal{M}_2 = (W_2, R_2, V_2)$ be two modal models. A non-empty relation $Z \subseteq W_1 \times W_2$ is called a *modal bisimulation* if the following hold:

(i) If w_1Zw_2 then w_1 and w_2 satisfy the same propositional variables;

(ii) If w_1Zw_2 and $w_1R_1v_1$, then there exists a $v_2 \in W_2$ such that $w_2R_2v_2$ and v_1Zv_2 (the *forth* condition);

(iii) If w_1Zw_2 and $w_2R_2v_2$, then there exists a $v_1 \in W_1$ such that $w_1R_1v_1$ and v_1Zv_2 (the *back* condition).

Throughout the thesis, I will be working with the notion of uniformly replacing propositional variables in a formula with other formulae.

Definition 2.3.5. Consider the basic modal language $Form(\Diamond, \Phi_n)$ over a finite set Φ_n of n propositional variables. A *substitution* is a map $\xi : \Phi_n \to \mathrm{Form}(\Diamond, \Phi_n)$.
The substitution induces the map $(\cdot)^\xi : \mathrm{Form}(\Diamond, \Phi) \to \mathrm{Form}(\Diamond, \Phi)$ known as *uniform substitution* which is recursively defined as follows:

$$\begin{aligned} \bot^\xi &= \bot \\ p^\xi &= \xi(p) \\ (\neg\phi)^\xi &= \neg(\phi^\xi) \\ (\phi \vee \theta)^\xi &= \phi^\xi \vee \theta^\xi \\ (\Diamond\phi)^\xi &= \Diamond(\phi^\xi) \end{aligned}$$

Given a substitution ξ and a formula ϕ, one can denote the substitution instance ψ of ϕ by $\psi := \phi^\xi$. An alternative notation for uniform substitution also used in this thesis is given by:

$$\phi^\xi \equiv \phi\left(\xi(p_1)/p_1, \ldots, \xi(p_n)/p_n\right).$$

I extend this concept to that of *constrained uniform substitution.*

Definition 2.3.6. Let ξ be a substitution map over the basic modal language $Form(\Diamond, \Phi_n)$, and let $X \subseteq \Phi_n$ be a subset of propositional variables. The substitution ξ restricted to X, called the *constrained uniform substitution*, is denoted by $\xi|_X$ and is defined by

$$(\xi|_X)(p) = \begin{cases} \xi(p) & \text{if } p \in X \\ p & \text{otherwise} \end{cases}$$

Example 2.3.1. Consider the substitution map ξ, defined by

$$\begin{aligned} \xi(p) &= q \wedge \Diamond r \\ \xi(q) &= p \\ \xi(r) &= q \vee \Box(p \wedge r) \end{aligned}$$

Let $X = \{q, r\}$. If $\phi = p \wedge q \wedge r$ then

$$\begin{aligned} \phi^{\xi} &= (p \wedge q \wedge r)^{\xi} = (q \wedge \Diamond r) \wedge p \wedge (q \vee \Box(p \wedge r)) \\ \phi^{\xi|_X} &= (p \wedge q \wedge r)^{\xi|_X} = p \wedge (q \vee \Box(p \wedge r)) \end{aligned}$$

I will now proceed to show that a logicatom universe $(\mathcal{U}, R_f)$ over PropCal_n can be uniquely represented by the pair (ξ, R_f) where $\xi : \Phi_n \to Form(\Phi_n)$ is a substitution map as introduced in Definition 2.3.5 and $R_f : Form(\Phi_n) \to \mathcal{P}(\Phi_n)$ is a relation generating function over PropCal_n as defined in Definition 2.2.2. This representation will enable me to analyse the various sought after characteristics of logicatom universes, as described in the thesis objective. In order to give the reader an intuitive feel of how I arrive at this alternative definition, I will first represent the logicatom universe of Example 2.2.2 using a model of the basic modal language.

Example 2.3.2. Define the model $\mathcal{M} = (W, R, V)$ with $W = \{\mu_1, \mu_2, \mu_3\}$ and $R = \{(\mu_1, \mu_2), (\mu_2, \mu_3), (\mu_3, \mu_2)\}$. The valuation V is specified by

$$\begin{aligned} V(p) &= \{\mu_1\} \\ V(q) &= \{\mu_2\} \\ V(r) &= \{\mu_3\} \end{aligned}$$

The model $\mathcal{M}$ completely determines the names (through the valuation mapping) and relationships (equivalent to the modal relationship) of the logicatom universe illustrated in Figure 2.3 of Example 2.2.2. To complete the specification, I require a method of representing the beliefs of each logicatom. This is accomplished by defining the substitution map $\xi : \Phi_3 \to Form(\Phi_3)$ with:

$$\begin{aligned} \xi(p) &= q \\ \xi(q) &= r \wedge \neg p \\ \xi(r) &= \neg p \wedge q \end{aligned}$$

By specifying the belief of a logicatom $\mu \in W$ as being $\xi\,(N(\mu))$, the logicatom universe defined in Example 2.2.2 has been completely specified. To summarise, the set W together with the valuation V determines the number of logicatoms and their corresponding names. The accessibility relation R identifies the relationships between the logicatoms, and the substitution ξ in turn determines the belief of each logicatom. Finally note that the proposition

$$\begin{aligned}\phi \;&:=\; ((p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge q \wedge \neg r) \vee (\neg p \wedge \neg q \wedge r))\\ &\wedge\quad (p \wedge \Diamond q)\\ &\wedge\quad (q \wedge \Diamond r)\\ &\wedge\quad (r \wedge \Diamond q)\end{aligned}$$

is globally true in $\mathcal{M}$ i.e. $\mathcal{M} \models \phi$.

The above example shows how I represent the information of a logicatom universe using a model of the basic modal language to represent the geometry (a co-ordinate system embodied in the name of each logicatom and a metric embodied in the relations between the logicatoms) and a substitution map to represent the beliefs. I will define *G-Models* as the class of such models that embody this geometric information. I will show that these G-models, together with a substitution map and a relation generating function, represent a unique logicatom universe. Further, I will show that these G-Models are in fact completely determined by the substitution map and relation generating function using the concept of a *G-defining* proposition. Formally, I define G-Models as:

Definition 2.3.7. A model $\mathcal{M} = (W, R, V)$ of the basic modal language over a finite set of propositional variables Φ_n is called a G-model iff the following conditions are satisfied

(i) $\forall p \in \Phi_n, \exists w \in W$ such that $w \in V(p)$ and $\forall w \in \mathcal{W}, \exists p \in \Phi$ such that $w \in V(p)$.

(ii) For $p, q \in \Phi_n$ with $p \neq q$, one has $V(p) \cap V(q) = \emptyset$.

(iii) For $p \in \Phi_n$ with $u, v \in V(p)$ and $w \in W$, if uRw then vRw.

Intuitively, condition (i) of the definition above states that every propositional variable is true in at least one world and every world satisfies at least one propositional variable. Condition (ii) states that every world satisfies at most one propositional variable. Finally condition (iii) implies that two worlds satisfying the same propositional variable, are modally equivalent. (This will become evident in the proofs that follow). To show how G-models are used in the representation of logicatom universes, I first prove that every G-model over Φ_n is bisimilar to a unique (up to isomorphism) G-model with exactly n worlds. In terms of the notation used, a bisimulation Z between G-models $\mathcal{M} = (W, R, V)$ and $\mathcal{M}' = (W', R', V')$ is *total* iff $\forall w \in W, \exists w' \in W'$ such that wZw'. The range of Z is denoted by $ran(Z) = \{w' \in W' \mid \exists w \in W \text{ st } wRw'\}$. The domain of Z denoted by $dom(Z) = \{w \in W \mid \exists w' \in W' \text{ st } wRw'\}$. The converse of a relation Z is the relation $Z^{\smile} = \{(w', w) \mid wZw'\}$.

Theorem 2.3.2. *Let $\mathcal{M} = (W, R, V)$ be a G-model over Φ_n. Then there exists a unique bisimilar G-model $\mathcal{M}^G = (W^G, R^G, V^G)$ with $\#W^G = n$*

Proof. Since $\mathcal{M}$ is a G-model, one has $\#W \geq n$. (This is enforced through conditions (i) and (ii) in Definition 2.3.7.) The case $\#W = n$ is trivially true. For the case $\#W > n$, define the G-model $\mathcal{M}^G = (W^G, R^G, V^G)$ as follows: Let $W^G = \{V(p) \mid p \in \Phi_n\}$. Define V^G by $V^G(p) = \{V(p)\}$ for every $p \in \Phi_n$. Finally, define R^G as follows: For every $p, q \in \Phi_n$, $V(p)R^GV(q)$ iff $\exists w \in V(p)$,$\exists v \in V(q)$ such that wRv. By construction, there are exactly n worlds with each world satisfying a unique proposition variable in Φ_n. All requirements of Definition 2.3.7 are therefore trivially satisfied. Define the bisimulation $Z \subset W \times W^G$ as follows: Select $w \in W$. Since $\mathcal{M}$ is a G-model, there exists a unique $p \in \Phi_n$ such that $w \in V(p)$. Then $wZV(p)$ with $w \in W$ and $V(p) \in W^G$. To show that Z is a bisimulation, I need to prove the three conditions specified in Definition 2.3.4. For requirement (i), select $w \in W$ and $V(p) \in W^G$ such that $wZV(p)$. By construction $w \in V(p)$ and $V(p) \in V^G(p)$. Since $\mathcal{M}$ is a G-model, w only satisfies the propositional variable p. (Conditions (i) and (ii) of Definition 2.3.7.) By construction $V(p)$ only satisfies the propositional variable p. Thus both w and $V(p)$ satisfy the same proposition variables. To show the forth condition, assume $wZV(p)$ and wRv for $w, v \in W$ and $V(p) \in W^G$. Select the unique $q \in \Phi_n$ such that $v \in V(q)$. (This is possible since $\mathcal{M}$ is a G-model). Now $vZV(q)$ by definition of Z and $V(p)R^GV(q)$ by definition of R^G, confirming the forth condition. To show the back condition, assume $wZV(p)$ and $V(p)R^GV(q)$ for $V(p), V(q) \in W^G$ and $w \in W$. By definition of R^G, there exists $w', v \in W$ with $w' \in V(p)$, $v \in V(q)$ and $w'Rv$. But by condition (iii) of Definition 2.3.7, one has $w, w' \in V(p)$ with $w'Rv$ implying wRv, as required.

To show uniqueness, let $\mathcal{M}_2^G = (W_2^G, R_2^G, V_2^G)$ be a bisimilar G-model over Φ_n with $\#W_2^G = n$ and bisimulation Z_2 such that Z_2 and $Z_2^\smile$ are total. In order to show that $\mathcal{M}^G$ is isomorphic to $\mathcal{M}_2^G$, define the function $f : W^G \to W_2^G$ as follows: Given $V(p) \in W^G$, select the unique $w_2 \in W_2^G$ that satisfies the propositional variable p. Existence of w_2 is guaranteed by the definition of G-models. Uniqueness is guaranteed due to the constraint that $W_2^G = n$. Then $f(V(p)) = w_2$. Since $\#W^G = n$, f is a bijection. By construction of f, condition (i) of Definition 2.3.2 is satisfied. To prove condition (ii), select $V(p), V(q) \in W^G$ such that $V(p)R^GV(q)$. By construction of $\mathcal{M}^G$, $\exists w \in V(p)$,$\exists v \in V(q)$ such that wRv. Now since Z_2 and $Z_2^\smile$ are total, $\exists w_2 \in W_2^G$ with wZ_2w_2. By the forth condition of Definition 2.3.4, $\exists v_2 \in W_2^G$ such that $w_2R_2^Gv_2$ and vZ_2v_2. Thus v and v_2 satisfy the same proposition variables, implying that $v_2 \in V_2^G(q)$. By construction of f, one has that $f(V(p))R_2^Gf(V(q))$ as required. The converse requirement that $f(V(p))R_2^Gf(V(q))$ implies $V(p)R^GV(q)$ is argued in the identical way using the total bisimulation $Z_2^\smile$. □

The above theorem shows that I can associate a unique G-model containing exactly n-worlds with any G-model over Φ_n. This unique G-model will, by construction, allow me to specify the amount of, and names of the logicatoms in a logicatom universe. As shown in Example 2.3.2, a substitution map allows one to specify the beliefs of each logicatom. Finally, a relation generating function will specify the relations between the logicatoms. This will naturally be required to induce the G-model accessibility relation. These requirements provide me with an alternative definition of a logicatom universe:

Definition 2.3.8. A logicatom universe over PropCal_n comprises the tuple $(\mathcal{M}^G, \xi, R_f)$ where $\mathcal{M}^G = (W^G, R^G, V^G)$ is a G-model, $\xi : \Phi_n \to \mathit{Form}(\Phi_n)$ is a substitution map and $R_f : \mathit{Form}(\Phi_n) \to \mathcal{P}(\Phi_n)$ a relation generating function satisfying the following property: If $p, q \in \Phi_n$ is such that $q \in R_f[\xi(p)]$, then $\forall w \in V(p)$,$\forall v \in V(q)$ one has that wR^Gv.

The constraint that needs to be satisfied in Definition 2.3.8 ensures that the relation generated by the relation generating function in the logicatom universe and the relation in the G-model are equivalent in terms of the mapping from the modal model to the logicatom universe. The rest of this subsection will focus on showing that the G-model itself is completely defined by the substitution map and relation generating function, allowing me to conclude this section by arriving at the final definition of a logicatom universe as a pair comprising a substitution map and a relation generating function only. Towards this end, I define the concept of G-Axioms and G-defining propositions.

Definition 2.3.9. Consider a model of the basic modal language $\mathcal{M} = (W, R, V)$ over $\Phi_n = \{p_1, \ldots, p_n\}$. Let $\Lambda = \{\phi \mid \mathcal{M} \models \phi\}$ be the set of globally true formulae in $\mathcal{M}$. I say the model $\mathcal{M}$ satisfies *G-Axioms* iff the following conditions hold:

Taut	Λ contains all propositional tautologies in $\mathit{Form}(\Phi_n)$
K	$(\Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)) \in \Lambda$
G1	$\left(\bigvee_{i=1}^{n} p_i\right) \in \Lambda$
G2	$\{p_i \to \neg p_j \mid 1 \leq i, j \leq n \text{ and } i \neq j\} \subset \Lambda$
G3	$\{\neg p_i \mid 1 \leq i \leq n\} \cap \Lambda = \emptyset$

Further, I specify the G-defining proposition $\phi^G_{\mathcal{M}}$ of the model $\mathcal{M}$ that satisfies the G-axioms as

$$\begin{aligned} \phi^G_{\mathcal{M}} := \left(\bigvee_{i=1}^{n} p_i\right) \quad &\wedge \quad \left(\bigwedge_{i,j=1;i\neq j}^{n} (p_i \to \neg p_j)\right) \\ &\wedge \quad \left(\bigwedge \{\psi \mid \psi \in \Lambda \cap \{p_i \to \Diamond p_j \mid 1 \leq i, j \leq n\}\}\right) \end{aligned} \tag{2.3.1}$$

Definition 2.3.9 specifies the class of models that I am interested in. The requirements **Taut** and **K** just mean that the models under consideration are a subset of the normal modal logic **K** (See Section 2.1.7). It will be shown that **G1**, **G2** and **G3** specify the unique naming constraint as well as constraining the minimum number of worlds in the model to be n. I will show that the G-defining proposition $\phi^G_{\mathcal{M}}$ of the model represents all the information required to specify the geometry of a logicatom universe i.e. the G-model. The next set of theorems will show one how the G-defining proposition and a G-model are related, providing me with the tools to create the final equivalent definition for logicatom universes.

Lemma 2.3.3. *Consider a model $\mathcal{M} = (W, R, V)$ in the basic modal language over Φ_n with $n \geq 2$. If $\mathcal{M}$ satisfies G-axioms, then $\#W \geq n$.*

Proof. Choose $w \in W$. From **G1** and **G2** one has

$$\mathcal{M}, w \models \bigvee_{i=1}^{n} p_i \tag{2.3.2}$$

$$\forall i, j \in \{1, \ldots, n\} \qquad \mathcal{M}, w \models p_i \rightarrow \neg p_j \tag{2.3.3}$$

Formula 2.3.2 implies that at least 1 propositional variable is true at w. Formula 2.3.3 implies that at most 1 propositional variable is true at w. For assume $\mathcal{M}, w \models p_i$ and $\mathcal{M}, w \models p_j$ for some $i \neq j$. Now $\mathcal{M}, w \models p_i$, Formula 2.3.3 and modus ponens imply $\mathcal{M}, w \models \neg p_k$ for every $k \neq i$, specifically for $k = j$ contradicting $\mathcal{M}, w \models p_j$.
Now assume $\#W < n$. Since only one propositional variable is true at any world with the rest false, and since $n \geq 2$, there exists $p_m \in \Phi_n$ such that $\mathcal{M}, w \models \neg p_m$ for any w. Thus $\mathcal{M} \models \neg p_m$ contradicting **G3**. □

I have shown that the G-axioms provide one with the minimum amount of worlds required in a G-model. I will now show that they specify the exact number (i.e. n worlds for a G-model over Φ_n) if one considers all bisimilar models and selects a model with the smallest number of worlds.

Theorem 2.3.4. *Consider two models $\mathcal{M} = (W, R, V)$, $\mathcal{M}' = (W', R', V')$ over Φ_n that satisfy G-axioms. Define $\Lambda = \{\phi \mid \mathcal{M} \models \phi\}$ and $\Lambda' = \{\phi \mid \mathcal{M}' \models \phi\}$. If*

$$\Lambda \cap \{p_i \rightarrow \Diamond p_j \mid 1 \leq i, j \leq n\} = \Lambda' \cap \{p_i \rightarrow \Diamond p_j \mid 1 \leq i, j \leq n\} \tag{2.3.4}$$

then M' is bisimilar to M. Further, the bisimulation Z and its converse $Z^\smile$ are total.

Proof. Define the relation $Z \subset W \times W'$ as follows: $\forall w \in W, w' \in W'$ wZw' iff w and w' satisfy the same proposition variables. In order to prove that this is a bisimulation, I need to show that Z is non-empty and satisfies the three conditions specified in Definition 2.3.4. Now from Lemma 2.3.3, I have that Z is non-empty since both models contain at least n worlds and each model has at least one world satisfying $\left(\bigwedge_{i=1, i\neq j}^{n} \neg p_i\right) \wedge p_j$ for every $j \leq n$. Thus every world in W is related to at least one world in W', and vice versa, showing that Z is non-empty, and further that Z and $Z^\smile$ are total. By construction, Z satisfies condition (i) of Definition 2.3.4.
To prove the forth condition, assume wZw' and wRv. Select the unique $p_w, p_v \in \Phi_n$ such that $w \in V(p_w)$ and $v \in V(p_v)$ respectively. This is possible since M, M' satisfy the G-axioms (thus every world satisfies one and only one propositional variable). Now $\mathcal{M}, v \models p_v$ and wRv implies $\mathcal{M}, w \models \Diamond p_v$. Since $\mathcal{M}, w \models p_w$, one can conclude that $\mathcal{M}, w \models p_w \rightarrow \Diamond p_v$. For any other $u \in W$, one has $\mathcal{M}, u \models p_w \rightarrow \Diamond p_v$ since $\mathcal{M}, u \models \neg p_w$ for $u \neq w$. Thus $\mathcal{M} \models p_w \rightarrow \Diamond p_v$ Using 2.3.4. one can conclude that

$$\mathcal{M}' \models p_w \rightarrow \Diamond p_v \tag{2.3.5}$$

Since wZw' implies w and w' satisfy the same proposition variables, one has $\mathcal{M}', w' \models p_w$, which together with 2.3.5 and modus ponens implies that $\mathcal{M}', w' \models \Diamond p_v$. Thus $\exists v' \in W'$ such that $w'R'v'$ and $\mathcal{M}', v' \models p_v$. This implies that v and v' satisfy the same proposition variables allowing one to

conclude vZv' and thus the forth condition, as required.
Proving the back condition is a symmetrical argument. Assume wZw' and $w'R'v'$. Select the unique $p_{w'}, p_{v'} \in \Phi_n$ such that $w' \in V(p_{w'})$ and $v' \in V(p_{v'})$ respectively. Now $\mathcal{M}', v' \models p_{v'}$ and $w'R'v'$ implies $\mathcal{M}', w' \models \Diamond p_{v'}$. Since $\mathcal{M}', w' \models p_{w'}$, one can conclude that $\mathcal{M}', w' \models p_{w'} \to \Diamond p_{v'}$. For any other $u' \in W'$, one has $\mathcal{M}', u' \models p_{w'} \to \Diamond p_{v'}$ since $\mathcal{M}', u' \models \neg p_{w'}$ for $u' \neq w'$. Thus $\mathcal{M}' \models p_{w'} \to \Diamond p_{v'}$ Using 2.3.4 one can conclude that

$$\mathcal{M} \models p_{w'} \to \Diamond p_{v'} \tag{2.3.6}$$

Since wZw' implies w and w' satisfy the same propositional variables, one has $\mathcal{M}, w \models p_{w'}$, which together with 2.3.6 and modus ponens implies that $\mathcal{M}, w \models \Diamond p_{v'}$. Thus $\exists v \in W$ such that wRv and $\mathcal{M}, v \models p_{v'}$. This implies that v and v' satisfy the same proposition variables allowing one to conclude vZv' and thus the back condition as required. □

Lemma 2.3.5. *Consider two models $\mathcal{M} = (W, R, V)$, $\mathcal{M}' = (W', R', V')$ over Φ_n that satisfy G-axioms. Define $\Lambda = \{\phi \mid \mathcal{M} \models \phi\}$ and $\Lambda' = \{\phi \mid \mathcal{M}' \models \phi\}$. Further, let $\phi^G_{\mathcal{M}}$, $\phi^G_{\mathcal{M}'}$ be the G-defining propositions for $\mathcal{M}$ and $\mathcal{M}'$ respectively. Then*

$$\begin{aligned} & \vdash_K \phi^G_{\mathcal{M}} \leftrightarrow \phi^G_{\mathcal{M}'} \\ \textit{iff} \quad & \Lambda \cap \{p_i \to \Diamond p_j \mid 1 \leq i, j \leq n\} = \Lambda' \cap \{p_i \to \Diamond p_j \mid 1 \leq i, j \leq n\} \end{aligned}$$

Proof. Assume $\Lambda \cap \{p_i \to \Diamond p_j \mid 1 \leq i, j \leq n\} = \Lambda' \cap \{p_i \to \Diamond p_j \mid 1 \leq i, j \leq n\}$. Then by 2.3.1, one has $\vdash_K \phi^G_{\mathcal{M}} \leftrightarrow \phi^G_{\mathcal{M}'}$. Conversely assume $\Lambda \cap \{p_i \to \Diamond p_j \mid 1 \leq i, j \leq n\} \neq \Lambda' \cap \{p_i \to \Diamond p_j \mid 1 \leq i, j \leq n\}$ and $\vdash_K \phi^G_{\mathcal{M}} \leftrightarrow \phi^G_{\mathcal{M}'}$. Without loss of generality, one can assume that $\exists i, j$ with $1 \leq i, j \leq n$ such that $(p_i \to \Diamond p_j) \in \Lambda$ and $(p_i \to \Diamond p_j) \notin \Lambda'$. Thus

$$\begin{aligned} & \vdash_K \phi^G_{\mathcal{M}} \leftrightarrow \phi^G_{\mathcal{M}'} \\ \Rightarrow \quad & \vdash_K \left(\phi^G_{\mathcal{M}} \wedge (p_i \to \Diamond p_j)\right) \leftrightarrow \left(\phi^G_{\mathcal{M}'} \wedge (p_i \to \Diamond p_j)\right) \\ \Rightarrow \quad & \vdash_K (p_i \to \Diamond p_j)) \leftrightarrow \bot \end{aligned}$$

resulting in the contradiction sought after. □

Theorem 2.3.4 together with Lemma 2.3.5 shows that the G-defining proposition of a model that satisfies G-axioms, determines the model up to bisimulation. The next set of theorems steer one towards the isomorphism required to generate a unique G-model.

Theorem 2.3.6. *Let $\mathcal{M} = (W, R, V)$, $\mathcal{M}' = (W', R', V')$ be models over Φ_n that satisfy G-axioms. If the models have the same G-defining proposition (i.e. $\vdash_K \phi^G_{\mathcal{M}} \leftrightarrow \phi^G_{\mathcal{M}'}$) and $\#W = \#W' = n$, then $\mathcal{M}$ is isomorphic to $\mathcal{M}'$.*

Proof. Since $\mathcal{M}, \mathcal{M}'$ are models over Φ_n with equivalent G-defining propositions, Theorem 2.3.4 and Lemma 2.3.5 imply that they are bisimilar with bisimulation Z. Define the map $f : \mathcal{M} \to \mathcal{M}'$ by $f(w) = w'$ iff wZw', with Z as defined in Theorem 2.3.4. Since $Z^\smile$ is total, f is surjective. Select $w_1, w_2 \in W$ such that $f(w_1) = f(w_2)$. By definition of f, w_1 and w_2 satisfy the same proposition variables. Since there are only n worlds in W, one must conclude that $w_1 = w_2$ implying that f is injective, and therefore a bijection.

By definition of Z, $w \in V(p)$ implies $f(w) \in V'(p)$. Conversely, one also has $f(w) \in V'(p)$ implies $w \in V(p)$, thus satisfying condition (i) of Definition 2.3.2.
To prove condition (ii) in 2.3.2, assume wRv. By definition of f, one has $wZf(w)$. By condition (ii) of (the bisimulation) Definition 2.3.4, there exists a $v' \in W'$ such that vZv' and $f(w)R'v'$. But vZv' implies $v' = f(v)$, proving that $f(w)R'F(v)$. Conversely assume $f(w)R'f(v)$. Now $wZf(w)$ and condition (iii) of Definition 2.3.4 implies that there exists $v \in W$ such that $vZf(v)$ and wRv as required. Since f is bijective, it follows that it is an isomorphism. □

Theorem 2.3.7. *Let $\mathcal{M} = (W, R, V)$ be a model over Φ_n that satisfies G-axioms and $\phi^G_{\mathcal{M}}$ its G-defining proposition. Then there exists a bisimilar model $\mathcal{M}' = (W', R', V')$ with an equivalent G-defining proposition and $\#W = n$.*

Proof. Construct the model $\mathcal{M}'$ over Φ_n as follows: Define $W' = \{w_p \mid p \in \Phi_n\}$ containing n worlds indexed by the proposition variables. Define V' by $V'(p) = \{w_p\}$ for any $p \in \Phi_n$. Define R' as follows: Choose $p, q \in \Phi_n$. If $\forall w \in V(p)$, $\exists v \in V(q)$ such that wRv, then $w_pR'w_q$.
Now $\mathcal{M}'$ satisfies G-axioms by construction: One has $M', w_p \models p$ implying $M', w_p \models \bigvee_{i=1}^{n} p_i$. This is true for any $p \in \Phi_n$, thus satisfying **G1**. Similarly, by definition $M', w_p \models \neg q$ for any $q \in \Phi_n$ with $p \neq q$. Thus $M', w_p \models p \rightarrow \neg q$. Further for $r \neq p$, $M', w_r \models p \rightarrow \neg q$ since $M', w_r \models \neg p$. Thus $M' \models p \rightarrow \neg q$. for $p, q \in \Phi_n$ with $p \neq q$ satisfying **G2**. **G3** is trivially satisfied since $M', w_p \models p$ for every $p \in \Phi_n$.
To show equivalent G-defining propositions, assume $\mathcal{M} \models p \rightarrow \Diamond q$. This implies that for every $w \in V(p)$, there exists a $v \in V(q)$ such that wRv. To show this, select $w \in V(p)$. Then $\mathcal{M}, w \vdash p$. This together with $\mathcal{M} \models p \rightarrow \Diamond q$ and modus ponens implies $\mathcal{M}, w \vdash \Diamond q$. By definition, this implies that there exists a world v such that wRv and $\mathcal{M}, v \vdash q$, concluding the statement above. By definition of R' above, one has $w_pR'w_q$. Now since $\mathcal{M}', w_p \vdash p$ and $\mathcal{M}', w_q \vdash q$, one can conclude that $\mathcal{M}', w_p \vdash \Diamond q$. Finally since $V'(p) = w_p$, one has $\mathcal{M}' \vdash p \rightarrow \Diamond q$.
Conversely assume $\mathcal{M}' \models p \rightarrow \Diamond q$. Since $V'(p) = \{w_p\}$ and $V'(q) = \{w_q\}$ (the proposition variables are true at one unique world only) one can deduce that $w_pR'w_q$. By definition of R', $\forall w \in V(p)$, $\exists v \in V(q)$ such that wRv. Thus $\mathcal{M}, w \vdash \Diamond q$ for every $w \in V(p)$. Thus $\mathcal{M} \vdash p \rightarrow \Diamond q$.

I have shown that $\mathcal{M}'$ satisfies G-axioms and has an equivalent G-defining proposition to $\mathcal{M}$. By Theorem 2.3.4, I have that $\mathcal{M}'$ is bisimilar to $\mathcal{M}$, and by construction $\#W' = n$. □

The theorem above concludes my objective: Firstly it shows that I can associate a unique model that satisfies G-axioms with a G-defining proposition i.e. the model containing the least worlds. Secondly, by construction, this model is a G-model, since it clearly satisfies the conditions specified in Definition 2.3.7. Now, all I need to do is prove that every G-model satisfies G-axioms, and I will have that every G-model can be represented by a proposition in $Form(\Diamond, \Phi_n)$.

Theorem 2.3.8. *Let $\mathcal{M} = (W, R, V)$ be a G-model over Φ_n. Then $\mathcal{M}$ satisfies G-axioms.*

Proof. To show **G1**, select $w \in W$. Since $\mathcal{M}$ is a G-model, there exists $p \in \Phi_n$ such that $w \in V(p)$. Thus $\mathcal{M}, w \models p$, implying that $\mathcal{M}, w \models \bigvee_{i=1}^{n} p_i$. Since w was arbitrary, one can conclude that $\mathcal{M} \models \bigvee_{i=1}^{n} p_i$, as required.
To show **G2**, select $p_i, p_j \in \Phi_n$ with $p_i \neq p_j$. Select $w \in W$. Since $\mathcal{M}$ is a G-model, there exists

$p \in \Phi_n$ such that $w \in V(p)$.

Case(i): Assume $p = p_i$. Thus $\mathcal{M}, w \models p_i$. Since $\mathcal{M}$ is a G-model, and $p_j \neq p_i$, one has $V(p) \cap V(p_j) = \emptyset$. Thus $w \notin V(p_j)$. Thus $\mathcal{M}, w \models \neg p_j$ resulting in $\mathcal{M}, w \models p_i \rightarrow \neg p_j$

Case(ii): Assume $p \neq p_i$. Since $\mathcal{M}$ is a G-model, one has $V(p) \cap V(p_i) = \emptyset$. Thus $w \notin V(p_i)$ implying $\mathcal{M}, w \vdash \neg p_i$. One can thus conclude that $\mathcal{M}, w \vdash \neg p_i \vee \neg p_j$ which is equivalent to $\mathcal{M}, w \vdash p_i \rightarrow \neg p_j$

Since w was arbitrary, one has $\mathcal{M} \vdash p_i \rightarrow \neg p_j$ for $i \neq j$ as required.

To show **G3**, assume there exists $p_i \in \Phi_n$ such that $\mathcal{M} \vdash \neg pi$. Since $\mathcal{M}$ is a G-model, there exists a $w \in W$ such that $w \in V(p_i)$. Thus $\mathcal{M}, w \models p_i$, arriving at a contradiction since $\mathcal{M} \vdash \neg pi$ implies $\mathcal{M}, w \vdash \neg pi$. □

I have completed the representation proofs and shown that every G-model can be represented by a G-defining proposition in $Form(\Diamond, \Phi_n)$.

Example 2.3.3. The geometry of the logicatom universe in Example 2.3.2 is completely determined by the G-defining proposition

$$\begin{aligned} \phi \quad &:= \quad ((p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge q \wedge \neg r) \vee (\neg p \wedge \neg q \wedge r)) \\ &\wedge \quad (p \wedge \Diamond q) \\ &\wedge \quad (q \wedge \Diamond r) \\ &\wedge \quad (r \wedge \Diamond q) \end{aligned}$$

The disjunction $((p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge q \wedge \neg r) \vee (\neg p \wedge \neg q \wedge r))$ captures the unique naming constraint while the remainder of the proposition captures the relation.

The advantage of proving that every G-model can be represented by a G-defining proposition is that I will be able to show that the information inherent in the G-model of Definition 2.3.8 is actually superfluous. Firstly note that the relationship between the modal relation, the relation generating function and the substitution map of Definition 2.3.8 can be rewritten using a G-defining proposition ϕ_G as follows:

$$\vdash_K \phi_G \rightarrow (p \rightarrow \Diamond q) \text{ iff } q \in R_f(\xi(p)) \tag{2.3.7}$$

Constraint 2.3.7 allows one to completely determine the G-defining proposition ϕ_G from the pair (ξ, R_f) by

$$\phi^G := \left(\bigvee_{i=1}^{n} p_i\right) \wedge \left(\bigwedge_{i,j=1;i\neq j}^{n} (p_i \rightarrow \neg p_j)\right) \wedge \left(\bigwedge_{i=1}^{n} \bigwedge_{p_j \in R_f(p_i)} p_i \rightarrow \Diamond p_j\right) \tag{2.3.8}$$

The term $\left(\bigvee_{i=1}^{n} p_i\right) \wedge \left(\bigwedge_{i,j=1;i\neq j}^{n} (p_i \rightarrow \neg p_j)\right)$ captures the unique naming constraint while the term $\left(\bigwedge_{i=1}^{n} \bigwedge_{p_j \in R_f(p_i)} p_i \rightarrow \Diamond p_j\right)$ captures the relation. This allows me to specify the final definition of a logicatom universe as:

Definition 2.3.10. A logicatom universe over PropCal_n is defined as the pair (ξ, R_f), where $\xi : \Phi_n \to \mathit{Form}(\Phi_n)$ is a substitution map and $R_f : \mathit{Form}(\Phi_n) \to \mathcal{P}(\Phi_n)$ is a relation generating function over PropCal_n.

The importance of this elegant definition will become evident in the next section. Finally, the transition function can now be viewed as a function that maps the substitution map to a new substitution map with the local constraint imposed. Formally, define $\Sigma = \{\xi \mid \xi : \Phi_n \to \mathit{Form}(\Phi_n)\}$ comprising the set of all substitutions over Φ_n. The transition function is a map $T : \Sigma \to \Sigma$ satisfying the 'local' constraint. This allows me to state an equivalent definition of a theory network:

Definition 2.3.11. A theory network over PropCal_n consists of the tuple (ξ, R_f, T) where (ξ, R_f) is a logicatom universe over PropCal_n and $T : \Sigma \to \Sigma$ is a local transition function mapping a substitution map to a new substitution map according to the local constraint:

$$[T(\xi)]\,(p) = T\,(\xi(p), \xi(q) \mid q \in R_f(\xi(p)))$$

i.e. the transition function operating on $p \in \Phi_n$ is a function of the belief of the logicatom named p and the beliefs of all logicatoms that are related to p.

Once again, the theory network (ξ, R_f, T) can be represented as a series of logicatom universes $((\xi, R_f), (T(\xi), R_f), (T(T(\xi)), R_f), \ldots)$. In terms of notation, I will usually define a theory network by (ξ_0, R_f, T) and represent each application of the transition function using a time parameter $t \in \{0, 1, 2, \ldots\}$ i.e. $\xi_{t+1} := T(\xi_t)$. Thus $(T(T(\xi_0)), R_f)$ would be represented as (ξ_2, R_f). The following examples encompass transition functions that satisfy the local constraint specified.

Example 2.3.4. The transition function I have used in Example 2.2.3 is that of the constrained uniform substitution transition function defined in Definition 2.3.6. In these examples, given a logicatom $\mu \in \mathcal{U}$, we consider all logicatoms $R(\mu)$ in the image set of μ, using the local relation generated by the relation generating function R_f of the logicatom universe $(\mathcal{U}, R_f)$. The beliefs of logicatoms in $R(\mu)$ are then uniformly substituted in $B(\mu)$ to create a new belief. Now it has been shown that a logicatom universe $(\mathcal{U}, R_f)$ can be equivalently represented as (ξ, R_f) for some substitution map $\xi : \Phi_n \to \mathit{Form}(\Phi_n)$. The *constrained uniform substitution* local transition function $T : \Sigma \to \Sigma$ is then defined by:

$$[T(\xi)](p) := \xi(p)^{\xi|_{R_f(p)}} \tag{2.3.9}$$

or using the time parameterised notation,

$$\xi_{t+1}(p) := \xi_t(p)^{\xi_t|_{R_f(p)}} \tag{2.3.10}$$

One can use the logical connectives to create a large number of these transition maps such as

$$[T(\xi)](p) := \xi(p)^{\xi|_{R_f(p)}} \vee \bigwedge_{q \in R_f(p)} \xi(q)^{\xi|_{R_f(p)}}$$

or using the time parameterised notation,

$$\xi_{t+1}(p) := \xi_t(p)^{\xi_t|_{R_f(p)}} \vee \bigwedge_{q \in R_f(p)} \xi_t(q)^{\xi_t|_{R_f(p)}}$$

Finally, a simple example of a transition function is given by

$$[T(\xi)](p) := \xi(p)^{\gamma} \tag{2.3.11}$$

or using the time parameterised notation,

$$\xi_{t+1}(p) := \xi_t(p)^{\gamma} \tag{2.3.12}$$

where γ is some substitution map. As will be shown in Section 2.5.3, transition functions of this type allow one to replicate the dynamics of cellular automaton systems using theory networks.

I now have the tools to go and investigate if the theory networks described formally satisfy all the requirements loosely specified in Section 2.2. I end this subsection with a note on the notation used. When referring to logicatom universes, I will use either Definition 2.2.3 or Definition 2.3.10 depending on the context. When I analyse theory networks at a particular point in time i.e. a logicatom universe, I will ignore the time subscript. On the other hand, if I analyse theory networks over time, the time subscript will be attached to the substitution map, and be incremented by 1 on each application of the transition function. In these cases, I will also parameterise logicatoms $\mu(t)$ with the time parameter. The relation of a logicatom universe will always be assumed to be generated by a relation generating function, and will therefore not have a time parameter attached to it.

2.4 Knowledge within theory networks

I will now analyse two specific examples of theory networks, with the objective of showing how knowledge can be modelled using theory networks. Section 2.4.1 focuses on modelling geometric knowledge. Section 2.4.2 shows how the dynamics of theory networks can be viewed in context with dynamical theories of logic i.e. belief revision.

2.4.1 Beliefs describing G-Models

I have created an entity (the logicatom) that lives in a space (the logicatom universe) and believes *'something'* about the space it lives in. The objective of this section is to show how the current definitions enable one to define what this *'something'* actually is. I will focus on various examples motivated by geometrodynamics, that display the essence of how to construct a theory network that models a particular type of knowledge. Now from a geometrodynamical perspective, I would require

an example of a logicatom believing something about the geometry of the space wherein it lives. The previous section provides the tools to show how this is actually possible. Using our previous results, if a logicatom believes a *G-defining* proposition, its belief can be uniquely represented as a G-model. In order to facilitate this, I simply extend the definition of a logicatom over PropCal_n to one over the basic modal language $ML(\Diamond, \Phi_n)$.

Definition 2.4.1. A *logicatom* μ over $ML(\Diamond, \Phi_n)$ is defined as an element of the cross product $\Phi_n \times \mathit{Form}(\Diamond, \Phi_n)$, represented as an ordered pair (p, ϕ) where $p \in \Phi_n$ is a propositional variable and $\phi \in \mathit{Form}(\Diamond, \Phi_n)$ is a modal formula of the basic modal language.

In extending the definition of a relation generating function over PropCal_n, to one over the basic modal language $ML(\Diamond, \Phi_n)$, I specify that the function generates sets of pairs (p, q) of propositional variables $p, q \in \Phi_n$. The reason for this will become clear in the examples below.

Definition 2.4.2. A *relation generating function* over $ML(\Diamond, \Phi_n)$ is defined as a function $R_f : \mathit{Form}(\Diamond, \Phi_n) \to \mathcal{P}(\Phi_n \times \Phi_n)$ that maps modal formulae to sets of propositional variable pairs.

The definition of a logicatom universe over PropCal_n is now easily extended to one over the basic modal language $ML(\Diamond, \Phi_n)$.

Definition 2.4.3. A logicatom universe over $ML(\Diamond, \Phi_n)$ is defined as the pair (ξ, R_f), where $\xi : \Phi_n \to \mathit{Form}(\Diamond, \Phi_n)$ is a substitution map and $R_f : \mathit{Form}(\Diamond, \Phi_n) \to \mathcal{P}(\Phi_n \times \Phi_n)$ is a relation generating function over $ML(\Diamond, \Phi_n)$.

Similarly for theory networks, one has

Definition 2.4.4. A theory network over $ML(\Diamond, \Phi_n)$ consists of the tuple (ξ, R_f, T) where (ξ, R_f) is a logicatom universe over $ML(\Diamond, \Phi_n)$ and $T : \Sigma \to \Sigma$ is a local transition function mapping a substitution map $\xi \in \Sigma$ to a new substitution map $T(\xi) \in \Sigma$ that satisfies the local constraint:

$$[T(\xi)](p) = T(\xi(p), \xi(q) \mid (p, q) \in R_f(\xi(p)))$$

Now if a logicatom over $ML(\Diamond, \Phi_n)$ believes a G-defining proposition, one can view the belief as a purely geometrical structure i.e. a G-model. Example 2.4.1 shows one interpretation of this geometrical state, as being what the logicatom believes the geometry of its space to be.

Example 2.4.1. Define the relation generating function $R_f : \Omega \times \Phi_n \to \mathcal{P}(\Phi_n \times \Phi_n)$ over $ML(\Diamond, \Phi_n)$ by

$$R_f(\phi) = \{(p, q) \mid \vdash_K \phi \to (p \to \Diamond q)\}$$

The induced local relation R in the logicatom universe is defined by

$$\mu R \nu \text{ iff } (N(\mu), N(\nu)) \in R_f[B(\mu)]$$

Intuitively, this says that logicatom μ is related to ν iff logicatom μ believes it is related to logicatom ν. Consider the logicatom universe $\mathcal{U}$ over PropCal_4 defined in the Figure 2.5, where the belief of each logicatom named p, q, r, s is a G-defining propositions in $\Phi_p = \{p, q, r\}$, $\Phi_q = \{p, q, r, s\}$, $\Phi_r = \{q, r, s\}$ and $\Phi_s = \{p\}$ respectively.

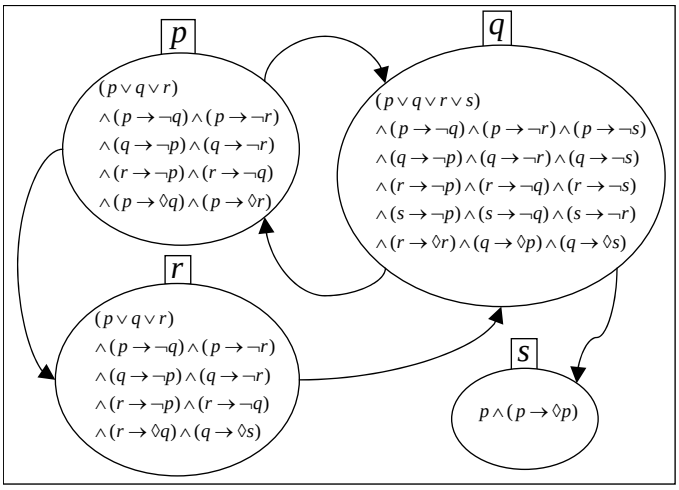

Figure 2.5: A logicatom universe with beliefs that are G-defining propositions

Now by construction, each logicatom belief is a G-defining formula. This allows one, as shown in Figure 2.6 (over the page), to interpret each logicatom belief as a representation of the geometrical structure of a logicatom universe i.e. a G-Model.

I will now proceed to construct an example of a theory network over $ML(\Diamond, \Phi_n)$. I will define the transition function in such a way so as to ensure the beliefs of the logicatoms are always G-defining propositions. Towards this end, I will define a concept that will be used throughout the remainder of this thesis when constructing relation generating functions and transition maps for theory networks; that of a proposition $\phi \in Form(\Diamond, \Phi_n)$ *requiring* a propositional variable $p \in \Phi_n$.

Definition 2.4.5. A propositional variable $p \in \Phi_n$ *is required* in a proposition $\psi \in Form(\Diamond, \Phi)$ iff $\forall \phi \in Form(\Diamond, \Phi)$ satisfying $\vdash_K \phi \leftrightarrow \psi$, variable p occurs in ϕ.

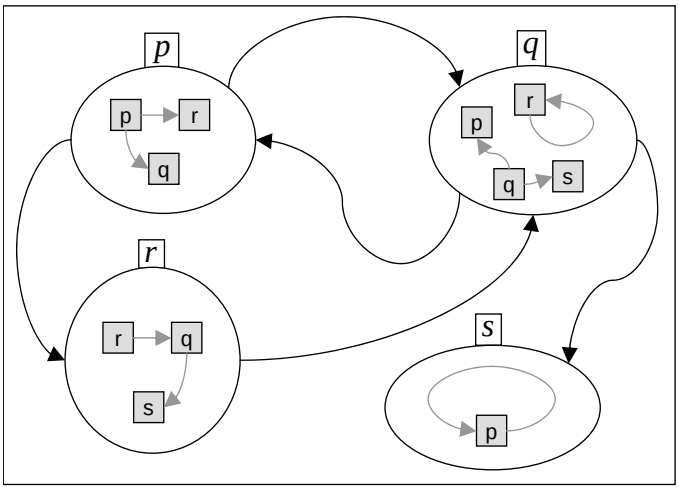

Figure 2.6: How each logicatom views its universe

Example 2.4.2. Let $\phi := (\Diamond q \wedge p) \vee (\Box \neg q \wedge p)$. Then q is not required in ϕ and p is. This can be shown by noticing that

$$\begin{aligned} & \vdash_K \phi \leftrightarrow (\Diamond q \wedge p) \vee (\Box \neg q \wedge p) \\ \text{iff} \quad & \vdash_K \phi \leftrightarrow (\Diamond q \wedge p) \vee (\neg \Diamond q \wedge p) \text{ since } \Diamond p \leftrightarrow \neg \Box \neg p \\ \text{iff} \quad & \vdash_K \phi \leftrightarrow p \text{ since } \vdash_K ((p \wedge \phi) \vee (p \wedge \neg \phi)) \leftrightarrow p \end{aligned}$$

Returning to theory networks over $ML(\Diamond, \Phi_n)$, first note that any G-defining proposition $\phi^G \in ML(\Diamond, \Phi_n)$ can be expressed as

$$\phi^G := \phi_c \wedge \left(\bigwedge_{(p_i, p_j) \in \Phi_n \times \Phi_n} \phi_{ij} \right)$$

where

$$\phi_c := \left(\bigvee_{i=1}^{n} p_i \right) \wedge \left(\bigwedge_{i,j=1; i \neq j}^{n} (p_i \rightarrow \neg p_j) \right)$$

represents the co-ordinate structure (the unique names of each logicatom) and

$$\phi_{ij} := p_i \rightarrow \Diamond p_j \text{ or } \phi_{ij} := \top$$

represent the relational structure. Finally, given a set of G-defining propositions $\Theta = \{\phi_1, \phi_2, \ldots, \phi_k\}$ over Φ_n and propositional variables $p, q \in \Phi_n$, define the subset

$$\Theta_{(p,q)} := \{\phi \in \Theta \mid \vdash_K \phi \rightarrow (p \rightarrow \Diamond q)\} \tag{2.4.1}$$

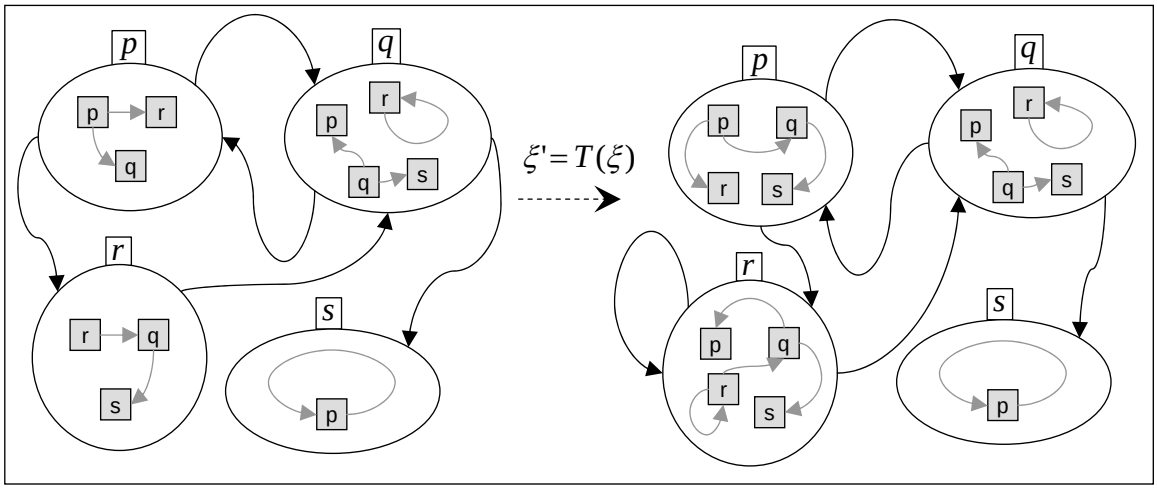

Figure 2.7: A transition rule for logicatoms with beliefs that are G-defining propositions

Example 2.4.3. Let (ξ, R_f, T) be a theory network over $ML(\Diamond, \Phi_4)$ where (ξ, R_f) is the logicatom universe over $ML(\Diamond, \Phi_4)$ defined in Example 2.4.1 and $\Phi_4 = \{p, q, r, s\}$. Define the local transition function $T : \Sigma \to \Sigma$ mapping a substitution map $\xi \in \Sigma$ to a new substitution map $T(\xi) \in \Sigma$ as follows: Given $x \in \Phi_4$, the set $\Theta^x = \{\xi(l) \mid (x, l) \in R_f(\xi(x))\}$ contains the beliefs (G-defining propositions) of all logicatoms related to the logicatom named x.
The set $\Phi^x = \{l \in \Phi_4 \mid \exists \psi \in \Theta^x \cup \xi(x)$ such that l is required in $\psi\}$ contains all the propositional variables required in every G-defining proposition in Θ^x and $\xi(x)$. Now

$$[T(\xi)](x) = \phi_c(x) \wedge \left(\bigwedge_{(u,v) \in \Phi^x \times \Phi^x} \phi_{uv}(x) \right)$$

where $\phi_c(x)$ represents the co-ordinate structure over Φ^x and

$$\phi_{uv}(x) := \begin{cases} u \to \Diamond v & \text{if } \frac{\#\Theta^x_{u,v}}{\#\Theta^x} \geq \frac{1}{2} \text{ or } \vdash_K \xi(x) \to (u \to \Diamond v) \\ \top & \text{otherwise} \end{cases} \tag{2.4.2}$$

The above transition rule behaves as follows: If the majority of logicatoms (more than half of the logicatoms) that I (a logicatom) am related too believe a G-defining proposition that implies $u \to \Diamond v$, then I will believe this on the next iteration. If I (a logicatom) believe a G-defining proposition that implies $u \to \Diamond v$, then I will believe this on the next iteration. The effects of this transition function operating on Example 2.4.1 is shown in Figure 2.7.

The examples defined thus far should show the reader how useful theory networks are in modelling various concepts / ideas.

2.4.2 Beliefs describing monotonic rules

Consider a belief represented by a proposition ϕ in some logic. Belief change encapsulates the problem of incorporating a new 'fact' represented by the proposition ψ in the same language. Dynamical theories of logic such as logic programming have been formulated in several frameworks to analyse this type of problem. Logic programming is a field of artificial intelligence that uses logic directly as a programming language, focusing on the representation of knowledge in flux, unlike that of classical logic. In particular, logic program updating was born in a seminal paper by Marek and Truszczyński [48] who introduced a language for specifying updates to knowledge bases known as revision programs. This field of research has proved very useful in modelling dynamic knowledge. A detailed review of the subject is available in [54, 5, 4, 45, 46]. Another framework is that of probabilistic reasoning [31] that combines both logic and probabilities, providing one with a formalism that admits the power of logic as well as allowing one to express rules with different levels of firmness and to change beliefs in response to a dynamic environment. In the *belief revision* framework, belief change is implemented using a set of constraints (known as postulates) on an operator $\circ$ which modifies the set ϕ of currently held beliefs to produce a new set $\phi \circ \psi$ implying the new information ψ. In particular, the AGM (Alchourrón, Peter Gärdenfors and David Makinson) postulates [3] form the starting point in this field. More recent work has shown that these postulates only cater for specific types of belief revision such as when one is obtaining new information from a static world, and not for cases where the world described is dynamic [60].

Default logic is an example of non-monotonic reasoning. Non-monotonic reasoning is best explained using a direct quote from McCarthy [49]: "Consider putting an axiom in a common sense database asserting that birds can fly. Clearly the axiom must be qualified in some way since penguins, dead birds and birds whose feet are encased in concrete can't fly. A careful construction of the axiom might succeed in including the exceptions of penguins and dead birds, but clearly one can think up as many additional exceptions like birds with their feet encased in concrete as one likes. Formalised nonmonotonic reasoning provides a way of saying that a bird can fly unless there is an abnormal circumstance and reasoning that only the abnormal circumstances whose existence follows from the facts being taken into account will be considered." -John McCarthy
Nonmonotonic reasoning gained momentum in 1980 with the publication of an issue of the Artificial Intelligence Journal devoted exclusively to the subject. Various papers therein extended the classical logic system in various ways. For example, McDermott and Doyle [50] introduced a modal operator $\langle M \rangle$ into first-order logic. $\langle M \rangle p$ was interpreted as meaning that 'p is consistent with what is

known'. Auto-epistemic logic [56] is one of the nonmonotonic logics that can be obtained from the approach of McDermott and Doyle. It was in this same issue that Reiter [68] developed another example of a non-monotonic logic known as default logic. Default logic consists of classical first order logic together with default rules.

Definition 2.4.6. A *default rule* d takes on the form

$$d := \frac{\psi : \phi_1, \phi_2, \ldots, \phi_n}{\eta}$$

The default comprises

the *prerequisite* $\text{preq}(d) := \psi$

the consistency requirements or *justifications* $\text{just}(d) := \phi_1, \phi_2 \ldots, \phi_n$

and the *consequent* $\text{cons}(d) := \eta$.

The prerequisite ψ together with the consequent η is known as the *monotonic rule* $\frac{\psi}{\eta}$. The default rule $\frac{\psi:\phi_1,\phi_2,\ldots,\phi_n}{\eta}$ is said to consist of a monotonic rule $\frac{\psi}{\eta}$ and the consistency requirements $\{\phi_1, \phi_2, \ldots, \phi_n\}$.

The default rule $\frac{\psi:\phi_1,\phi_2,\ldots,\phi_n}{\eta}$ is interpreted as follows: 'If ψ is derivable and for all i $(1 \leq i \leq n)$, $\neg\phi_i$ is not derivable, then derive η'.

Example 2.4.4. Consider the default with free variables: 'If x is a quaker and it cannot be proved that x is not a pacifist then deduce that x is a pacifist'. This is represented as $\frac{\text{quaker}(x):\text{pacifist}(x)}{\text{pacifist}(x)}$.

Definition 2.4.7. *Rule-based reasoning* encompasses a set V of propositions in classical logic together with a set $\mathcal{R}$ of monotonic rules $\frac{\phi_1,\ldots,\phi_n}{\eta}$.
$V \vdash_R \phi$ holds iff there is a finite sequence of propositions $\phi_1, \ldots, \phi_n$ with $\phi_n := \phi$ such that for every $i \in \{1, 2, \ldots, n\}$, either

- $V \cup \{\phi_1, \ldots, \phi_{k-1}\} \vdash \phi_k$

where the inference $\vdash$ defined in classical propositional calculus or

- there exists a rule $\frac{\psi_1,\ldots,\psi_k}{\phi_i} \in R$ such that $\{\psi_1, \ldots, \psi_k\} \subseteq \{\phi_1, \ldots, \phi_{i-1}\}$

The following simple example shows one how rule-based reasoning works.

Example 2.4.5. From $\{p\}$ and the monotonic rules $\{\frac{\neg\neg p}{q}, \frac{\neg\neg q,p}{r}\}$, $\{r\}$ can be derived i.e. $\{p\} \vdash_R r$

1. $\neg\neg p$: since $\{p\} \vdash \neg\neg p$ in PropCal
2. q : using the rule $\frac{\neg\neg p}{q}$ and 1.
3. $\neg\neg q$: since $\{q\} \vdash \neg\neg q$ in PropCal

4. r : using the rule $\frac{\neg\neg q, p}{r}$ and 3.

A *default theory* is defined to be a pair (D, W) where D is a set of default rules and W a set of first order propositions. I will end at this point, except to note that interpreting the default rules as mappings from some incomplete theory to a more complete extension of the theory motivated important concepts such as *extensions* in default logic. Even though this has been a very superficial review on the extensive field of belief revision (and in particular, default logic), it provides me with the tools required to show that the dynamics of certain theory networks can interpreted in terms of rule-based reasoning.

Lemma 2.4.1. *Consider a theory network $\mathcal{T} = (\xi, R_f, T)$ over PropCal$_n$ with $\Phi_n = \{p_1, p_2, \ldots, p_n\}$. The local relation generating function R_f is defined by $R_f(\phi) = \{p_i \in \Phi_n \mid \vdash \phi \rightarrow p_i\}$ and the transition rule encapsulated using constrained uniform substitution i.e. $T(\xi)(p_i) = \xi(p_i)^{\xi|_{R_f(p_i)}}$. For every logicatom $p_i \in \Phi_n$, define the set of monotonic rules:*

$$\mathcal{R}(p_i) = \left\{ \frac{p_j}{\xi(p_j)} \mid p_j \in R_f(p_i) \right\} \tag{2.4.3}$$

Then

$$\xi(p_i) \vdash_{\mathcal{R}(p_i)} [T(\xi)](p_i) \tag{2.4.4}$$

Proof. One needs to prove that $[T(\xi)](p_i)$ is derivable from $\xi(p_i)$ and the monotonic rules $\mathcal{R}(p_i)$. This is done by first noting that

$$\vdash \xi(p_i) \leftrightarrow \left(\bigwedge_{p_j \in R_f(\xi(p_i))} p_j \wedge \eta \right) \tag{2.4.5}$$

with η not requiring p_j for any $p_j \in R_f(\xi(p_i))$. This relation is obtained by noting that if $\vdash \phi \rightarrow q$ then $\vdash \phi \leftrightarrow (q \wedge \psi)$ with ψ not requiring q. An inductive argument will result in expression 2.4.5. Applying the constrained uniform substitution transition function results in

$$\vdash [T(\xi)](p_i) \leftrightarrow \left(\bigwedge_{p_j \in R_f(\xi(p_i))} \xi(p_j) \wedge \eta \right) \tag{2.4.6}$$

Now using 2.4.5 and $\wedge$ elimination in propositional calculus, one has

$$\xi(p_i) \vdash_{\mathcal{R}(p_i)} \eta \tag{2.4.7}$$

and

$$\xi(p_i) \vdash_{\mathcal{R}(p_i)} p_j \tag{2.4.8}$$

for every $p_j \in R_f(\xi(p_i))$. Now

$$\xi(p_i) \vdash_{\mathcal{R}(p_i)} p_j \tag{2.4.9}$$

together with the monotonic rule

$$\frac{p_j}{\xi(p_j)} \in \mathcal{R}(p_i) \tag{2.4.10}$$

allows one to deduce

$$\xi(p_i) \vdash_{\mathcal{R}(p_i)} \xi(p_j) \tag{2.4.11}$$

for every $p_j \in R_f(\xi(p_i))$, and thus

$$\xi(p_i) \vdash_{\mathcal{R}(p_i)} \bigwedge_{p_j \in R_f(\xi(p_i))} \xi(p_j) \tag{2.4.12}$$

This together with expression 2.4.7 gives one

$$\xi(p_i) \vdash_{\mathcal{R}(p_i)} \bigwedge_{p_j \in R_f(\xi(p_i))} \xi(p_j) \wedge \eta \tag{2.4.13}$$

Finally, using 2.4.6, one deduces $\xi(p_i) \vdash_{\mathcal{R}(p_i)} [T(\xi)](p_i)$ as required □

The above lemma tells one that the updated belief of every logicatom in the theory network defined can be viewed as just a consequence of monotonic rules, specified by the substitution map ξ and relation generating function R_f, applied on the current belief. Thus one can see that selecting a particular transition function and local relation allows one to put theory network dynamics in context with belief revision, in particular, rule-based reasoning. This important example of a theory network will arise again in Chapter 4, albeit in a different guise.

2.5 Theory networks and cellular automata

Birds do it!
Bees do it!
Even theory networks do it! ...

This section details the relationship between theory networks and cellular automata. Section 2.5.1 provides an overview of the vast field of cellular automata. For the purpose of this thesis I shall simply define a cellular automaton system and focus on its application to modelling self-replicating systems. Section 2.5.3 then shows how theory networks can emulate the dynamics of cellular automaton systems. Theory networks thus inherit a large portion of research done within cellular automaton systems, in particular, the applications of cellular automata systems to various fields of study. This is important in terms of the construction objective stated, for the following main reason: Equating memes with the quantum of knowledge, one can view a particular belief as being made up of a whole set of memes. Now some beliefs are constructed in such a way, so as to make self-reproduction a fundamental part of the belief e.g. an evangelical religion. Thus any mathematical structure that claims to model memes must be rich enough to model entities (made up of memes) that have the capability of making copies of themselves.

2.5.1 The cellular automaton system

Ulam proposed the concept of cellular automata. Ulam's idea was to construct a mathematically defined model of the physical universe in which one could build a wide range 'machines'. The model would have cells with connections and a set of locally defined laws that held at every point in the space. These locally defined laws encompassed transition functions that specified how the state of a cell changed with respect to the state of all its neighbouring cells. Conceptually, cellular automata are portrayed using an array of identical processing units called cells that are interconnected throughout the cellular space in some regular manner [67]. Each cell can typically be in any one of two or more possible states that can change over synchronous time $t = 0, 1, 2, 3, \ldots$. I will formally define the cellular space using the concept of a lattice geometry.

Definition 2.5.1. An n-dimensional *lattice geometry* (L, U) consists of the set $L \subseteq \mathbb{Z}^n$ of cells and a set $U \subseteq L \times L$ of edges. The edges define the neighbours of the cells. One can generate the edge set U using the concept of direction vectors. Given $x \in L$, a direction vector $d_x \in \mathbb{Z}^n$ is such that $(x, x + d_x) \in U$. One can thus equivalently define an n-dimensional lattice geometry as (L, D) where $D = \{d_x | x \in L\}$ is the set of direction vectors of every cell x.

Example 2.5.1. Examples of various lattice geometries are provided by Meyer [53]. In particular, the two dimensional honeycomb lattice geometry is shown Figure 2.8 below:

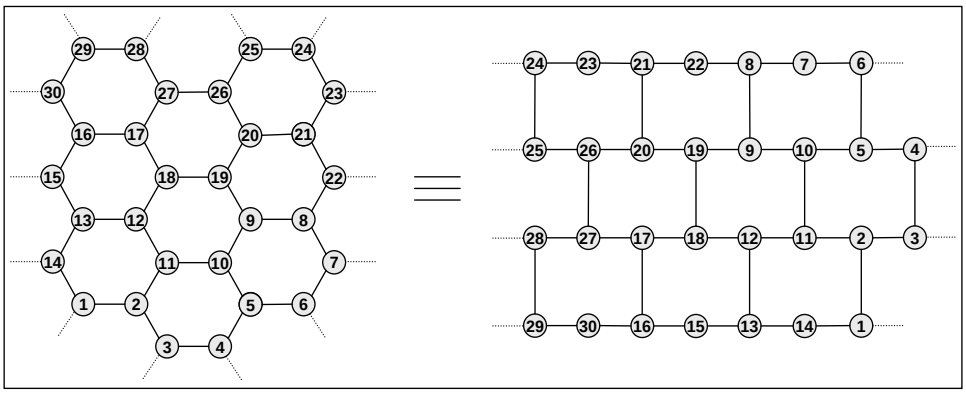

Figure 2.8: Two equivalent representations of the honeycomb lattice geometry

The set of cells $L = \mathbb{Z}^2$. In order to define the direction vectors, Meyer defines the notion of even and odd cells. A cell $(x, y) \in \mathbb{Z}^2$ is odd iff $x + y$ is odd, otherwise the cell is even. The direction vectors are then defined as follows: If $\overrightarrow{x}$ is even then $y_x \in \{(1, 0); (-1, 0); (0, 1)\}$ else if $\overrightarrow{x}$ is odd then $y_x \in \{(1, 0); (-1, 0); (0, -1)\}$ This allows one to invoke the following 'co-ordinate system' for the honeycomb lattice:

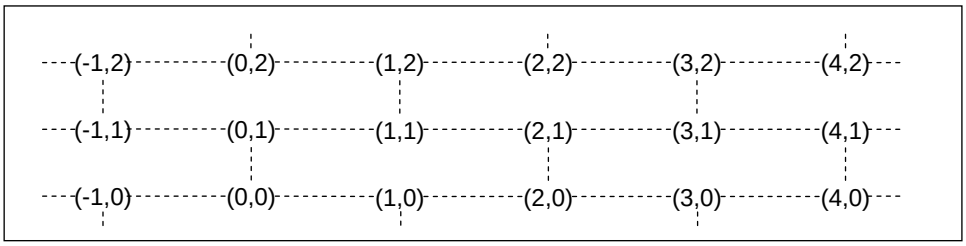

Figure 2.9: A co-ordinate system for the honeycomb lattice

Cellular automata models are usually defined on an isotropic space. Towards this end, the notion of a *Regular Lattice Geometry* is defined. A n-dimensional lattice geometry is *regular* if it can be transformed into itself by translations in n independent directions i.e. the space is isotropic. Using the definition of direction vectors, I can define a *Regular Lattice Geometry* as a lattice geometry L together with a direction set $D = \{y_i | 1 \leq i \leq n; y_i \in \mathbb{Z}^n\}$ of linearly independent vectors such that $\forall x \in L$, $\forall y \in D$, $(x, x + y) \in U$ and $\forall (x, z) \in U$, $\exists y \in D$ such that $z = x + y$.

Example 2.5.2. The triangular lattice is an example of a regular lattice geometry. The set of cells is once again $L = \mathbb{Z}^2$.

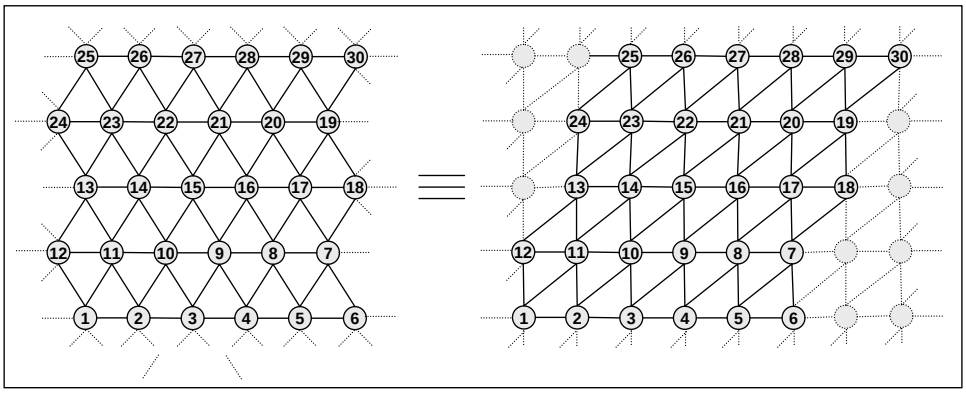

Figure 2.10: Two equivalent representations of the triangular lattice

The set of direction vectors is defined by $D = \{(1,0); (-1,0); (0,1); (0,-1); (1,1); (-1,-1)\}$

Definition 2.5.2. A *Cellular Automaton System* consists of a lattice geometry of cells (L, D) together with a field $\omega : \mathbb{N} \times L \rightarrow S$. The natural numbers $\mathbb{N}$ label the discrete synchronous time-steps. S represents a finite set of possible states that each cell can be in and includes a single quiescent state (i.e. a null state). The field ω defines the state of each cell at a given point in time. ω evolves according to some local rule (referred to as the transition function) $\omega_{t+1}(x) = F\left(\omega_t(x+d) \mid d \in D(x) \cup \overrightarrow{0}\right)$ where $D(x)$ is the set of direction vectors defining the local neighbourhoods of x, and $\overrightarrow{0}$ represents the 0-vector in L. The field ω_0 is constrained to map a finite number of cells to a non-quiescent state. The transition function F is constrained to map cells in the quiescent state and whose neighbours are all in the quiescent state to the quiescent state.

Note that the last two constraints imply that ω_t will always map a finite number of cells to a non-quiescent state. These constraints thus enable one to model cellular automaton systems using finite lattices. This finite property will prove important when I show that theory networks can simulate cellular automaton systems.

Example 2.5.3. A popular example of a cellular automaton system is Conway's [6] model called 'Life'. The regular lattice geometry used by Conway's life is the double triangular lattice geometry. Here $L = \mathbb{Z}^2$ and $D = \{(1,0); (-1,0); (0,1); (0,-1); (1,1); (-1,-1); (1,-1); (-1,1)\}$

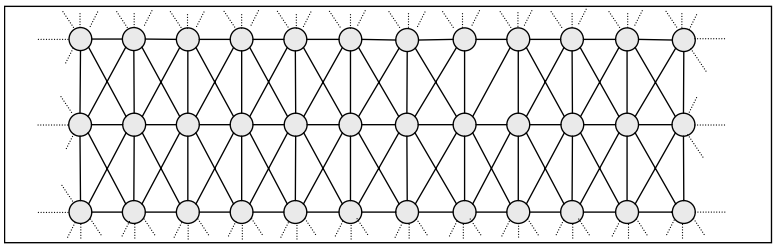

Figure 2.11: The double triangular lattice used in Conway's life

The set $S = \{0, 1\}$ consists of two states with the transition function defined as follows:

1. If at time t, a cell is in state 0 and exactly 3 of its local neighbours are in state 1, then the state of the cell at time $t + 1$ is 1, otherwise it remains 0
2. If at time t, a cell is in state 1 and either 2 or 3 neighbours are in state 1, then the state of the cell at time $t + 1$ remains 1, otherwise it becomes 0

A classic example of the 'emergent behaviour' evidenced in cellular automata is that of the 'glider' in Conway's life [39]: A particular configuration of cells in state 1 move across the lattice and maintain their shape with a period of four time-steps.

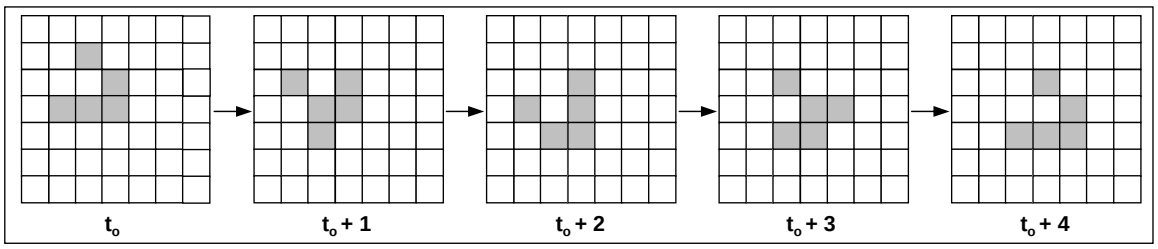

Figure 2.12: The 'Glider' in Conway's life

A host of other structures can be created, and through their interactions, it can be shown that life can simulate a universal Turing machine [6], allowing one to deduce that it can be programmed to perform any desired calculation.

Depending on the properties of the transition function F and the field ω, one can define various types of cellular automata including *deterministic*, *indeterministic* [52] and *probabilistic* cellular automata [9]. There are various equivalent definitions [40] and representations [76] of cellular automata, providing one with the tools to abstractly analyse these structures. For example, given a finite set

S and a dimension d, one defines the *lattice* as the set $L = \mathbb{Z}^d$ and the *shift space* as the topological product $S^L = \{f : L \to S\}$ where S is given the discrete topology and S^L the product topology. Each lattice direction determines a natural *shift operator* σ_i. Then a purely dynamical definition of a cellular automaton system is as a continuous map $G : S^L \to S^L$ that commutes with every σ_i.

2.5.2 Self-reproducing automata

"Von Neumann was interested in the general question: What kind of logical organisation is sufficient for an automaton to be able to reproduce itself? The question is not precise and admits to trivial versions as well as interesting ones. Von Neumann had the familiar natural phenomena of self-reproduction in mind when he posed it, but he was not trying to simulate the self-reproduction of a natural system at the levels of genetics and biochemistry. He wished to abstract from the natural self-reproduction problem its logical form" [9]. The notion of using cellular automata as the platform in which to model self-reproduction was suggested to von Neumann by Ulam [9]. Von Neumann then proceeded to show that self-reproducing automatons could be designed, by constructing a two-dimensional automaton system with 29 states per cell. The lattice used was the 'box lattice geometry'. Here $L = \mathbb{Z}^2$ and $D = \{(1,0);(-1,0);(0,1);(0,-1);\}$. The transition function was deterministic and the same for every cell in the space. His genius came in specifying a particular cellular configuration that would self-replicate. Von Neumann was able to exhibit a universal Turing machine embedded in a cellular array. Further, his universal computer was modified so that as output, it could construct in the cellular array, any configuration described on its input tape. This universal constructor would construct any machine described on the tape and in addition, also construct a copy of the input tape, attaching it to the constructed machine [43]. Self reproduction happened when the machine described on the tape was the universal constructor itself. The actual structure occupied tens of thousands of cells and was never practically implemented. Since this seminal work, various researchers have studied the algorithms needed to support self-replicating systems [43, 65, 67]. The machines have been significantly simplified to structures that occupy relatively few cells (orders of ten) with the number of states each cell can take on significantly reduced. Some of the most exciting new research has been in showing how self-replicating structures can *emerge* in cellular automaton spaces. In the cases mentioned thus far, models of self-replication were initialised with an original copy of the structure that would self-replicate, the transition function working for this specific structure only. Chou and Reggia [66] showed that certain cellular automaton models have self-replicating structures emerging from an initial *random* configuration.

2.5.3 Projected theory networks

Theory networks by construction, share many traits with cellular automata systems. The logicatoms are analogous to the cells, the local relation specifies the lattice geometry while the cell states can correspond to the various beliefs. However, the fundamental difference is that the neighbourhood of a logicatom is dynamic. As opposed to structurally dynamic cellular automata [41], the structure changes in theory networks are all determined locally and not by some global function. Theory networks are in fact more general structures: I will prove this by showing that theory networks can imitate the dynamics of cellular automata systems. The example below will show the methodology I will follow to achieve this objective.

Example 2.5.4. Consider a cellular automata system defined on a 1-dimensional lattice geometry, with $S = \{0, 1\}$ i.e. the cells can be in 1 of 2 states. The automata are indexed using the integers $\mathbb{Z}$. The transition rule F is defined as follows: If at time t , a cells left neighbour is in state 1 and the cell is in state 0, then the state of the cell at $t + 1$ is 1, otherwise it is 0. This simple cellular automata system is shown in the diagram below, with an initial configuration and the 3 succeeding time steps.

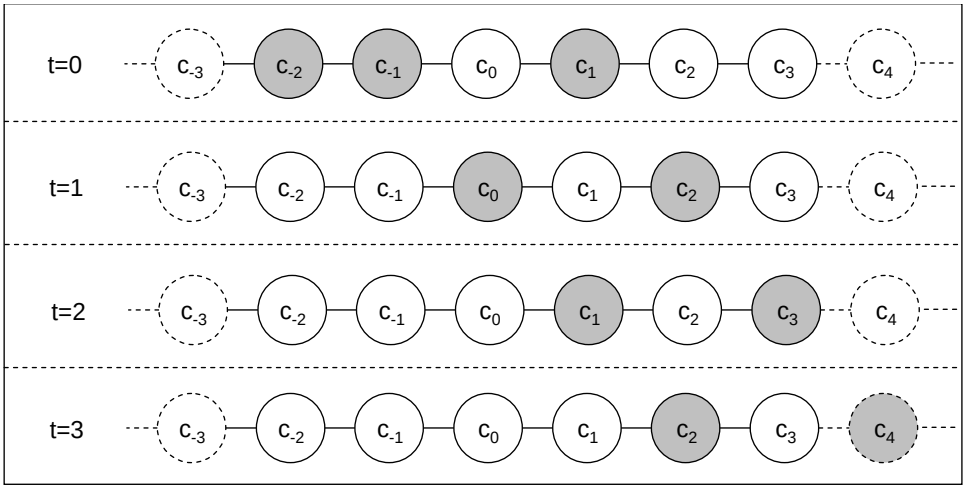

Figure 2.13: The 1-dimensional 'left neighbour' cellular automata system

Observe that since the state space S of the cellular automaton system is just $\{0, 1\}$, one can represent the transition function using propositional logic. Using $\omega_t(x)$ to represent the state of cell x at time t define

$$\omega_{t+1}(x) := \neg\omega_t(x) \wedge \omega_t(x - 1) \tag{2.5.1}$$

This observation allows one to link theory networks with cellular automaton systems. In particular, define the theory network (ξ_0, R_f, T) over $\mathrm{PropCal}_{2n}$ for some large $n \in \mathbb{N}$, as follows: Let $\{\mu_i = (p_i, \phi_i) \mid -n \leq i \leq n\}$ be the logicatoms in the universe. The relation generating map

$R_f : Form(\Phi_{2n}) \to \mathcal{P}(\Phi_{2n})$ is defined by [3]

$$R_f(\phi) = \{p \in \Phi_{2n} \mid \phi \text{ requires } p\}$$

The substitution map ξ_0 is defined by

$$\xi_0(p_i) = \begin{cases} \neg p_i \wedge p_{i-1} & \text{if } i \neq -n \\ \neg p_{-n} \wedge p_n & \text{if } i = -n \end{cases} \tag{2.5.2}$$

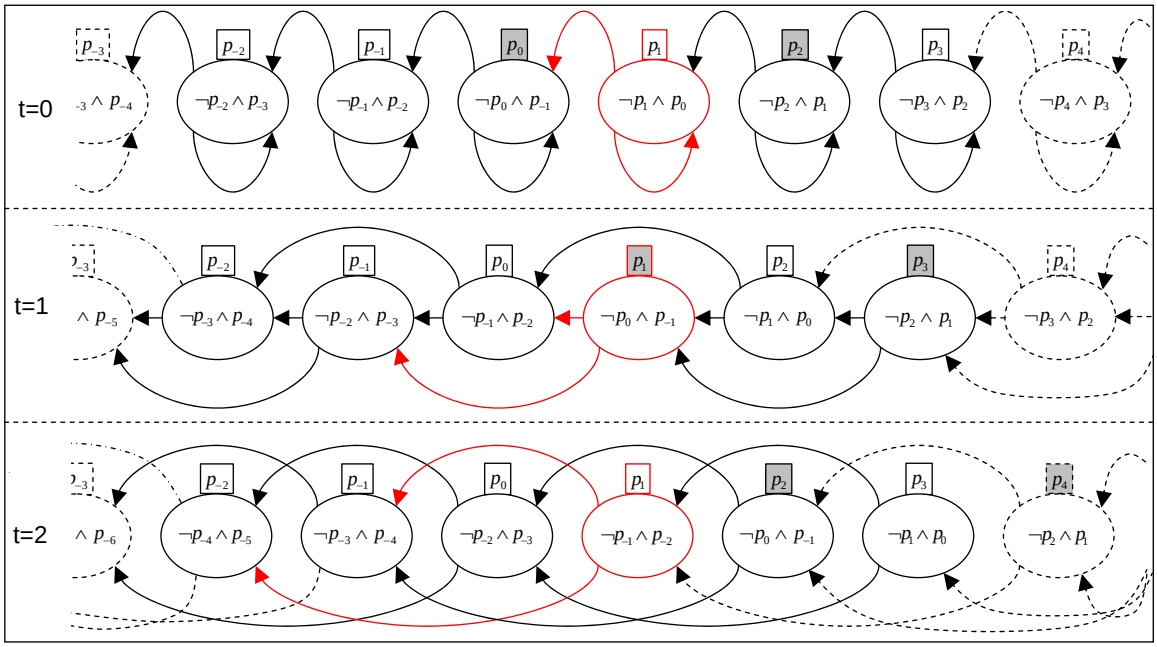

Figure 2.14: The associated cellular automata system

Finally, the transition function T is defined by [4]

$$[T(\xi_t)](p_i) = \xi_{t+1}(p_i) = \xi_t(p_i)^{\xi_0} \tag{2.5.3}$$

The initial configuration of logicatoms, together with the following 2 time steps is shown in Figure 2.14. In order to see that the theory network above simulates the defined cellular automaton system, define the valuation function $V : \Phi_n \to \{0, 1\}$ by

$$V(p_i) = \begin{cases} 1 & \text{if } i \in \{-1, -2, 1\} \\ 0 & \text{otherwise} \end{cases}$$

Inductively extend this valuation to propositions in $Form(\Phi_n)$. Now define the state of a logicatom μ to be $V(B(\mu))$. One sees that only the logicatoms named p_0 and p_2 are in state 1 initially. The

[3] See Definition 2.4.5
[4] See 2.3.12 in Example 2.3.4

next iteration sees the logicatoms named p_1 and p_3 in state 1, the rest in state 0. The third iteration sees logicatoms named p_2 and p_4 in state 1, the rest in state 0. What is in effect happening is that the beliefs are being updated in such a way so as to specify the state of the logicatom in terms of the initial configuration. Thus at time $t = 1$ in the diagram above, the logicatom named p_1 has belief $\neg p_0 \wedge p_{-1}$ which may be interpreted as saying that it will be in state 1 iff initially p_0 was in state 0 and p_{-1} was in state 1.

I will use the above example to illustrate how a theory network can simulate a cellular automaton system.

Definition 2.5.3. Let $\mathcal{T} = (\xi_0\, R_f, T)$ be a theory network over PropCal_n. Let $\sigma : Form(\Phi_n) \to S$ be a function mapping propositional formulae to a finite state space S that includes a single quiescent state (a null state). The pair $(\mathcal{T}, \sigma)$ is known as a *projected theory network*.

The motivation behind adding the state function σ to a theory network tuple in Definition 2.5.3 is that it will allow one to analyse the dynamics of a particular projection of the beliefs of each logicatoms. In particular, the 'state' of a logicatom μ named p at time t in a projected theory network $(\xi_0\, R_f, T, \sigma)$ will be $\sigma(\xi_t(p))$ In Example 2.5.4, this projection was just the usual valuation map in PropCal_n (See section 2.1). The state space in that example was none other than {TRUE, FALSE}. This concept of projecting to a state will be used again in Chapter 4. I will now show how a restricted class of projected theory networks simulate a particular class of cellular automata systems.

Theorem 2.5.1. *Let $\mathcal{C} = (L, \omega, S, F)$ be a cellular automaton system with a finite lattice geometry of cells L, a field ω describing the initial states of the cells, the set of states $S = \{0, 1\}$ restricting the cells to one of two states and a transition function F. Then there exists a projected theory network $\mathcal{T}$ that simulates $\mathcal{C}$.*

Proof. To prove the result, I represent the cellular automaton system $\mathcal{C}$ using propositional logic. Let $\Phi_L = \{p_y \mid y \in L\}$ and define $\Phi_x = \{p_z \mid z = x \text{ or } z \in D(x)\} \subseteq \Phi_L$, where $D(x)$ is the set containing all the neighbours of x. The following construction process associates a proposition $\phi_x \in Form(\Phi_x)$ with the transition function defined at cell x. If the cell x has $n - 1$ neighbours, there will be 2^n possible configurations of states that the cell together with all its neighbours can be in. Each of these configurations implies the next state of the cell. Consider the configurations that imply the next state of cell x to be 1. Call this set of configurations $\mathcal{C}$. Now for every configuration $c \in \mathcal{C}$, define the unique proposition $\bigwedge_{z \in \Phi_x} q_c(z)$ where

$$q_c(z) := \begin{cases} p_z & \text{if cell } z \text{ is in state 1 in configuration } c \\ \neg p_z & \text{if cell } z \text{ is in state 0 in configuration } c \end{cases} \tag{2.5.4}$$

Now define

$$\phi_x^F := \bigvee_{c \in \mathcal{C}} \bigwedge_{z \in \Phi_x} q_c(z) \tag{2.5.5}$$

All I have done is used propositional logic to describe a formula that only uses boolean variables. By construction, any valuation $V : \Phi_L \to \{0, 1\}$ that represents the state of the cells at time t (i.e. Cell

x is in state 1 iff $V(p_x) = 1$) will map ϕ_x^F to the next state of cell x. Intuitively, I have rewritten the cellular automata transition function F and specified it using a set of propositions ϕ_x^F, one for every cell. I define the substitution function

$$\xi_0(p_x) = \phi_x^F \tag{2.5.6}$$

that embodies the cellular automaton transition function F. To observe this, define the valuation $V_0 : \Phi_L \to \{0, 1\}$ by

$$V_0(p_x) = \omega_t(x) \tag{2.5.7}$$

where $\omega_t(x)$ is the field describing the states of the cells at time t. Then

$$V_0((p_x)^{\xi_0}) = V_0(\phi_x) = \omega_{t+1}(x) \tag{2.5.8}$$

Now construct the projected theory network (ξ_0, R_f, T, σ) over PropCal_L as follows: The initial substitution map ξ_0 is given by

$$\xi_0(p_x) = \phi_x^F \tag{2.5.9}$$

The relation generating function embodies the *requires relation* defined using Definition 2.4.5.

$$R_f(\phi) = \{p_x \mid \phi \text{ requires } p_x\} \tag{2.5.10}$$

The transition function T is given by

$$\xi_{t+1}(p_x) = [T(\xi_t)](p_x) = [\xi_t(p_x)]^{\xi_0} \tag{2.5.11}$$

for every $p_x \in \Phi_L$. The state map $\sigma : Form(\Phi_L) \to S$ with state space $S = \{0, 1\}$ is given by:

$$\sigma(\phi) := V_0(\phi) \tag{2.5.12}$$

where V_0 represents the valuation $V_0(p_x) = w_0(x)$ inductively extended to all propositions in $Form(\Phi_n)$. $w_0(x)$ is just the initial cell states of the cellular automata system. To prove the result, I will show that $\omega_t(x) = \sigma(\xi_{t-1}(p_x))$, for every $x \in L$ and all times $t \geq 1$. This will be done by induction. Firstly note that

$$\sigma(\xi_0(p_x)) = \sigma(\phi_x) = V_0(\phi_x) = \omega_1(x) \tag{2.5.13}$$

showing the case for $t = 1$. Now assume that the case holds for $t = k$.

$$\sigma(\xi_{k-1}(p_x)) = V_0(\xi_{k-1}(p_x)) = \omega_k(x) \tag{2.5.14}$$

Then

$$\sigma(\xi_k(p_x)) = \sigma([\xi_{k-1}(p_x)]^{\xi_0}) = V_0([\xi_{k-1}(p_x)]^{\xi_0}) = \omega_{k+1}(x) \tag{2.5.15}$$

using 2.5.7 and 2.5.8. In particular, 2.5.15 is obtained by observing that if $\phi, \psi \in Form(\Phi_n)$ are propositions such that $V(\phi) = V(\psi)$, then $V(\phi^\xi) = V(\psi^\xi)$. (The uniform substitution inference rule in the Hilbert style system specified in Section 2.1). Now since $V_0(\xi_{t-1}(p_x)) = V_0(p_x)$, one has $V_0([\xi_{t-1}(p_x)]^{\xi_0}) = V_0((p_x)^{\xi_0})$ as required. □

Theorem 2.5.1 shows that theory networks can simulate the dynamics of a 2-state ($S = \{0, 1\}$) cellular automaton system. But 'Life' is a 2-state cellular automaton system that can simulate a universal Turing machine [6]. This implies that theory networks have the capability of simulating a universal Turing machine. Since both have the capabilities of simulating universal Turing machines, one can conclude that they are mathematically equivalent from a simulation perspective. However, as I believe I have showed (and will continue to show) in this thesis, the expressive language available in theory networks (by construction) makes them a superior modelling platform.

2.6 Summary of Chapter 2

The logicatom and its dynamic universe, the theory network was defined. In there most general format, these logicatoms have been defined over the modal language in description logic, making them an ideal platform with which to model the dynamics of knowledge. It was shown how the beliefs of logicatoms can be represented as purely geometrical structures i.e. G-defining propositions. A case study also showed how the dynamics of certain theory networks can be viewed in terms of rule-based reasoning. Finally projected theory networks were defined, allowing one to analyse how the dynamics of a specified projection of a logicatom's belief evolved over time. The existence of projected theory networks that emulate the dynamics of 2-state cellular automaton systems was affirmed, thus showing their equivalence with cellular automaton systems and their ability to simulate a universal Turing machine and consequently, model self-replicating structures.

Chapter 3

Incorporating Natural Selection

"Natural selection is the main agent controlling the composition of a species during the course of time, eliminating certain variants and thus preventing change in some directions, making other variants more prevalent and hence producing evolutionary change in other directions" [1].

In this chapter I will mathematically derive the necessary and sufficient requirements for it to be said that natural selection regulates the dynamics of a space. I will then proceed to investigate if the dynamics of theory networks can be said to be regulated by natural selection. The results obtained will lead me to the construction of a mathematical space, known as an inverted theory network, whereupon I prove that there exists an inverted theory network that is regulated by natural selection.

Section 3.1 and 3.2 provide an overview of the principle of natural selection and its applications in various fields of study. In section 3.3, I mathematically formalise the concepts introduced, and provide all the necessary proofs to imply that a structure is regulated by natural selection. As an aside, Section 3.4 shows how this formalism can be applied to explain why sex evolved. In section 3.5, I achieve the construction objective by defining a mathematical space that is regulated by natural selection. Section 3.6 summarises the results of this chapter.

3.1 The biological requirements for natural selection

The theory of evolution, proposed by Charles Darwin in his influential work 'The origin of species' [14], explains the development of new living organisms from pre-existing ones. Darwin's theory can be broken up into 4 distinct parts [18], namely:

1. The Theory of Evolution: This states that natural living organisms have been evolving since the dawn of life [44].

2. The Theory of Common Descent: This purports that all observed species on earth are the modified descendants of one or a few simple organisms [14].

3. The Theory of Gradation: This asserts that the evolutionary process comprises small gradual changes in the organisms from one generation to the next.

4. The Theory of Natural Selection: This defines the process that governs evolution. Natural selection can be viewed as a regulating principle in which the environment (encompassing factors such as climate, competition with other organisms, availability of certain types of food etc) determines which members of a species will successfully reproduce and thus pass on their traits to the next generation.

Now all living organisms can be viewed as a duality of their genotype (the underlying genetic coding), and their phenotype (their manner of response to the environment, encapsulated in their behaviour, physiology and morphology) [26]. The relationship between a genotype and a phenotype is very complex in natural organisms. A single gene may simultaneously affect several phenotypic traits (pleiotropy) and a single phenotypic characteristic may be determined by the simultaneous interaction of a number of genes (polygeny). Natural selection operates on the phenotypic expression of a genotype with the consequence of regulating the frequency of this genotype in a population. Natural selection thus facilitates the rise of highly complex organisms from simple ones, through the continued preferential survival and reproduction of those members of the environment that have traits best suited to deal with the environment [44]. The neo-Darwinian argument asserts that natural selection is the predominant mediating evolutionary 'force' that prevails in shaping the phenotypic characters of organisms in nature. Within the context of theory networks, the primary question that requires answering, if one is to prove (as is required by the construction objective) that natural selection regulates the dynamics of a space is: What are the necessary and sufficient requirements for natural selection to occur? Darwin stated that natural selection occurs because individuals have varying traits, some of these traits are linked to differences in fitness (longevity, fecundity) and some of those traits are heritable [34, 47].

More formally,

Definition 3.1.1. Evolutionary biology states that the necessary and sufficient requirements for it to be said that natural selection regulates a space, are [47]:

1. Heritability: This embodies the requirement that the traits must be inheritable.
2. Differential Fitness: The requirement of differential fitness means that the variation in traits affects reproduction and survival rates of the individuals, i.e. fitness.
3. Phenotypic Variation: The requirement of variation means that different individuals in the environment have different phenotypic traits.

3.2 Natural selection in action

In recent years, the regulating principle of natural selection has been applied to research spheres over and above that of evolutionary biology. Holland [38] used the principle to devise his genetic algorithms for mathematical optimisation. Evolutionary computation has become an intensely researched field [26]. Edelman [20] used the principle in his theory of 'Neural Darwinism', explaining how the neural structure of the brain is constructed and maintained. As mentioned earlier, Dawkins [15] proposed it as the regulating principle of memes. In this section, I will review three applications of natural selection that are relevant to this thesis.

3.2.1 Cosmological natural selection

In Chapter 1, I hypothesised that natural selection is the regulating principle required by Wheeler for the observed physical laws and initial conditions to arise out of the "unpredictable outcome of billions and billions of elementary quantum phenomena". In other words, I am saying that physical law as we know it *evolved* by natural selection i.e. Einstein's theory of general relativity is just one set of the fittest laws that have survived the battle thus far. In order to show that this is not as ludicrous a claim as many will initially think it is, I review an analogous (yet different) hypothesis on how natural selection played a role in observable physics: Smolin's Cosmological Natural Selection.

Q. Smith and L. Smolin independently suggested a mechanism for the evolution of universes by natural selection. Smolin [72] asked the question: "Why are the laws of physics and initial conditions of the universe such that stars exist?". Stars account largely for the variety of phenomena observed in the universe - they synthesise the higher nuclei and keep large regions of the universe far from thermal equilibrium. On the other hand, stars exists due to (at least) two fundamental constants of physics : the ratio of 10^{40} between the gravitational and other interactions of nucleons and the fact that the neutron-proton mass difference is less than nuclear binding energies. Thus a theory that

explains these fundamental constants would explain why stars exist. Smolin proposed the following explanation as to why these constants are at the current observed values.
Firstly, he proposed two postulates regarding spacetime singularities formed by gravitational collapse as predicted by the Einstein field equations:

1 The singularities give birth to new universes.

2 The physical constants, represented by the parameter p_{new} of the new universe are perturbed by a small amount relative to the original constants p_{old}.

Smolin also assumed that the parameters are such that they always cause the new universe to recollapse, thus always ensuring that the new universe has at least one descendent, although I do not find this assumption necessary for the argument. The parameter p influences the number of singularities that a universe can have, and thus the number of descendants - the requirement of differential fitness. The assumption that the physical constants of the descendent universes are small perturbations of the parent universe's constants, gives one the requirement of heritability and variation. Thus one can conclude that natural selection regulates the frequency of these parameters in the space comprising all these universes. It intuitively follows that the parameters that determine the maximum amount of singularities in a given universe will after some time, dominate this space (although time would need to be defined in this context). This then leads one full circle in answering the initial question: "The parameters of particle physics are such that most changes in their values should lead to decreases in the expected number of black holes in the universe." This is the "most intriguing passage" in Lee Smolin's theory [75] - the fact that the hypothesis may be confirmed or refuted through a combination of astrophysical observations and theory. Smolin analyses these parameters [73] and proposes three observational tests [74].

My hypothesis, although similar, is fundamentally different. I propose that the laws of physics governing the universe evolved by natural selection. In Smolin's cosmological natural selection, physical law evolves as the parameters change from one universe to the next, but remains constant within each universe. In my scenario, the physics of the early universe would be different from the physics we observe today. This topic will be expanded upon in Chapter 4. For now, it suffices to say that the application of natural selection to physics is a serious field of study.

3.2.2 The human thought process

Goertzel [29] reintroduced the hypothesis of the evolving mind. The formulation of this hypothesis is that the "mental process is itself a form of evolution: That when you think, remember and feel, the

process going on in your head is actually a process of evolution by natural selection" [29]. This is not a new hypothesis, but can be traced back to Darwin and H. Spencer (1873) in 'The study of Sociology'. In this hypothesis, one views the complex process of thought, in which the total environment (all your current beliefs, feelings, memories etc) determines which new thoughts or beliefs survive to create a new environment. From my perspective, Dawkins' hypothesis incorporates this theory. Whether the meme's environment is a society of many people or a single human brain, the regulating principle should in principle remain identical. Modelling the dynamics of beliefs within a society is thus equivalent to modelling the dynamics of beliefs within an individual. However, the relevant point is that natural selection as the regulating principle of knowledge arises in various contexts.

3.2.3 The clay hypothesis and the origin of life

If one accepts Darwin's theory of evolution, one can conceptually step backwards in time and ask: What was the structure of the early organisms or replicators? Biology teaches us that the nucleic acids (the genotype) are responsible for replication and mutation. The proteins mediate the functions of the cell (the phenotype). This complex yet elegant structure could clearly not be a starting point [13]. Research originally focused on finding simple independently reproducible subsystems constructed using organic material. However in 1966 an alternative suggestion was made by A.G. Cairns-Smith. He hypothesised that the early replicators were constructed from inorganic material and were in fact the ordinary crystals observed in nature. He went on to propose that the crystals of choice for the first genetic material might be non-other than clay minerals [13].

In order to understand this hypothesis, the analogies between the behaviour of crystals and biological organisms need to be clarified. Firstly, crystals replicate in the sense that new crystal layers form from solution onto pre-existing layers. As crystals replicate, environmental conditions can cause changes / mutations in the structure. An important aspect of this is that the mutations in the structure will continue to be replicated thereafter - the requirements of heritability and variation are satisfied. Finally, a mutation in the structure can cause the momentum of the crystallisation process of the solution to increase (or decrease), resulting in more (or less) copies of the mutated crystal per unit time - the property of differential fitness. One can conclude that natural selection regulates the space of crystals.
The structure of an individual crystal is analogous to the genotype of an organism, while the phenotype encompasses properties of the structure which include

- Fecundity Properties: The frequency of the reproduction rate measured.

- Interaction Properties: How the crystal structure interacts within the environment (the surrounding solution which can include other crystal structures, temperature, availability of chemicals/elements in the solution etc) according to currently accepted physical laws (chemistry, quantum mechanics etc)

As candidates for the early replicators, clay minerals have very positive properties. Clay minerals are hydrous aluminosilicates that are characterised by crystal sizes less than 2μm in diameter. Clays have amazingly diverse microstructures comprising a wide variety of crystals [13]. The abundance of clays on earth ($\pm 50\%$ of sedimentary rocks) is a result of synthesis occurring due to the crystallisation of solutions that are in constant supply from rock weathering. The important point is that the process of clay synthesis is a crystallisation process i.e. a self-assembly, and that there is ample of this structure on the earth. To move from clay minerals to the biological replicators of today, one needs to map an evolutionary path from a purely inorganic replicator to the organic structures evident at present. The clay hypothesis states that this transition occurred

- as inorganic structures mutated to allow the addition of organic elements into their structures
- and environmental conditions (which could be as simple as the availability of this organic molecule) gave the structure a competitive advantage over other structures.

How organic elements can change the phenotype of a crystal is evidenced in the following statement by A.G. Cairns [13]: "If you want to control the direction of growth of crystals, one of the standard ways of doing it is by adding organic molecules so that they absorb on certain faces and prevent certain faces growing faster than other ones." Research into the clay hypothesis (see [13]) shows detailed scientific argument substantiating the hypothesis. If we accept this hypothesis, we can see that natural selection is a powerful regulating principle in that it provides an emergent path from a purely physical model (crystals and their interactions according to the governing laws of quantum theory) to the biological model (genes and their phenotypes).

3.3 The mathematics of natural selection

I offer you a choice - either believe that natural selection as the regulating principle in nature is an absolute truth, or abandon mathematics.

This section mathematically formalises the concept of 'natural selection regulating the dynamics of a space'. This results in a proof showing the necessary and sufficient requirements for the process of natural selection to occur. My objective will be achieved by showing that a single equation governs

any system regulated by natural selection. The foundation of this equation was laid in the derivation of the Price equation [61, 62].

3.3.1 The Price Equation

Price [61, 62, 32] derived his equation as follows: Let P_1 and P_2 be populations (sets) of a single species, such that P_1 contains all parents of P_2, and P_2 consists of all the offspring of P_1. Price's objective was to measure how the frequency of a particular trait **T** changes as one moves from the parent to the child population. Assume there are N individuals $\{x_1, x_2, \ldots, x_N\}$ in P_1 and let $q(i)$ represent the frequency of trait **T** in individual x_i. For simplicity purposes, one can think of q as a boolean function $q : P_1 \to \mathbf{2}$ on P_1 in the sense that an individual x_i has the trait (i.e. $q(i) = 1$) or not (i.e. $q(i) = 0$). In Price's original argument, $0 \leq q(i) \leq 1$, but this generality does not affect the essence of the final results I am seeking. Let $z(i)$ be the number of offspring of individual x_i. For simplicity purposes, only consider asexual reproduction. (This will be generalised to sexual and 'polysexual' reproduction in the subsequent section.) Let $g'(i)$ be the number of offspring of individual x_i that have the trait **T** and define $q'(i) = g'(i)/z(i)$. Let $Q_1 = \frac{\sum_{i=1}^{N} q(i)}{N} = \overline{q}$ and Q_2 be the frequency of trait **T** in the parent and child population respectively, where $\overline{q}$ represents the arithmetic mean of the sequence $(q(1), q(2), \ldots, q(N))$. Finally, define $\triangle q(i) = q'(i) - q(i)$. Now Price observed that

$$\begin{aligned} Q_2 &= \frac{\sum_{i=1}^{N} g'(i)}{\sum_{i=1}^{N} z(i)} = \frac{\sum_{i=1}^{N} z(i)q'(i)}{\sum_{i=1}^{N} z(i)} = \frac{\sum_{i=1}^{N} z(i)q'(i)}{N\overline{z}} \\ &= \frac{\sum_{i=1}^{N} z(i)q(i)}{N\overline{z}} + \frac{\sum_{i=1}^{N} z(i)\triangle q(i)}{N\overline{z}} \end{aligned} \tag{3.3.1}$$

The covariance between two sequences $v = (v_1, v_2, \ldots, v_m)$ and $w = (w_1, w_2, \ldots, w_m)$ is defined by

$$\mathrm{Cov}(v, w) = \frac{\sum_{k=1}^{m} v_k w_k}{m} - \overline{vw} \tag{3.3.2}$$

On substitution, Equation 3.3.1 becomes

$$\begin{aligned} Q_2 &= \frac{\overline{zq} + \mathrm{Cov}(z, q)}{\overline{z}} + \frac{\sum_{i=1}^{N} z(i)\triangle q(i)}{N\overline{z}} \\ &= \overline{q} + \frac{\mathrm{Cov}(z, q)}{\overline{z}} + \frac{\sum_{i=1}^{N} z(i)\triangle q(i)}{N\overline{z}} \end{aligned}$$

Price thus arrived at the relation

$$\begin{aligned} \triangle Q &= Q_2 - Q_1 \\ &= \frac{\mathrm{Cov}(z,q)}{\overline{z}} + \frac{\sum_{i=1}^{N} z(i)\triangle q(i)}{N\overline{z}} \\ &= \frac{\rho(z,q)\sigma(q)\sigma(z)}{\overline{z}} + \frac{\sum_{i=1}^{N} z(i)\triangle q(i)}{N\overline{z}} \end{aligned} \tag{3.3.3}$$

where $\rho(z,q)$ is the Pearson correlation coefficient and $\sigma(q)$,$\sigma(z)$ are the variances of the series q, z respectively. I have used the statistical relation

$$\mathrm{Cov}(v,w) = \rho(v,w)\sigma(v)\sigma(w) \tag{3.3.4}$$

for any series v, w. Price's derivation showed that the change in frequency of a trait **T** from the parent to the child population is dependent on the *variance* of the trait **T** in the parent population (i.e. the $\sigma(q)$ term) and $\rho(z,q)$, the correlation between a parent having the trait and the amount of offspring (i.e. *Differential Fitness*). Price however stopped at this point, arguing that the $\frac{\sum_{i=1}^{N} z(i)\triangle q(i)}{N\overline{z}}$ term would be insignificant under certain conditions, and only focused on the covariance term. I will now reconsider this argument and show how all of Darwin's requirements for natural selection can be expressed in a single equation.

3.3.2 The equation of natural selection

Evolutionary theory [47] states that in order for natural selection to regulate the frequency of a characteristic (trait) in a population of individuals, one requires:

I : Heritability - This characteristic must be inheritable.

II : Differential Fitness - This trait must affect the reproduction and survival rates of the individuals, i.e. their fitness.

III: Phenotypic Variation - Different individuals must have varying traits that result in different morphologies, physiologies and behaviours.

The biological sciences assume that if these principles hold, a population will undergo evolutionary change. I will proceed to formally prove that these requirements are necessary but not sufficient for natural selection to occur. In order to proceed, I need to define the space wherein natural selection occurs, and what it is that natural selection acts on. In the biological sciences, Dawkins [16] defines the unit of selection *"as the entity for whose benefit adaptations may be said to exist."* In Darwin's

seminal work [14], natural selection acts upon individuals. Competing theories proposing gene selection [15], individual selection, group selection, whole environments and combinations thereof [47] have been advocated. In fact, the question *"What is the unit of selection?"* encapsulates a central theoretical problem of evolutionary biology [16]. This question falls outside the scope of this thesis due to the fact that the principle of natural selection is independent of the unit of selection. In particular, Lewontin [47] argues how the principles of natural selection can be applied to molecules, cells, populations, ... in fact, any set of entities that have the variation, reproduction and heritability properties mentioned above.

I consider a finite set $\mathcal{W}$ that comprise the units of selection. I follow Mayr's terminology and call the elements of $\mathcal{W}$ *selectons* [16]. What these selectons actually represent is determined by the context. Thus in the biological case, if one assumes a particular gene sequence determines an observable phenotype, then a selecton $w \in \mathcal{W}$ could actually be a sequence of genes. In Smolin's Cosmological Natural Selection [72], a selecton would be a universal constant that together with other constants determine the physics of a universe.
I define the set of replicators $\mathcal{R} = \mathcal{P}(\mathcal{W})$ as the power set of the units of selection. I am thus assuming that a set of selectons completely determines a replicator. $\mathcal{R}$ will represent the complete set of replicators with lower case Greek letters $\phi, \psi, \ldots$ denoting individual replicators.
Finally, I define a binary relation S over $\mathcal{R}$ that will model the replication process. The idea is that for any 2 replicators $\phi, \psi \in \mathcal{R}$, $\phi S \psi$ means ψ is a successor (child) of predecessor (parent) ϕ. The relation S gives rise to the set of all successors $S(\phi) = \{\psi \mid \phi S \psi\}$ and predecessors $S^{\smile}(\psi) = \{\phi \mid \phi S \psi\}$ of replicators ϕ and ψ respectively.

Consider $H, H' \subset \mathcal{R}$ satisfying $\bigcup_{\phi \in H} S(\phi) = H'$ and $\bigcup_{\psi \in H'} S^{\smile}(\psi) \subseteq H$. These constraints state that H' comprises all the successors of H and H contains all predecessors of H'. One doesn't necessarily have equality in the latter case since some replicators in H might have no successors.
$C(\psi) = \#S^{\smile}(\psi)$ represents the number of predecessors of replicator ψ. The function

$$z(\phi) = \sum_{\psi \in S(\phi)} \frac{1}{C(\psi)} \tag{3.3.5}$$

designates the amount of successors of a replicator ϕ. The assumption here is that all predecessors of a replicator contribute equally to the successor population. Thus if a successor has m predecessors, then each predecessor will contribute $\frac{1}{m}$ replicators toward the successor population.

The frequency of a selecton w in replicator ϕ is defined by the function

$$q(\phi, w) = \begin{cases} 1 & \text{if } w \in \phi \\ 0 & \text{otherwise} \end{cases} \tag{3.3.6}$$

For biologists, it is important to note that since w can only have a frequency 0 or 1 in a replicator, the biological example of alleles will require the use of w_0 and w_1 as two copies of the gene with the allele on in one and off in the other.

The weighted selecton frequency $q'(\phi, w)$ of w in all successors of ϕ is given by

$$z(\phi)q'(\phi, w) = \sum_{\psi \in S(\phi)} \frac{q(\psi, w)}{C(\psi)} \tag{3.3.7}$$

In the case where $C(\psi)$ is a constant for all replicators ψ (e.g. sexual reproduction in natural systems), this reduces to the usual definition of selecton frequency in a population i.e. $q'(\phi, w) = \frac{\sum_{\psi \in S(\phi)} q(\psi, w)}{\#S(\phi)}$. Finally, define the frequency of a selecton w in the sets H and H' as $Q(H, w) = \frac{\sum_{\phi \in H} q(\phi, w)}{\#H}$ and $Q(H', w) = \frac{\sum_{\psi \in H'} q(\psi, w)}{\#H'}$ respectively.

My objective is to represent $Q(H', w)$ in terms of $Q(H, w)$. I am interested in what aspects affect the change of frequencies of a selecton in the populations. Towards this end, I define elements $\psi_1, \psi_2 \in H'$ as siblings $\psi_1 \sim \psi_2$ iff $S^\smile(\psi_1) = S^\smile(\psi_2)$ i.e. they have the same set of predecessors. This relation is transitive ($\psi_1 \sim \psi_2$ and $\psi_2 \sim \psi_3 \rightarrow \psi_1 \sim \psi_3$), symmetric ($\psi_1 \sim \psi_2 \rightarrow \psi_2 \sim \psi_1$) and reflexive ($\psi_1 \sim \psi_1$), and hence an equivalence relation on H'. One can thus uniquely partition H' into the disjoint union of the equivalence classes of siblings : $H' = \bigcup_{i=1}^{n} \Psi_i$ with each Ψ_i containing the set of siblings with common predecessors.

Since $(\psi_1 \sim \psi_2) \Rightarrow (C(\psi_1) = C(\psi_2))$ one can define $C(\Psi_i) = C(\psi)$ for any $\psi \in \Psi_i$.

Now any function $f : \mathcal{R} \times \mathcal{W} \rightarrow \mathbb{R}$ satisfies the property

$$\begin{aligned} \sum_{\phi \in H} \sum_{\psi \in S(\phi)} f(\psi, w) &= \sum_{i=1}^{n} \sum_{\phi \in H} \left(\sum_{\psi \in S(\phi) \bigcap \Psi_i} f(\psi, w) \right) \\ &= \sum_{i=1}^{n} C(\Psi_i) \sum_{\psi \in \Psi_i} f(\psi, w) \end{aligned} \tag{3.3.8}$$

One obtains this by noting that if $\phi_1, \phi_2 \in H$ is such that $\phi_1, \phi_2 \in S^\smile(\Psi_i) = \bigcup_{\psi \in \Psi_i} S^\smile(\psi)$ then $S(\phi_1) \bigcap \Psi_i = S(\phi_2) \bigcap \Psi_i = \Psi_i$. Further, the sum is over exactly $C(\Psi_i)$ copies of this same set. Using 3.3.8, I have

$$\sum_{\phi \in H} z(\phi) = \sum_{\phi \in H} \sum_{\psi \in S(\phi)} \frac{1}{C(\psi)} = \#H' \tag{3.3.9}$$

Similarly

$$\sum_{\phi\in H} z(\phi)q'(\phi,w) \quad = \quad \sum_{\phi\in H}\sum_{\psi\in S(\phi)} \frac{q(\psi,w)}{C(\psi)} = \sum_{\phi\in H'} q(\phi,w) \tag{3.3.10}$$

The selecton frequency in the successor set H' is thus given by

$$Q(H',w) = \frac{\sum\limits_{\psi\in H'} q(\psi,w)}{\#H'} = \frac{\sum\limits_{\phi\in H} q'(\phi,w)z(\phi)}{\sum\limits_{\phi\in H} z(\phi)} \tag{3.3.11}$$

I first consider the case of asexual reproduction only. I thus constrain the generality by assuming that $\forall\psi \in H'$ one has $C(\psi) = 1$ i.e. all successors have one predecessor. Decomposing H into the 2 disjoint subsets $H_0 = \{\phi \in H \mid q(\phi,w) = 0\}$ and $H_1 = \{\psi \in H \mid q(\psi,w) = 1\}$ whose union is H and observing that $\sum\limits_{\phi\in H_0} f(\phi,w) = \sum\limits_{\phi\in H} [1-q(\phi,w)]\, f(\phi,w)$ and $\sum\limits_{\phi\in H_1} f(\phi,w) = \sum\limits_{\phi\in H} q(\phi,w)f(\phi,w)$ is true for any function $f : \mathcal{R}\times\mathcal{W} \to \mathbb{R}$ allows me to rewrite 3.3.11 as:

$$\begin{aligned}
Q(H',w) &= \frac{\sum\limits_{\phi\in H_0} z(\phi)q'(\phi,w) + \sum\limits_{\phi\in H_1} z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H} z(\phi)} \\
&= \left(\frac{\sum\limits_{\phi\in H_0} z(\phi)}{\sum\limits_{\phi\in H} z(\phi)}\right)\left(\frac{\sum\limits_{\phi\in H_0} z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H_0} z(\phi)}\right) + \left(\frac{\sum\limits_{\phi\in H_1} z(\phi)}{\sum\limits_{\phi\in H} z(\phi)}\right)\left(\frac{\sum\limits_{\phi\in H_1} z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H_1} z(\phi)}\right) \\
&= \left(\frac{\sum\limits_{\phi\in H} [1-q(\phi,w)]\, z(\phi)}{\sum\limits_{\phi\in H} z(\phi)}\right)\left(\frac{\sum\limits_{\phi\in H} [1-q(\phi,w)]\, z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H} [1-q(\phi,w)]\, z(\phi)}\right) \\
&+ \left(\frac{\sum\limits_{\phi\in H} q(\phi,w)z(\phi)}{\sum\limits_{\phi\in H} z(\phi)}\right)\left(\frac{\sum\limits_{\phi\in H} q(\phi,w)z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H} q(\phi,w)z(\phi)}\right)
\end{aligned} \tag{3.3.12}$$

Now note that the term $\frac{\sum\limits_{\phi\in H}[1-q(\phi,w)]z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H}[1-q(\phi,w)]z(\phi)}$ in the equation above calculates the ratio of the number of replicators in H' containing w (the $[z(\phi)q'(\phi,w)]$ term in the numerator) that have predecessors not containing w (the $[1-q(\phi,w)]$ term in the numerator) to the number of replicators in H' that have predecessors that don't contain w. If I define $p_m(w)$ to be the probability that an arbitrary successor ψ with predecessors $S^{\smile}(\psi) = \{\phi_1,\phi_2,\ldots,\phi_C\}$ will have the selecton w given that none of its predecessors in $S^{\smile}(\psi)$ have the selecton, then I can deduce for the case of asexual

reproduction that

$$p_m(w) = \mathcal{E}\left[\frac{\sum_{\phi\in H}\left[1-q(\phi,w)\right]z(\phi)q'(\phi,w)}{\sum_{\phi\in H}\left[1-q(\phi,w)\right]z(\phi)}\right] \tag{3.3.13}$$

where $\mathcal{E}$ represents the expectation value.

In order to analyse the term $\frac{\sum_{\phi\in H} q(\phi,w)z(\phi)q'(\phi,w)}{\sum_{\phi\in H} q(\phi,w)z(\phi)}$, I define the probability of inheritance $p_i(w)$ to be the probability that an arbitrary successor ψ with predecessors $S^{\smile}(\psi) = \{\phi_1, \phi_2, \ldots, \phi_C\}$ *has the selecton* w given that one of its predecessors in $S^{\smile}(\psi)$ has the selecton i.e. conditional probability that a selecton w occurs in a successor ψ given that it occurred in at least one of its predecessors $\phi \in S^{\smile}(\psi)$. This happens due to mutation and inheritance and is thus the probability that the selecton mutated $[p_m(w)]$ or that the selecton was inherited $[p_i(w)]$ and no mutation occurred $[1-p_m(w)]$.

$$p_i(w)\left[1-p_m(w)\right] + p_m(w) = \mathcal{E}\left[\frac{\sum_{\phi\in H} q(\phi,w)z(\phi)q'(\phi,w)}{\sum_{\phi\in H} q(\phi,w)z(\phi)}\right] \tag{3.3.14}$$

Finally, using an identical observation to that of Price in his formulation of the Price Equation [61, 62], I have

$$\frac{\sum_{\phi\in H} q(\phi,w)z(\phi)}{\sum_{\phi\in H} z(\phi)} = \rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}} + Q(H,w) \tag{3.3.15}$$

Here $\rho(q,z) = \frac{\#H\sum_{\phi\in H} q(\phi,w)z(\phi) - \sum_{\phi\in H} q(\phi,w)\sum_{\phi\in H} z(\phi)}{\sqrt{\left[(\#H)\sum_{\phi\in H} q(\phi,w)^2 - \left(\sum_{\phi\in H} q(\phi,w)\right)^2\right]\left[(\#H)\sum_{\phi\in H} z(\phi)^2 - \left(\sum_{\phi\in H} z(\phi)\right)^2\right]}}$ is the Pearson Correlation Coefficient, $\sigma(q) = \sqrt{\frac{(\#H)\sum_{\phi\in H} q(\phi,w)^2 - \left(\sum_{\phi\in H} q(\phi,w)\right)^2}{(\#H)^2}}$ is the population standard deviation and $\bar{q} = \frac{\sum_{\phi\in H} q(\phi,w)}{\#H}$ is the population mean. I substitute Equations 3.3.13, 3.3.14 and 3.3.15 into Equation 3.3.12 and arrive at the Equation of Natural Selection for the case where all successors have one and only one predecessor:

$$\frac{Q(H',w) - p_m(w)}{(1-p_m(w))} = p_i(w)\left(\rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}} + Q(H,w)\right) \tag{3.3.16}$$

First note that Equations 3.3.14 and 3.3.13 implicitly assume normal distributions for $p_i(w)$ and $p_m(w)$. (The validity of these equations is numerically illustrated in the appendix.)

Now Equation 3.3.16 describes how the frequency of a particular selecton changes over time. Since natural selection regulates this process, I can conclude that any asexual reproductive system regulated by natural selection will obey the above equation. The reader should note that the terms

$p_i(w)$, $p_m(w)$, $\sigma(q)$ and $\rho(q,z)$ are the mathematical analogues of heritability, variation and differential fitness respectively.

Now I consider the more general case where $\forall\psi \in H'$, $C(\psi) = C$ (i.e. all successors have *exactly* C predecessors). For this case, I use the intuitiveness of Equation 3.3.16. Define $p_w(H)$ to be the probability that an arbitrary successor $\psi \in H'$ has a parent in H containing the selecton under consideration. Using this definition together with the definition of $p_m(w)$ and $p_i(w)$, I can say that the frequency distribution $Q(H', w)$ for replicators constrained to having only one predecessor is given by the probability that the selecton mutated *or* that it did not mutate *and* was inherited from a parent having the selecton; i.e.

$$Q(H', w) = [1 - p_m(w)]\, p_i(w) p_w(H) + p_m(w) \tag{3.3.17}$$

This is exactly of the form of Equation 3.3.16 if I define $p_w(H) := \rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}} + Q(H,w)$. This states that the probability of a parent having a selecton is determined by the frequency of the selecton in the parent population and the covariance term specified by Price [61].
I now have the tools required to derive the form of Equation 3.3.16 for the case of constant C parents. Consider any successor ψ with C predecessors. The probability that a predecessor does not have the selecton is given by $1 - p_w(H)$. The probability that all its predecessors don't have the selecton is given by $[1 - p_w(H)]^C$. The probability that at least one of ψ's predecessors has the selecton is given by $1 - [1 - p_w(H)]^C$. Thus the probability that ψ has the selecton that was inherited with no mutation occurring is given by $[1 - p_m(w)]\, p_i \left[1 - [1 - p_w(H)]^C\right]$ The probability the ψ has the selecton due to mutation is given by $p_m(w)$. Thus

$$Q(H', w) = [1 - p_m(w)]\, p_i(w) \left[1 - [1 - p_w(H)]^C\right] + p_m(w) \tag{3.3.18}$$

The generalised form of the Equation 3.3.16 for replicators with exactly C predecessors is thus given by:

$$\frac{Q(H', w) - p_m(w)}{(1 - p_m(w))} = p_i(w) \left(1 - \left[1 - \rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}} - Q(H,w)\right]^C\right) \tag{3.3.19}$$

One could arrive at Equation 3.3.19 using an identical argument as in the asexual case, except that Equation 3.3.13 is replaced by:

$$p_m(w) + p_i(w)(1 \quad - \quad p_m(w))(1 - \left[1 - Q(H,w) - \rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}}\right]^{C-1}) \\ = \quad \mathcal{E}\left[\frac{\sum\limits_{\phi \in H}\left[1 - q(\phi,w)\right] z(\phi) q'(\phi,w)}{\sum\limits_{\phi \in H}\left[1 - q(\phi,w)\right] z(\phi)}\right] \tag{3.3.20}$$

I conclude by deriving the equation of natural selection for the most general case - that of multiple non-constant parents i.e. $\forall \psi \in H', C(\psi) \geq 1$. To achieve this, define $p_C(H)$ as the probability that an arbitrary successor in $\psi \in H'$ has exactly C predecessors in H. Now if I consider any successor $\psi \in H'$, I can deduce that the probability that one of its predecessors has the selecton is given by the weighted sum $\sum\limits_{C=1}^{\#H} p_C(H)\left(1 - \left[1 - p_w(H)\right]^C\right)$ where $\sum\limits_{C=1}^{\#H} p_C(H) = 1$. Using the identical argument implemented in the derivation of 3.3.19, I can conclude that the most general form of the equation is given by:

$$\frac{Q(H',w) - p_m(w)}{(1 - p_m(w))} = p_i(w) \sum_{C=1}^{\#H} p_C(H)\left(1 - \left[1 - \rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}} - Q(H,w)\right]^C\right) \tag{3.3.21}$$

I conclude that any reproductive system regulated by natural selection will obey Equation 3.3.21.

For the interested (or sceptical) reader, the appendix contains numerical simulations confirming Equations 3.3.20 and 3.3.14, relating the probability of mutation and inheritance with the other variables of the derivation.

3.3.3 The mathematical requirements of natural selection

The derived equations provide me with the tools to formalise the necessary and sufficient requirements for natural selection to regulate (and *continue* regulating) the selecton frequency within a space. I proceed by proving necessity. In the case where $p_i(w)$ and $p_m(w)$ are constant for all generations, and all replicators have the same number of parents, I will show that the requirements of

I Heritability, as specified by $p_i(w) > 0$

II Differential fitness, as specified by $\rho(q,z) \neq 0$

III Variation, as specified by $\sigma(q) > 0$

are necessary conditions for natural selection to regulate the frequency of a selecton w. I proceed by proving the contrapositive. Consider any selecton w.

I': The requirement of *Heritability* is embodied in $p_i(w)$ term. Setting $p_i(w) = 0$ would result in $Q(H', w) = p_m(w)$ for this and all subsequent generations. The change in the selecton frequency of w between parent and child populations would thus have a constant expectation value of 0. The actual change in frequency of the selecton w between generations would all be due to a normal random distribution and thus one can conclude that natural selection does not regulate the frequency of this selecton

II': The requirement of *Differential Fitness* is encapsulated in the $\rho(q, z)$ term. Setting $\rho(q, z) = 0$ results in the frequency of w in subsequent generations converging rapidly to the following fixed points, for the cases of asexual and sexual reproduction specified below:

$$Q(H^{\infty}, w) = \frac{p_m}{1 - p_i + p_m p_i} \text{ for } c = 1 \tag{3.3.22}$$

$$Q(H^{\infty}, w) = 1 + \frac{\sqrt{4(1-p_m)^2 p_i^2 - 4(1-p_m)^2 p_i + 1} - 1}{2p_i(1-p_m)} \text{ for } c = 2 \tag{3.3.23}$$

For arbitrary C, one solves a polynomial equation of degree C. I deduce that the selecton frequency change between parent and child populations will tend towards a constant expectation value of 0 resulting once again in the frequency being determined solely by random variables. I can conclude that after a few generations, natural selection will not regulate the frequency of this selecton.

III': The requirement of *Variation* is determined by all three terms $p_m(w)$, $p_i(w)$ and $\sigma(q)$. Setting $p_m(w) = 0$, $p_i(w) = 1$ and $\sigma(q) = 0$ will imply no variation in this and subsequent generations. This case is handled in [II']. (Note that if one of the conditions $p_m(w) > 0$, $p_i(w) < 1$ or $\sigma(q) > 0$ hold, then variation will persist, even if $\sigma(q) = 0$ is true in any one generation.)

This establishes that the necessary requirements for natural selection are exactly those specified by Darwin, for the case where $p_i(w)$ and $p_m(w)$ are constant through all generations. Now one can argue that $p_i(w)$ should be constant for all generations - the probability that a successor will inherit a selecton from one of its parents only depends on the replication process. However, this is not necessarily true for $p_m(w)$, since evolution occurs through the process of cumulative selection [15]. To clarify this, first note that a selecton can be viewed as the phenotype of some gene complex. Consider the following toy model: Assume there exists a species A that has no method of viewing light, a species B that can see in black and white and a species C that can see in colour. To 'mutate'

from B to C is far more likely that to mutate from A to C. Thus the probability of mutation of the selecton 'view light in colour' changes depending on its environment i.e. the other selections in the replicator. Thus one can argue that the proof in [II'] above is flawed for this case. However, if the mutation probability is very small, then one can still say that the average expectation value of change in [II'] is zero, thus retaining the validity of the argument.

For the case of sufficiency, one notes that Equation 3.3.19 has infinitely many fixed points that satisfy the three requirements. For the case $c = 1$, these fixed point are specified by:

$$Q(H^{\infty}, w) = \frac{p_i \left[1 - p_m(w)\right] \rho(q, z)\sigma(q)\frac{\sigma(z)}{\bar{z}} + p_m(w)}{1 - p_i(w)(1 - p_m(w))} \tag{3.3.24}$$

For the other cases (e.g. $C > 1$), one solves a polynomial equation of degree C. (These fixed points are called *evolutionary stable systems* in biology.) Now it is important to note that these fixed points exist only if the replication relation S (and hence the term $\rho(q, z)$) is a deterministic function of the space of replicators i.e. the correlation with fitness depended only upon the characteristics of all the other replicators in the population. The reason that natural systems do not cease to evolve is due to the fact that replicators are linked to non-deterministic chaotic systems e.g. the weather. If S was a non-deterministic function, then Equation 3.3.19 would have no fixed points that satisfied the three criteria specified by Darwin. I thus formulate the fourth criterion that has always been implicitly assumed in evolutionary biology:

IV: Non-Deterministic Replication - The reproduction and survival rates of individuals is a non-deterministic process.

As an aside, note that the derivation of Equation 3.3.19 would proceed identically as shown if one had defined the function q by $q(\phi, U) = \begin{cases} 1 & \text{if } U \subseteq \phi \\ 0 & \text{otherwise} \end{cases}$ where $U \subseteq \mathcal{W}$ is a set of selectons. This shows that natural selection can operate on individual selectons, sets of selectons, individuals or groups as argued by Lewontin[47].

I will conclude this section by emphasising an important point. It is assumed that one cannot say that natural selection is the regulator of the selecton frequencies if the frequency is determined solely by statistical distributions i.e. the expectation value of the frequency remains constant through all generations. Now if $p_m(w) \geq p_i(w)$, one cannot say that natural selection is the dominant regulating principle, even if the expectation value changed with time. Randomness would play a greater role in determining the frequency. Thus the requirement of heritability should be stated as

$$p_i(w) - p_m(w) > 0 \tag{3.3.25}$$

.

3.4 The evolution of sex

This section will detail an application of the equation of natural selection to a topic outside the scope of this thesis and should be viewed as an (interesting) aside to the main objective. The topic of interest is the evolution of sex. Can Equation 3.3.19 facilitate in the explanation of why sex evolved? Biology has currently produced no satisfactory answer to this question. Sex is costly. Sex implies recombination of gene complexes, resulting in the break up of *favourable* gene complexes. Mathematically, this is explained by using the probability of inheritance. In the asexual scenario, if a replicator has a particular selecton, then $p_i \approx 1$. In the sexual scenario, if a replicator has a selecton, then $p_i \approx \frac{1}{2}$. Thus in the sexual case the species must be fit enough to survive *and* reproduce, but then give up half of that for its partners half. Further, asexual species do not have to invest energy in finding a mate. On the other hand, asexual species have a higher rate of extinction relative to sexual species. Various arguments have been put forth to explain this. Group selectionists argue that sex increases a groups ability to respond to a changing environment and is thus selected for. Other postulates include that sex is advantageous when rapid evolution is necessary. There are various other arguments as to why sex is costly as well as theoretical arguments that attempt to explain why sex evolved [80]. I will attempt to explain why sex evolved using the natural selection equation derived in the previous section.

In order to analyse the problem statement, I investigate the relationship between the change in frequency of a selecton and its correlation with fitness. The frequency of a selecton in an asexual population is governed by Equation 3.3.16. Defining $\Delta Q(w) = Q(H', w) - Q(H, w)$, I rewrite Equation 3.3.16 as

$$\Delta Q(w) = p_i(w)\left[1 - p_m(w)\right]\left[\frac{\rho(q,z)\sigma(q)\sigma(z)}{\overline{z}} + Q(H,w)\right] \quad + \quad p_m(w) - Q(H,w) \tag{3.4.1}$$

For what values of $\rho(q, z)$ does the selecton frequency increase (or stay constant) from one generation to the next i.e. $\Delta Q(w) \geq 0$? I solve for $\rho(q, z)$ in Equation 3.4.1 and arrive at:

$$\rho(q,z) \geq \frac{\overline{z}}{\sigma(q)\sigma(z)}\left[\frac{Q(H,w) - p_m(w)}{p_i(w)\left(1 - p_m(w)\right)} - Q(H,w)\right] \tag{3.4.2}$$

Figure 3.1 below shows a plot of the simulated results for Equation 3.4.2 [1] The x axis represents the initial selecton frequency. The y axis represents the minimum value the correlation coefficient $\rho(q,z)$ can have to still ensure that the selecton frequency does not decrease. The results are plotted for 2 values of $p_m(w) \in \{\frac{1}{100}, \frac{5}{100}\}$ with the other simulation parameters (see appendix) set to: $n_{\text{Predecessors}} = 1000$, $n_{\text{Successors}} = 2000$, $C = 1$, $\text{Parent}_{\text{Constant}} = \text{True}$, $p_i = 0.95$.

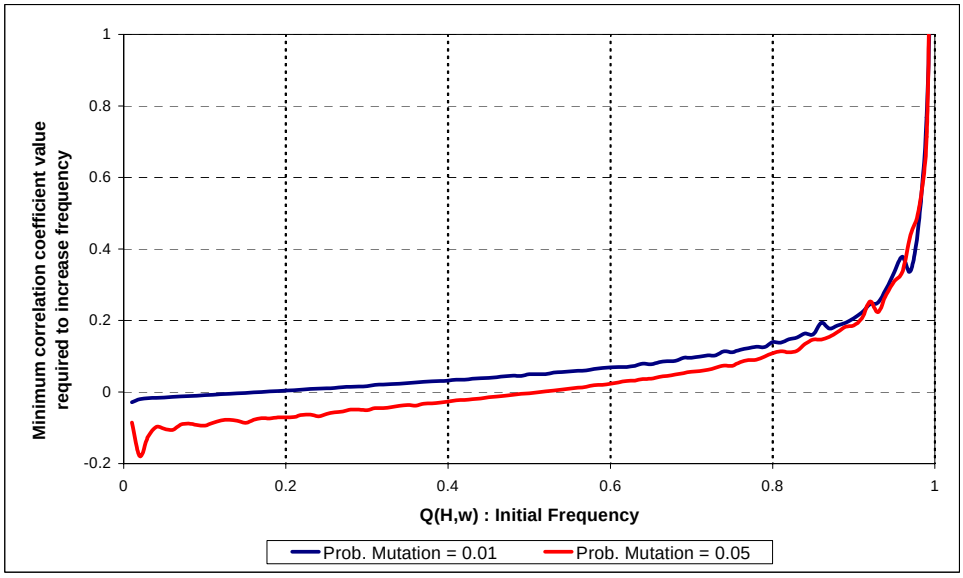

Figure 3.1: The simulated minimum value of the correlation coefficient $\rho_{\min}(q,z)$ that will guarantee an increase in the selecton frequency

The results are very intuitive: The frequency in most cases will only increase if the correlation coefficient $\rho(q,z) > 0$. In the case where $p_m = 0.05$, one sees that the selecton will remain and increase in the population even if one has a slightly negative correlation associated with this selecton i.e. the selecton detrimentally influences fecundity. This is because the mutation rate p_m will replenish the selectons (that are selected out) in the population.

I repeat the process for sexual reproducing species. Using Equation 3.3.18 with $C = 2$ one derives:

$$\rho(q,z) \leq \frac{\overline{z}}{\sigma(q)\sigma(z)} \left(1 - Q(H,w) + \sqrt{1 - \left[\frac{Q(H,w) - p_m(w)}{p_i(w)\,(1 - p_m(w))}\right]}\right) \tag{3.4.3}$$

$$\rho(q,z) \geq \frac{\overline{z}}{\sigma(q)\sigma(z)} \left(1 - Q(H,w) - \sqrt{1 - \left[\frac{Q(H,w) - p_m(w)}{p_i(w)\,(1 - p_m(w))}\right]}\right) \tag{3.4.4}$$

[1] I use the simulation program described in Appendix A to plot this relation for various fixed values of $p_m(w)$ and $p_i(w)$.

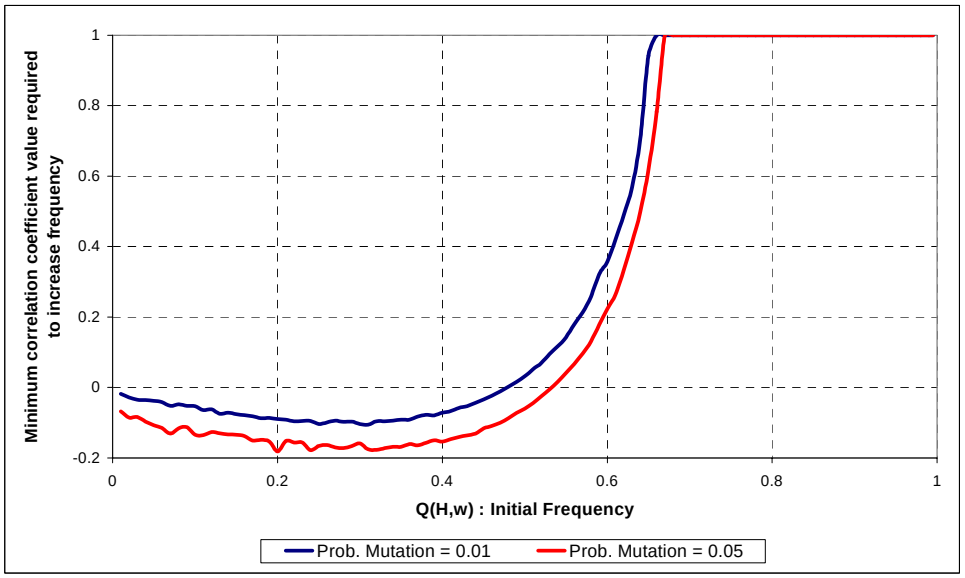

Figure 3.2: The simulated minimum value of the correlation coefficient $\rho_{\min}(q, z)$ that will guarantee an increase in the selecton frequency for sexual reproduction

Figure 3.2 shows the simulated results for the same 2 values of $p_m(w) \in \{\frac{1}{100}, \frac{5}{100}\}$, and with the other parameters being identical to that of the asexual case : $n_{\text{Predecessors}} = 1000$, $n_{\text{Successors}} = 2000$, $\text{Parent}_{\text{Constant}}$ = True; except that $C = 2$ and $p_i = 0.65$. (The justification for dropping $p_i(w)$ down from 0.95 in the asexual case to 0.65 in this sexual case is given as follows: p_i represent the probability of inheritance given that at *least* one parent has the selecton. Now if both parents have the selecton, the probability that the successor will have the selecton should be equal to that of the asexual case i.e. 0.95. If one parent has the selecton, one can argue that the probability drops by a factor of 2. i.e. 0.475. The the probability of inheritance given that at least one of the predecessors has the selecton should be somewhere between these two values. I chose $p_i = 0.65$ which is less than the average, so as not to distort the comparison I am about to make.) Now compare the results for $p_m = 0.01$ for sexual vs asexual reproduction.

To interpret the results, note that complexity in species arises due to cumulative selection - not understanding this concept led to Sir Fred Hoyle's memorable misunderstanding of the theory of natural selection [17]. Hoyle compared natural selection to a hurricane blowing through a junkyard and chancing to assemble a Boeing 747. Cumulative selection entails small incremental changes in every generation that leads to the complexity observable today. Consider the following toy model. One has three traits $\mathbf{T}_1$, $\mathbf{T}_2$, $\mathbf{T}_3$ with complexity increasing from $\mathbf{T}_1$ to $\mathbf{T}_2$ to $\mathbf{T}_3$. Further, assume $\mathbf{T}_1$ is required for $\mathbf{T}_2$ to evolve, and $\mathbf{T}_2$ is required for $\mathbf{T}_3$ to evolve. Now assume $\mathbf{T}_1$ is beneficial

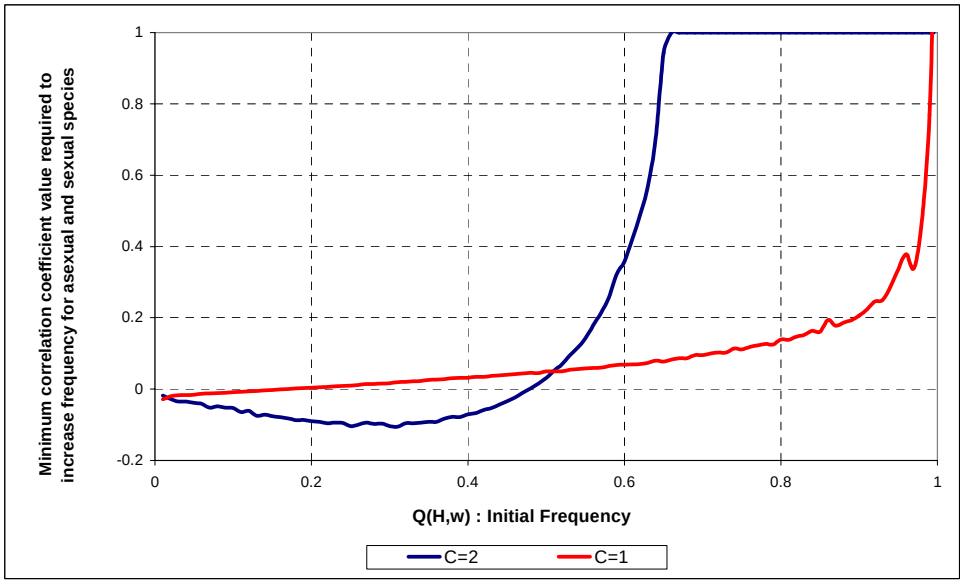

Figure 3.3: The simulated minimum value of the correlation coefficient $\rho_{\min}(q, z)$ that will guarantee an increase in the selecton frequency for sexual and asexual reproduction

to a species fecundity, $\mathbf{T}_2$ is detrimental and $\mathbf{T}_3$ is beneficial again. The analysis above shows that $\mathbf{T}_2$ will always be selected out in asexual reproducing populations. However, in sexual reproducing species, it can still remain and increase in the population even though it is detrimental to fecundity. This will allow $\mathbf{T}_3$ to evolve in sexual species, but not in asexual species. This is why sexual species are observed to be more complex than asexual species. Now use this argument for the characteristic of sexual reproduction itself. Sexual reproduction is a trait that could have initially been detrimental to the species. However, it was allowed to remain in the species for long enough to allow complexity to arise, resulting in the trait of sexual reproduction being beneficial, as is observed today.

3.5 Formal structures regulated by natural selection

The most important assumption in the derivation of the equation of natural selection was that the entities in the space replicate and that they are defined by a set of selectons. The notion of replication comes naturally in logicatoms, since one can view the relation as determining the next set of parents. Thus, to complete the construction objective, I need to define what the units of selection are in logicatoms. Subsection 3.5.1 describes how I represent the 'selectons' of a particular class of logicatoms. This representation allows me to formally state the hypothesis that needs to be proven to complete my construction objective. Subsection 3.5.2 details the search for theory networks

regulated by natural selection. Finally, Subsection 3.5.3 shows how the construction objective is achieved using 'inverted theory networks'.

3.5.1 The logicatom's selectons

Brink and Rewitzky [8] emphasised the paradigm triangle between a logic, its algebra and the semantics, unifying the various ways of looking at the same reality. This paradigm is very useful when attempting to analytically simulate structures built using propositional logic, such as the proposed theory networks.[2] The semantic representation also fulfills the requirement of representing the 'selectons' of logicatoms. As mentioned in Section 2.1, the semantics of propositional logic, known as possible world semantics, uses the idea that any proposition can be represented by the set of worlds in which it is true. Soundness and completeness of propositional calculus imply that two propositions will be logically equivalent iff they are true in the same set of worlds. A world w is just a valuation $w : \mathit{Form}(\Phi) \to \mathbf{2} = \{0, 1\}$ defined inductively on the structure of $\phi \in \mathit{Form}(\Phi_n)$. In terms of notation, I will denote the set of all possible valuations by W. Given a proposition $\psi \in \mathit{Form}(\Phi_n)$, the set $[\psi]$, called the meaning of ϕ contains all worlds in which ψ is true. i.e. $[\psi] = \{w \in W \mid w(\psi) = 1\}$. The propositional tautology $\top$ is then represented by the whole set W, while the contradiction $\bot$ is represented by the empty set $\emptyset$.

Example 3.5.1. In $\mathrm{PropCal}_3$, one has 2^3 possible worlds depicted in the table below.

p_1	p_2	p_3	**Valuation**
0	0	0	w_0
1	0	0	w_1
0	1	0	w_2
1	1	0	w_3
0	0	1	w_4
1	0	1	w_5
0	1	1	w_6
1	1	1	w_7

The table shows the mapping $w_i : \Phi_3 \to \mathbf{2}$ which extends to a valuation on the set of all formulae $\mathit{Form}(\Phi_3)$ using the following iterative definitions:

$$\begin{aligned} w(\neg\phi) &= W - w(\phi) \\ w(\phi \vee \psi) &= w(\phi) \cup w(\psi) \\ w(\phi \wedge \psi) &= w(\phi) \cap w(\psi) \end{aligned}$$

All propositions can be represented by the set of worlds in which they are true.
Thus $[p_1] = \{w_1, w_3, w_5, w_7\}$ and $[p_1 \wedge p_2] = \{w_3, w_7\}$.

[2] See Appendix A

One now has the tools to display a more intuitive representation of logicatom universes from a 'natural selection' perspective. In particular, I identify the selectons of a logicatom as being the worlds comprising the belief. To see this, consider the following example of a logicatom universe $\mathcal{U}$ over PropCal_3.

Example 3.5.2. Depicted in Figure 3.4 below is a logicatom universe over PropCal_3 where we denote the beliefs using well formed formulae, as has been the representation up to now. The relation generating function $R_f : Form(\Phi_3) \to \mathcal{P}(\Phi_3)$ is defined by $R_f(\phi) = \{p_i \in \Phi_3 \mid \vdash \phi \to p_i\}$.

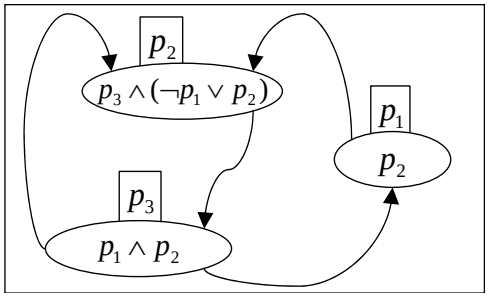

Figure 3.4: Representing the beliefs using propositions

The semantic representation of the propositions allows one to depict the beliefs of the logicatoms as in Figure 3.5 below:

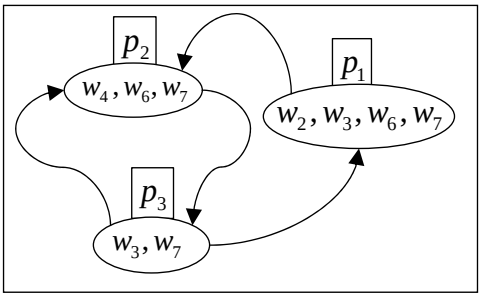

Figure 3.5: Representing the beliefs using world semantics

From this perspective, a world is the unit of selection. A set of worlds determines the belief of a logicatom. Thus one analogises 'genes' with 'worlds'. It is important to note that as in the case of genes in biology, one can have the properties of pleiotropy (a single world can simultaneously affect relations with several logicatoms) and polygeny (a relation between two logicatoms may be determined by multiple worlds) depending on the choice of the relation generating function. An example will clarify this point.

Example 3.5.3. Consider the logicatom named p_3 in the logicatom universe defined in Example 3.5.2. Its belief is semantically represented by $[\neg p_1 \wedge p_2] = \{w_2, w_6\}$.The relation generating function R_f generates it relation to the logicatoms named p_1 and p_3. Now to evidence pleiotropy, consider adding the world w_1 to the semantic representation of the belief. Now $\{w_1, w_2, w_6\} = [(p_1 \wedge \neg p_2 \wedge p_3) \vee (\neg p_1 \wedge p_2)]$ which would result in the logicatom being related to no other logicatom i.e. a single world can simultaneously affect relations with several logicatoms.
For polygeny, consider the logicatom named p_1. Its belief is semantically represented by $[p_2] = \{w_2, w_3, w_6, w_7\}$. Assume I remove any one of these worlds. One has

$$\begin{aligned}
\{w_3, w_6, w_7\} &= [p_2 \wedge (p_1 \wedge \neg p_3 \vee p_3)] \\
\{w_2, w_6, w_7\} &= [p_2 \wedge (\neg p_1 \wedge \neg p_3 \vee p_3)] \\
\{w_2, w_3, w_7\} &= [p_2 \wedge (p_1 \wedge p_3 \vee \neg p_3)] \\
\{w_2, w_3, w_6\} &= [p_2 \wedge (\neg p_1 \wedge p_3 \vee \neg p_3)]
\end{aligned}$$

Thus the removal of any one world leaves the relation unchanged i.e. it will be related to p_2 in all the cases.

Mapping the formalism of theory networks to the natural selection variables, one has:

(a) The individual logicatoms will be viewed as replicators

(b) The individual worlds comprising the belief of a logicatom will denote the selectons

(c) The local relation of the logicatom universe represents the successor-predecessor relation required in the formalisation of natural selection. The relation is however 'inverted'. If $\mu R \nu$, I view ν as a parent of child μ. The reason is that the belief of ν will update the belief of μ, meaning that selectons will 'flow' from ν to μ i.e. the belief of a logicatom determines its 'parents'.

3.5.2 In search of evolving theory networks

For the remainder of this chapter, I only consider the class of theory networks (ξ, R_f, T) over PropCal_n with the transition function defined by $\xi_{t+1}(p) := [T(\xi_t)](p) = (\xi_t(p))^{\xi_t|_{R_f(p)}}$. This is just the constrained uniform substitution transition function, as defined in Definition 2.3.6. Various classes of relation generating functions will be analysed, in search of evolving theory networks. I proceed to define a few concepts that are necessary in achieving the construction objective. Firstly, in terms of notation, I will refer to any variable $p \in \Phi_n$ as an affirmed variable, while $\neg p \in \text{Form}(\Phi_n)$ will be known as a negated variable.

Definition 3.5.1. The class of *implication relations* comprises one of the following 3 relations defined on a logicatom universe:

(a) **Affirmed Implication Relation:** $\mu R \nu$ iff $\vdash B(\mu) \rightarrow N(\nu)$ and $\nvdash B(\mu) \leftrightarrow \bot$. The generating function for this relation is given for any $\phi \in Form(\Phi_n)$ with $\phi \neq \bot$ by
$R_f(\phi) = \{p \in \Phi_n \mid \vdash \phi \rightarrow p\}$

(b) **Negated Implication Relation:** $\mu R\nu$ iff $\vdash B(\mu) \rightarrow \neg N(\nu)$ and $\nvdash B(\mu) \leftrightarrow \bot$. The generating function is given for any $\phi \in Form(\Phi_n)$ with $\phi \neq \bot$ by
$R_f(\phi) = \{p \in \Phi_n \mid \vdash \phi \rightarrow \neg p\}$

(c) **Affirmed or Negated Implication Relation:** $\mu R\nu$ iff $\vdash B(\mu) \rightarrow N(\nu)$ or $\vdash B(\mu) \rightarrow \neg N(\nu)$ and $\nvdash B(\mu) \leftrightarrow \bot$. The generating function is given for any $\phi \in Form(\Phi_n)$ with $\phi \neq \bot$ by $R_f(\phi) = \{p \in \Phi_n \mid \vdash \phi \rightarrow \neg p$ or $\vdash \phi \rightarrow p\}$

Now consider the proposition $\psi_1 := (p \wedge \neg q) \vee (\neg p \wedge q)$. It is obvious that both variables p and q are required to express this proposition. Further, if I restrict the connectives to $\wedge$, $\vee$ and $\neg$, then p, $\neg p$, q and $\neg q$ are required to express this proposition. On the other hand, consider $\psi_2 := p \vee (\neg p \wedge q)$. Once again both variables are required. However in this case, I can express ψ_2 without using $\neg p$ since ψ_2 is equivalent to $q \vee (p \wedge \neg q)$. I will now formally define this concept independent of the logical operators used.

Definition 3.5.2. Only the affirmed variable p is required in proposition ϕ iff

(a) p is required in ϕ and

(b) $\exists \alpha, \beta \in Form(\Phi_n - \{p\})$ such that $\nvdash \alpha \leftrightarrow \beta$ and $\vdash \phi \leftrightarrow [(\alpha \wedge p) \vee \beta]$ respectively.

Analogously,

Definition 3.5.3. Only the negated variable $\neg p$ is required in proposition ϕ iff

(a) p is required in ϕ and

(b) $\exists \alpha, \beta \in Form(\Phi_n - \{p\})$ such that $\nvdash \alpha \leftrightarrow \beta$ and $\vdash \phi \leftrightarrow [(\alpha \wedge \neg p) \vee \beta]$ respectively.

This provides me with the final set of local relation definitions used in the remainder of this chapter.

Definition 3.5.4. The *minimum variable* class of relations comprises the following 4 relations:

(a) **Affirmed Or Negated Relation**: $\mu R\nu$ iff $N(\nu)$ is required in $B(\mu)$.

(b) **Affirmed And Negated Relation**: $\mu R\nu$ iff $N(\nu)$ *and* $\neg N(\mu)$ are required in $B(\mu)$

(c) **Only Affirmed Relation**: $\mu R\nu$ iff *only* $N(\nu)$ is required in $B(\mu)$

(d) **Only Negated Relation**: $\mu R\nu$ iff *only* $\neg N(\nu)$ is required in $B(\mu)$.

One can also form combinations of the above such as: $\mu R\nu$ iff $N(\nu)$ is required in $B(\mu)$ and it is not the case that only $\neg N(\nu)$ is required in $B(\mu)$. I will use the local relations defined (and combinations thereof) to investigate whether various examples of theory networks are regulated by natural selection. Towards this end, I need to show that the selectons (worlds) within each logicatom satisfy the requirement of heritability, differential fitness, variation and non-deterministic replication. As was evident in Section 3.3, the mathematics of natural selection uses statistics as

its building blocks. I thus use counting arguments as the foundations of my proofs. This will allow me to calculate the probability that a logicatom is related to k other logicatoms (for a given local relation in the theory network). Counting arguments will also be used to calculate the probability that a particular world is an element of a logicatom's belief (using possible world semantics) given that it is an element of a related logicatom's belief.

3.5.2.1 Calculating the parameters of natural selection

Sections 3.5.2.2, 3.5.2.3 and 3.5.2.4 are riddled with tedious combinatoric proofs (apologies to the reader) that calculate the various probabilities sought after. To top it all off, the work results in a negative answer - a theory network cannot be regulated by natural selection since, being a deterministic system, it cannot satisfy the requirement of non-deterministic replication. However, the results of this section are important for the following reasons:

(a) Some results will be reused when simulating pregeometries in Chapter 4

(b) The probability parameters of certain theory networks will be used to show interesting characteristics of rule-based reasoning from a natural selection perspective.

(c) The methods of proof shown are good tools for any person wishing to do further research in this topic.

I will use this section to basically explain the approach followed and summarise the results. With regards to $prob_i(w)$, the probability of inheritance of a selecton (world) w, one needs to calculate the probability that a logicatom μ is related to say k other logicatoms, and given that one of these logicatoms has the selecton (i.e. the world w is an element of the meaning of it's belief), μ will have the selecton after the transition function has been applied (i.e. the world w will be an element of the meaning of the updated belief of μ). The first calculation involves the relation generating function, since this determines the local relations. The second calculation involves the constrained uniform substitution function, since this determines the updated belief. Table 3.1 shows the results of these calculations for the various relation generating functions considered.

Constrained subset of $Form(\Phi_n)$ containing propositions that	**Number of propositions is given by**
imply at least one affirmed (negated) variable	$F(n) = \begin{cases} \left(\sum_{k=1}^{n-1} \binom{n}{k} \left(2^{2^{n-k}} - F(n-k)\right)\right) + 2 & \text{if } n > 1 \\ 2 & \text{if } n = 1 \end{cases}$
imply exactly k affirmed (negated) variables	$G(n,k) = \begin{cases} \binom{n}{k} \left(2^{2^{n-k}} - F(n-k)\right) & \text{if } k < n \\ 1 & \text{if } k = n \end{cases}$
imply at least one affirmed or negated variable	$H(n) = \begin{cases} \sum_{k=1}^{n-1} 2^k \binom{n}{k} \left(2^{2^{n-k}} - H(n-k)\right) + 2^n + 1 & \text{if } n > 1 \\ 3 & \text{if } n = 1 \end{cases}$
imply exactly k affirmed or negated variables	$I(n,k) = \begin{cases} 2^k \binom{n}{k} \left(2^{2^{n-k}} - H(n-k)\right) & \text{if } n > 1 \text{ and } k < n \\ 2^n & \text{if } n \geq 1 \text{ and } k = n \end{cases}$
imply at least one affirmed (negated) variable and are true in exactly k worlds	$\#\Pi_n^{\pm}(k) = \sum_{t=1}^{n} K^{\pm}(n,k,t)$
imply at least one affirmed or negated variable and are true in exactly k worlds	$\#\Pi_n(k) = 2\sum_{t=1}^{n} K(n,k,t)$
don't imply any affirmed or negated variables and are true in exactly k worlds	$J(n,k) = \binom{2^n}{k} - 2 * \sum_{t=1}^{n} K(n,k,t)$

Table 3.1: Counting arguments of constrained subsets of $Form(\Phi_n)$

where

$$K^{\pm}(m+1,k,s+1) = \binom{2^m}{k} - \sum_{l=1}^{s} K^{\pm}(n-1,k,l)$$

and

$$K(n,k,t) = \begin{cases} \binom{2^{n-1}}{k} - 2\sum_{l=1}^{t-1} K(n-1,k,l) & \text{if } n > 1 \text{ and } k \leq 2^{n-1} \\ 1 & \text{if } n = k = t = 1 \\ 0 & \text{otherwise} \end{cases}$$

These formulae allow us to build expressions for the probabilities sought after. For example, given $\phi \in Form(\Phi_n)$ such that ϕ implies no negated or affirmed variables, the probability that $w \in [\phi]$ for any world $w \in W$ is given by

$$prob(w \in [\phi]) \quad = \quad \frac{\sum_{k=1}^{2^n} J(n,k)k2^{-n}}{\sum_{k=1}^{2^n} J(n,k)}$$

This is observed noting that the number of propositions, not implying any affirmed or negated variable, and are true in exactly k worlds, is given by $J(n,k)$. The probability that an arbitrary world is in a set with k elements is given by $\frac{k}{2^n}$. The result follows.
Table 3.2 summarises the results of calculating these probabilities for various classes of theory networks over PropCal$_n$ for large n, as well as showing wether the requirements of natural selection are satisfied or not. The theory networks under consideration are theory networks governed by the constrained uniform substitution transition rule $(\xi_{t+1}(p) := [T(\xi_t)](p) = (\xi_t(p))^{\xi_t|_{R_{f(p)}}})$ with the relation generating function specified in the first column.

Local Relation	$prob_i(w)$	$prob_m(w)$	**Heritability**	**Differential Fitness**	**Variation**
Negated Implication Relation	0		×		
Affirmed Implication Relation	$\frac{1}{2}$	0	✓	✓	×
Affirmed or Negated Implication Relation	$\frac{1}{4}$	$\frac{1}{4}$	×		
Only Affirmed Relation	$\frac{3}{4}$	$\frac{1}{4}$	✓	✓	✓

Table 3.2: The natural selection parameters and satisfied requirements of the theory network analysed

In the case of the negated implication relation, I calculate the probability of inheritance $prob_i(w)$ of an arbitrary world to be 0. Since heritability has not been met, I stop there. In the case of the affirmed implication relation, I have the property of inheritance, but since the mutation probability is 0, this leads to the requirement of variation not being satisfied. In the case of the affirmed or negated implication relation, I have $prob_i(w) = prob_m(w)$, which does not satisfy 3.3.25. Finally, the last case of the only affirmed relation satisfies Darwin's three requirements, but can still not be said to be regulated by natural selection since one does not have non-deterministic replication. I will conclude this section by focusing on the results of the theory network governed by the affirmed

implication relation and constrained uniform substitution transition function. As shown in Table 3.2, heritability and differential fitness are satisfied while variation is not. Now the dynamics of this theory network was interpreted in terms of rule-based reasoning in Section 2.4.2. I believe that this shows why rule based reasoning cannot model thought - variation is not maintained since $prob_m(w) = 0$, and variation is required for creative thought. The rule based consequent (the updated beliefs of a logicatom) has a meaning comprising a subset of the union of the meaning of its parents' belief. On the other hand, it is interesting to note that rule-based reasoning can be said to satisfy two of the requirements. I now proceed to detail the calculations that have been summarised in this section.

3.5.2.2 The requirement of heritability

The hereditary criterion specifies that if a predecessor has the selecton, there is a non-zero probability that the successor will have the selecton. I proceed to prove some theorems that will show how certain theory networks have the property of heritability. Consider an arbitrary logicatom $\mu(t)$ in a logicatom universe at time t and let $\{\nu_i(t) : i \in \{1, 2, \ldots, k\}\ ,\ k \leq n\}$ be a non-empty indexed set of logicatoms related to $\mu(t)$. Given $w \in W$, I am required to calculate

(a) $prob_i(w)$: the probability that $w \in [B(\mu(t+1))]$ given that it is an element of at least one parent's belief i.e. $\exists j \in \{1, 2, \ldots, k\}$ such that $w \in [B(\nu_j)]$.

(b) $prob_m(w)$: the probability that $w \in [B(\mu(t+1))]$ given that it is not an element of any parent's belief i.e. $\forall i \in \{1, 2, \ldots, k\}\ w \notin [B(\nu_i)]$

In the proofs that follow, I have used the alternative notation for uniform substitution that lends itself to these proofs. To remind the reader, let $\xi : \Phi_n \rightarrow Form(\Phi_n)$ be some substitution map, such that

$$\xi(p_i) = \psi_i$$

Given $\phi \in Form(\Phi_n)$ then

$$\phi(\psi_1/p_1, \psi_2/p_2, \ldots, \psi_n/p_n) \equiv \phi^{\xi}$$

To get a feel of how I approach this problem, I will show how theory networks using the negated implication relation do not satisfy the heritability requirement.

Theorem 3.5.1. *Theory networks over PropCal$_n$ using constrained uniform substitution as the transition function and the negated implication relation ($\mu R\nu$ iff $\vdash B(\mu) \rightarrow \neg N(\nu)$) have the property that $prob_i(w) = 0$ for any world $w \in W$:*

Proof. Consider an arbitrary logicatom $\mu(t)$ in a logicatom universe at time t and let $\{\nu_i(t) : i \in \{1, 2, \ldots, k\}\ ,\ k \leq n\}$ be a non-empty indexed set of logicatoms related to $\mu(t)$. For ease of notation, define

$$\begin{aligned} B(\mu(t)) &:= \phi \\ B(\nu_i(t)) &:= \psi_i \\ N(\nu_i(t)) &:= p_i \end{aligned}$$

Assume $w \in [\psi_j]$ for some $j \in \{1, 2, \ldots, k\}$. Now since logicatoms are related according to the negated implication relation, the following holds for the belief of $\mu(t)$:

$$\vdash \phi \leftrightarrow (\neg p_j \wedge \alpha) \tag{3.5.1}$$

where $\alpha \in Form(\Phi_n - \{p_j\})$ not requiring p_j. This is deduced from the fact that $\mu(t)R\nu_i(t)$ for every i, implying that $\phi \rightarrow \neg p_i$ for every i. Applying the constrained uniform substitution transition rule (See Definition 2.3.6), one obtains

$$B(\mu(t+1)) = \neg\psi_j \wedge \alpha \tag{3.5.2}$$

Since $w \notin [\neg\psi_j]$, one can deduce that $w \notin [B(\mu(t+1))]$. This is true for any w, concluding that $prob_i(w) = 0$. □

The above theorem allows one to conclude that theory networks governed by the negated implication relation are not regulated by natural selection. However, in the theorems that follow, it will not suffice just to prove that $prob_i(w) > 0$. For example, if $prob_i(w) = \frac{1}{4}$ and $prob_m(w) = \frac{1}{2}$, one cannot claim that the heritability condition is satisfied since this inheritance will be 'swamped' out by random mutation. I therefore need to show that $prob_i(w) - prob_m(w) > 0$ for any world w.

The next set of theory networks that will be analysed are those governed by the *affirmed implication relation*, the *affirmed or negated implication relation* and the *only affirmed relation*. The proofs require counting arguments that are detailed in the subsequent lemmas.

Lemma 3.5.2. *Consider the subset Θ_n^+ (respectively Θ_n^-) of Form(Φ_n) defined by*

$$\Theta_n^+ = \{\phi \in Form(\Phi_n) \mid \exists p \in \Phi_n \text{ with } \vdash \phi \rightarrow p\} \tag{3.5.3}$$

$$\Theta_n^- = \{\phi \in Form(\Phi_n) \mid \exists p \in \Phi_n \text{ with } \vdash \phi \rightarrow \neg p\} \tag{3.5.4}$$

Θ_n^+ (Θ_n^-) comprises all the formulae in Form(Φ_n) that imply at least one affirmed (negated) propositional variable in Φ_n. Define $F^+(n) = \#\Theta_n^+$ and $F^-(n) = \#\Theta_n^-$.
Then $F^+(n) = F^-(n) = F(n)$ where

$$F(n) = \begin{cases} \left(\sum_{k=1}^{n-1} \binom{n}{k} \left(2^{2^{n-k}} - F(n-k)\right)\right) + 2 & \text{if } n > 1 \\ 2 & \text{if } n = 1 \end{cases} \tag{3.5.5}$$

Proof. I proceed to prove the result for $F^+(n)$ since an identically symmetrical proof shows the result for $F^-(n)$. The formula is clearly true for $n = 1$ since PropCal_1 has 2 wff $\{\bot, p\}$ that imply the single propositional variable p. Now assume the formula is true for all integers less than or equal to n. For the case of $n+1$, one needs to count all the propositions in $Form(\Phi_{n+1})$ that imply at least one affirmed propositional variable in Φ_{n+1}. I build up this set using the following constructive process.

To construct formulae implying only one affirmed propositional variable, choose some $p \in \Phi_{n+1}$. Define the proposition $\phi := p \wedge \psi$ with $\psi \in Form(\Phi_{n+1} - \{p\})$ not implying any affirmed variable and not requiring p. Since there are exactly 2^{2^n} propositions up to logical equivalence in $Form(\Phi_{n+1} - \{p\})$, one can deduce that the number propositions not implying any propositional variable is $2^{2^n} - F^+(n)$ (by definition of $F^+(n)$). Thus the number of propositions in PropCal_{n+1} that imply p and no other affirmed variable is $2^{2^n} - F^+(n)$. This is true for any variable $p \in Form(\Phi_{n+1})$, thus resulting in $(n+1)(2^{2^n} - F^+(n))$ propositions in PropCal_{n+1} that imply exactly one propositional variable.

For formulae implying k affirmed propositional variables with $k \leq n$, proceed similarly. Select k variables $p_1, p_2, \ldots, p_k \in \Phi_{n+1}$. Define the proposition $\phi := \left(\bigwedge_{i=1}^{k} p_i\right) \wedge \psi$ with $\psi \in Form(\Phi_{n+1} - \{p_1, p_2, \ldots, p_k\})$ not implying any affirmed propositional variable and not requiring $p_1, \ldots, p_k$. There are exactly $2^{2^{n+1-k}}$ propositions up to logical equivalence in $Form(\Phi_{n+1} - \{p_1, p_2, \ldots, p_k\})$, and thus $2^{2^{n+1-k}} - F^+(n+1-k)$ propositions not implying any affirmed propositional variable. Further, there are $\begin{pmatrix} n+1 \\ k \end{pmatrix}$ ways of choosing k distinct variables from a set of $n+1$ variables. Thus, one obtains $\begin{pmatrix} n+1 \\ k \end{pmatrix}\left(2^{2^{n+1-k}} - F^+(n+1-k)\right)$ propositions in $Form(\Phi_{n+1})$ implying k affirmed propositional variables. For $k = n+1$, one has 1 proposition $(\bigwedge_{p \in \Phi_{n+1}} p)$ implying all the affirmed variables. This together with the contradiction ($\bot$) adds 2 to the sum, resulting in the formula above. □

Corollary 3.5.3. *Define the sets*

$$\Delta_n^+(k) = \{ \quad \phi \in Form(\Phi_n) \mid \exists p_1, \ldots, p_k \in \Phi_n \text{ with } p_i \neq p_j \text{ for } i \neq j \text{ such that} \\ \vdash \phi \rightarrow \bigwedge_{i=1}^{k} p_i \text{ and } \nvdash \phi \rightarrow q \text{ for any } q \in \Phi_n - \{p_1, \ldots, p_k\} \ \} \tag{3.5.6}$$

$$\Delta_n^-(k) = \{ \quad \phi \in Form(\Phi_n) \mid \exists p_1, \ldots, p_k \in \Phi_n \text{ with } p_i \neq p_j \text{ for } i \neq j \text{ such that} \\ \vdash \phi \rightarrow \bigwedge_{i=1}^{k} \neg p_i \text{ and } \nvdash \phi \rightarrow \neg q \text{ for any } q \in \Phi_n - \{p_1, \ldots, p_k\} \ \} \tag{3.5.7}$$

comprising all formulae that imply exactly k affirmed (respectively negated) propositional variables. Define the number of elements in the sets by $G^+(n,k) = \#\Delta^+(n,k)$ and $G^-(n,k) = \#\Delta^-(n,k)$ respectively. Then $G^+(n,k) = G^-(n,k) = G(n,k)$ where

$$G(n,k) = \begin{cases} \begin{pmatrix} n \\ k \end{pmatrix}\left(2^{2^{n-k}} - F(n-k)\right) & \text{if } k < n \\ \\ 1 & \text{if } k = n \end{cases} \tag{3.5.8}$$

Proof. Once again, I prove the result for the affirmed propositional variables only - the negated variable case being an identically symmetrical proof. The result is clearly true for n=1. Note that the contradiction ($\bot$) is omitted since this does not imply an exact amount of variables. For $n > 1$, choose k variables $\{p_1, \ldots, p_k\}$ from Φ_n with $k < n$
Define the formula $\phi := \left(\bigwedge_{i=1}^{k} p_i\right) \wedge \psi$ with $\psi \in Form(\Phi_n - \{p_1, \ldots, p_k\})$ not implying any affirmed propositional variables. Using Lemma 3.5.2, there are exactly $F(n-k)$ propositions that imply at least one affirmed variable. Thus there are $\left(2^{2^{n-k}} - F(n-k)\right)$ propositions ψ that imply no affirmed propositional variables. Since there are $\begin{pmatrix} n \\ k \end{pmatrix}$ ways of choosing k propositions out of Φ_n, one arrives at the above result for $k < n$. For $k = n$ the result is the single formula $\bigwedge_{i=1}^{n} p_i$ □

The above lemmas provide me with the tools to analyse theory networks governed by the affirmed (respectively negated) implication relation. For theory networks governed by the affirmed *or* negated implication relation, the following counting lemmas will be relevant.

Lemma 3.5.4. *Consider the subset* Υ_n *of* $Form(\Phi_n)$ *defined by*

$$\Upsilon_n = \{\phi \in Form(\Phi_n) \mid \exists p \in \Phi_n \text{ with } \vdash \phi \to p \text{ or } \vdash \phi \to \neg p\}$$

Υ_n *comprises all formulae that imply at least one affirmed or negated propositional variables. Define* $H(n) = \#\Upsilon_n$. *Then*

$$H(n) = \begin{cases} \sum_{k=1}^{n-1} 2^k \begin{pmatrix} n \\ k \end{pmatrix} \left(2^{2^{n-k}} - H(n-k)\right) + 2^n + 1 & \text{if } n > 1 \\ 3 & \text{if } n = 1 \end{cases} \tag{3.5.9}$$

Proof. The equation is clearly true for $n = 1$ where we have the 3 formulae $\{\bot, p, \neg p\}$ implying at least one affirmed or negated variable. Now assume the formula is correct for all integers less than or equal to n. To construct a proposition $\phi \in Form(\Phi_{n+1})$ implying exactly k affirmed or negated variables with $1 \leq k < n+1$, select k elements $\{p_1, \ldots, p_k\} \subseteq \Phi_{n+1}$. For each element p_i define q_i such that $q_i := p_i$ or $q_i := \neg p_i$. The proposition $\phi := \bigwedge_{i=1}^{k} q_i \wedge \psi$ with $\psi \in Form(\Phi_{n+1} - \{p_1, \ldots, p_k\})$ and ψ not implying any affirmed or negated variables in $\Phi_{n+1} - \{p_1, \ldots, p_k\}$ satisfies the requirement. Now there are exactly $2^{2^{n+1-k}}$ propositions up to logical equivalence in $Form(\Phi_{n+1} - \{p_1, \ldots, p_k\})$ and by assumption, exactly $H(n+1-k)$ of these propositions imply at least one affirmed or negated variable. Thus there are $2^{2^{n+1-k}} - H(n+1-k)$ propositions up to logical equivalence in $Form(\Phi_{n+1} - \{p_1, \ldots, p_k\})$ that do not imply any affirmed or negated propositional variable. Further, there are $\begin{pmatrix} n \\ k \end{pmatrix}$ possible ways of choosing k variable from Φ_n and 2^k possible ways of deciding whether the variable or its negation is selected. This is true for all $1 \leq k < n+1$. For $k = n+1$, there are 2^{n+1} possible propositions of the form $\bigwedge_{i=1}^{n+1} q_i$ where $q_i := p_i$ or $q_i := \neg p_i$. Finally the contradiction adds 1 more formula to the set. I thus have

$$H(n+1) = \sum_{k=1}^{n} 2^k \begin{pmatrix} n+1 \\ k \end{pmatrix} \left(2^{2^{n+1-k}} - H(n+1-k)\right) + 2^{n+1} + 1.$$

By induction, the proof is concluded. □

Corollary 3.5.5. *Define the set*

$$\begin{aligned}\Gamma_n(k) = \{ \quad & \phi \in \mathit{Form}(\Phi_n) \mid \exists p_1, \ldots, p_k \in \Phi_n \text{ with } p_i \neq p_j \text{ for } i \neq j \text{ such that} \\ & \forall i \in \{1, \ldots, k\} \ \vdash \phi \rightarrow p_i \text{ or } \vdash \phi \rightarrow \neg p_i \\ & \text{and } \nvdash \phi \rightarrow q \text{ and } \nvdash \phi \rightarrow \neg q \text{ for any } q \in \Phi_n - \{p_1, \ldots, p_k\} \ \}\end{aligned} \tag{3.5.10}$$

comprising all formulae that imply exactly k affirmed or negated propositional variables. Let $I(n,k) = \#\Gamma_n(k)$. Then

$$I(n,k) = \begin{cases} 2^k \begin{pmatrix} n \\ k \end{pmatrix} \left(2^{2^{n-k}} - H(n-k)\right) & \text{if } n > 1 \text{ and } k < n \\ \\ 2^n & \text{if } n \geq 1 \text{ and } k = n \end{cases} \tag{3.5.11}$$

Proof. This is clearly true for $n = 1$. (Note that once again, the contradiction ($\bot$) is omitted since this does not imply an exact amount of variables.) To construct every $\phi \in \Gamma_n(k)$ select k variables $\{p_1, \ldots, p_k\} \subseteq \Phi_n$ with $k < n$. For each variable p_i define q_i such that $q_i := p_i$ or $q_i := \neg p_i$. The proposition $\phi := \bigwedge_{i=1}^{k} q_i \wedge \psi$ with $\psi \in \mathit{Form}(\Phi_{n+1} - \{p_1, \ldots, p_k\})$ and ψ not implying any affirmed or negated variables in $\Phi_{n+1} - \{p_1, \ldots, p_k\}$ satisfies the requirements stipulated for $\Gamma_n(k)$. Now there are exactly $2^{2^{n-k}}$ propositions up to logical equivalence in $\mathit{Form}(\Phi_n - \{p_1, \ldots, p_k\})$ and by assumption, exactly $H(n-k)$ of these propositions imply at least one affirmed or negated variable. Thus there are $2^{2^{n-k}} - H(n-k)$ propositions up to logical equivalence in $\mathit{Form}(\Phi_n - \{p_1, \ldots, p_k\})$ that do not imply any affirmed or negated propositional variable. Further, there are $\begin{pmatrix} n \\ k \end{pmatrix}$ possible ways of choosing k variables from Φ_n and 2^k possible ways of deciding whether the variable or its negation is selected. For $k = n$, there are 2^n possible propositions of the form $\bigwedge_{i=1}^{n} q_i$ where $q_i := p_i$ or $q_i := \neg p_i$. The result follows. □

The counting lemmas and corollaries above allow one to answer questions such as: 'What is the probability that a logicatom μ is related to k logicatoms?', for the various classes of relations used. However, in order to calculate the probability of inheritance (and as shall be shown, mutation) one further needs to answer questions such as: 'Given $\eta \in \mathit{Form}(\Phi_n)$ with η not implying any affirmed propositional variable, what is the probability that $w \in [\eta]$ for some $w \in W$?' Now if there are no 'constraints' imposed on η, this probability would be exactly $\frac{1}{2}$, since any world w is in exactly half the propositions (up to logical equivalence) in $\mathit{Form}(\Phi_n)$. However, this is not the case when the sample of propositions considered is 'constrained', as is evidenced in the next example.

Example 3.5.4. Consider the set $Q = \mathit{Form}(\Phi_2) - \Upsilon_2$ (with Υ_2 defined in Lemma 3.5.4) of all propositions in PropCal$_2$ that do not imply an affirmed or negated propositional variable. Let the variables in Φ_2 be $\Phi_2 = \{p, q\}$. The number of propositions in $\mathit{Form}(\Phi_2)$ that imply at least 1 affirmed or negated variable is given by $H(2) = 9$. Thus $\#Q = 2^{2^2} - 9 = 7$. These are listed as

$$Q = \{q \rightarrow p, p \leftrightarrow q, \neg(p \leftrightarrow q), p \rightarrow q, \neg p \vee \neg q, p \vee q, \top\}$$

One needs to represent these propositions using the possible world semantics in order to calculate the probability that any world is in one of these formulae. Towards this end, define the following worlds in PropCal_2.

p	q	**Valuation**
0	0	w_0
1	0	w_1
0	1	w_2
1	1	w_3

The propositions in Q are thus represented by

ϕ : **Proposition in Q**	$[\phi]$: **Semantic Representation**	$\#[\phi]$
$q \rightarrow p$	$\{w_0, w_1, w_3\}$	3
$p \leftrightarrow q$	$\{w_0, w_3\}$	2
$\neg(p \leftrightarrow q)$	$\{w_1, w_2\}$	2
$p \rightarrow q$	$\{w_0, w_2, w_3\}$	3
$\neg p \vee \neg q$	$\{w_0, w_1, w_2\}$	3
$p \vee q$	$\{w_1, w_2, w_3\}$	3
$\top$	$\{w_0, w_1, w_2, w_3\}$	4

This allows one to calculate the probability that $w \in [\phi]$ for any $w \in \{w_0, w_1, w_2, w_3\}$ and $\phi \in Q$. In particular, one evaluates $p(w \in [\phi]) = \frac{5}{7}$.

The above example shows how a constraint on the sample set of propositions changes the probability that any world is contained in one of the formulae. In order to generalise this result, first note that for an arbitrary subset $U \subseteq W$, given $w \in W$, one has $p(w \in U) = \frac{\#U}{\#W}$. Thus knowing the size of the sets $[\phi]$ of constrained propositions ϕ will allow us to calculate the required probability. This observation also allows one to understand why the probability in the example above is greater than a half. The constraint imposed in Example 3.5.4 is more likely to eliminate a proposition whose meaning has fewer elements. The constraint that no affirmed or negated variable is implied will always include all propositions with a world set count greater than 2^{n-1} and exclude propositions with a world set count of 1. The counting lemmas below detail this concept. As before, I first consider the case of implying only affirmed (respectively negated) variables.

Lemma 3.5.6. *Denote the variables in Φ_n by $\Phi_n = \{p_1, p_2, \ldots, p_t, \ldots, p_n\}$. Define $K^+(n,k,t)$ as being the number of propositions in Form(Φ_n) that are true in exactly k worlds and imply the affirmed variable p_t while not implying any of the affirmed variables $p_{t-1}, p_{t-2}, \ldots, p_1$. Then*

$$K^+(n,k,t) = \begin{cases} \begin{pmatrix} 2^{n-1} \\ k \end{pmatrix} - \sum_{l=1}^{t-1} K^+(n-1,k,l) & \text{if } n > 1 \text{ and } k \leq 2^{n-1} \\ 1 & \text{if } n = k = t = 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.5.12}$$

Proof. The result is clearly true for $k = n = 1$ since there is only 1 formula (i.e. $\phi := p_1$ for $\Phi_1 = \{p_1\}$) that implies an affirmed variable and is true in 1 world only. In the general case, any formula $\phi \in \mathit{Form}(\Phi_n)$ with $\#[\phi] > 2^{n-1}$ will imply no affirmed propositional variable. This is deduced by noting that if a formula ϕ implies an affirmed variable p_t, the affirmed variable must be true in all the worlds in $[\phi]$. But every affirmed variable is true in only 2^{n-1} worlds, resulting in a contradiction. One therefore only needs to consider the case of $k \leq 2^{n-1}$.

For the case $n = 2$, one has the following world representation

p_2	p_1	**World**
0	0	w_0
1	0	w_1
0	1	w_2
1	1	w_3

For $t = 1$, consider the variable p_1 which is true in w_2 and w_3. There are $\left(\begin{array}{c} 2 \\ k \end{array} \right) = K^+(2, k, 1)$ possible ways of choosing propositions true in exactly k worlds and implying p_1.
For $t = 2$, there are $\left(\begin{array}{c} 2 \\ k \end{array} \right)$ possible ways of choosing propositions true in exactly k worlds and implying p_2.
In the case $k = 1$, I have double counted a proposition (i.e. $[p \wedge \neg q] = \{w_2\}$), resulting in

$$\begin{aligned} & \left(\begin{array}{c} 2 \\ k \end{array} \right) - \left(\begin{array}{c} 1 \\ 1 \end{array} \right) \\ = \; & \left(\begin{array}{c} 2 \\ k \end{array} \right) - K^+(1, k, 1) \\ = \; & K^+(2, k, 2) \end{aligned}$$

as required.
In the case $k = 2$, I have $1 = K^+(2, 2, 2)$ proposition true in 2 worlds implying p_2 and not implying p_1. This concludes the case for $n = 2$.

Now assume that

$$K^+(n, k, t) = \left(\begin{array}{c} 2^{n-1} \\ k \end{array} \right) - \sum_{l=1}^{t-1} K^+(n-1, k, l) \tag{3.5.13}$$

holds for $n \leq m$, for every $1 \leq k \leq 2^{n-1}$ and $1 \leq t \leq n$. Consider the case $n = m + 1$.
For the case $t = 1$, one has exactly $\left(\begin{array}{c} 2^m \\ k \end{array} \right)$ possible propositions that imply p_1 and are true in k worlds, as required. Assume Equation 3.5.13 holds for the case $n = m+1$, $1 \leq t \leq s$ and $1 \leq k \leq 2^m$. For the case $t = s+1$, one has exactly $\left(\begin{array}{c} 2^m \\ k \end{array} \right)$ propositions that imply p_{s+1} and are true in k worlds. The crux of the argument is to note that exactly $K^+(m, k, s), K^+(m, k, s-1), \ldots, K^+(m, k, 1)$ of the $\left(\begin{array}{c} 2^m \\ k \end{array} \right)$ propositions imply $p_s, p_{s-1}, \ldots, p_1$ respectively. This is deduced by observing that if $\phi \in \mathit{Form}(\Phi_m)$ is a proposition that implies the affirmed variable p_s and does not imply the affirmed variables $p_{s-1}, p_{s-2}, \ldots, p_1$, then $p_{m+1} \wedge \phi$ is a proposition in $\mathit{Form}(\Phi_{m+1})$ that implies the affirmed

variables p_{m+1}, p_s and does not imply the affirmed variables $p_{s-1}, p_{s-2}, \ldots, p_1$. Thus

$$K^+(m+1,k,s+1) = \left(\begin{array}{c} 2^m \\ k \end{array} \right) - \sum_{l=1}^{s} K^+(n-1,k,l) \tag{3.5.14}$$

showing that Equation 3.5.13 holds for $n = m+1$, for every $1 \leq k \leq 2^{n-1}$ and $1 \leq t \leq n$. The result follows by induction. □

Corollary 3.5.7. *Let $\Pi_n^+(k) = \{\phi \mid \#[\phi] = k\} \cap \Theta_n^+$ be the set of formulae in Form(Φ_n) that imply at least 1 affirmed variable and are true in exactly k worlds with $k \geq 1$.*
Then $\#\Pi_n^+(k) = \sum_{t=1}^{n} K^+(n,k,t)$

Proof. By definition of $K^+(n,k,t)$ in Lemma 3.5.6, the result follows. □

Define $K^-(n,k,t)$ to be the number of propositions in *Form*(Φ_n) that are true in exactly k worlds and imply the negated variable p_t while not implying any of the negated variables $p_{t-1}, p_{t-2}, \ldots, p_1$
An identical proof to Lemma 3.5.6 shows that $K^-(n,k,t) = K^+(n,k,t)$.

Lemma 3.5.8. *Denote the variables in Φ_n by $\Phi_n = \{p_1, p_2, \ldots, p_t, \ldots, p_n\}$. Define $K(n,k,t)$ as being the number of propositions in Form($\Phi_n^)$ that are true in exactly k worlds and imply the affirmed variable p_t while not implying any of the affirmed **or** negated variables $p_{t-1}, \neg p_{t-1}, p_{t-2}, \neg p_{t-2}, \ldots, p_1, \neg p_1$. Then*

$$K(n,k,t) = \begin{cases} \left(\begin{array}{c} 2^{n-1} \\ k \end{array} \right) - 2\sum_{l=1}^{t-1} K(n-1,k,l) & \text{if } n > 1 \text{ and } k \leq 2^{n-1} \\ 1 & \text{if } n = k = t = 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.5.15}$$

Proof. This proof is identical to Lemma 3.5.6, except that now one needs to excise the formulaes that imply affirmed or negated variables when counting. Since the number of propositions in *Form*(Φ_n) that are true in exactly k worlds and imply the **affirmed** variable p_t while not implying any of the affirmed **or** negated variables $p_{t-1}, \neg p_{t-1}, p_{t-2}, \neg p_{t-2}, \ldots, p_1, \neg p_1$ is equal to number of propositions in that are true in exactly k worlds and imply the **negated** variable $\neg p_t$ while not implying any of the affirmed **or** negated variables $p_{t-1}, \neg p_{t-1}, p_{t-2}, \neg p_{t-2}, \ldots, p_1, \neg p_1$, this introduces the factor of 2 in the recursive formulae, leading to the result. □

Corollary 3.5.9. *Let $\Pi_n(k) = (\{\phi \mid \#[\phi] = k\} \cap \Gamma_n)$ be the set of formulae in Form(Φ_n) that imply at least 1 affirmed **or** negated variable and are true in exactly k worlds with $k \geq 1$. Then*

$$\#\Pi_n(k) = 2\sum_{t=1}^{n} K(n,k,t) \tag{3.5.16}$$

Proof. By definition of $K(n,k,t)$, the result follows. □

Corollary 3.5.10. *The number of propositions in Form(Φ_n) true in exactly k worlds not implying any affirmed or negated variables is given by*

$$J(n,k) = \left(\begin{array}{c} 2^n \\ k \end{array} \right) - 2 * \sum_{t=1}^{n} K(n,k,t) \tag{3.5.17}$$

Proof. There are $\left(\begin{array}{c} 2^n \\ k \end{array} \right)$ propositions in $Form(\Phi_n)$ true in exactly k worlds. Thus there are $\left(\begin{array}{c} 2^n \\ k \end{array} \right) - \#\Pi_n(k)$ true in exactly k worlds not implying any affirmed or negated variables. The result follows. □

Lemma 3.5.11. *Let $\phi \in Form(\Phi_n) - \Upsilon_n$ be a proposition that implies no negated or affirmed variable. Then for any $w \in W$, one has*

$$\begin{aligned} prob(w \in [\phi]) &= \frac{\sum_{k=1}^{2^n} J(n,k) k 2^{-n}}{\sum_{k=1}^{2^n} J(n,k)} \\ &= \frac{\sum_{k=1}^{2^n} J(n,k) k 2^{-n}}{2^{2^n} - H(n)} \end{aligned} \tag{3.5.18}$$

Proof. The number of propositions in $Form(\Phi_n) - \Upsilon_n$ that are true in exactly k worlds is given by $J(n,k)$ in Lemma 3.5.10. The probability that an arbitrary world is in a set with k elements is given by $\frac{k}{2^n}$. The result follows. □

One now has the counting tools to calculate the various probabilities in the theory networks under consideration.

Theorem 3.5.12. *Theory networks over $PropCal_n$ having the constrained uniform substitution transition function and the local relation determined by the affirmed or negated implication requirement have the following properties for any world $w \in W$:*

(a) $prob_i(w) \simeq \frac{1}{4}$ for $n \geq 6$

(b) $prob_m(w) \simeq \frac{1}{4}$ for $n \geq 6$

Here $prob_i(w)$ is the probability that a logicatom will contain an arbitrary selecton (a world w) after an iteration, given that one of its parents has the selecton. Similarly $prob_m(w)$ is the probability that a logicatom will contain an arbitrary selecton (a world w) after an iteration, given that none of its parents has the selecton.

Proof. Consider an arbitrary logicatom $\mu(t)$ in a universe at time t and let
$\{\nu_i(t) : i \in \{1, 2, \ldots, k\}\ , k \leq n\}$ be a non-empty indexed set of logicatoms related to $\mu(t)$. For ease of notation, define

$$\begin{aligned} B(\mu)(t) &:= \phi \\ B(\nu_i)(t) &:= \psi_i \\ N(\nu_i)(t) &:= p_i \end{aligned}$$

The definition of the affirmed or negated implication requirement allows one to assert that

$$\vdash \phi \leftrightarrow \left(\bigwedge_{i=1}^{k} q_i \wedge \beta \right)$$

where $q_i := p_i$ or $q_i := \neg p_i$ and $\beta \in Form(\Phi_n - \{p_1, \ldots, p_k\})$ does not imply any affirmed or negated variable in $\Phi_n - \{p_1, \ldots, p_k\}$. Applying the constrained uniform substitution transition rule, one obtains

$$B(\mu(t+1)) := \bigwedge_{i=1}^{k} \eta_i \wedge \beta \tag{3.5.19}$$

with $\eta_i := \psi_i$ or $\eta_i := \neg\psi_i$ depending on whether $q_i := p_i$ or $q_i := \neg p_i$ respectively.

To prove (a), assume $\exists j \in \{1, \ldots, k\}$ such that $w \in [\psi_j]$. Consider the case when $k < n$. Now $w \in [B(\mu(t+1))]$ iff $w \in [\psi_i]$ for every i and $w \in [\beta]$. The table below shows the various cases that can occur for $i \in \{1, \ldots, k\} - \{j\}$

	$q_i := p_i$	$q_i := \neg p_i$
$w \in [\eta_i]$	1	0
$w \in [\neg\eta_i]$	0	1

In 2 out of 4 cases, one has that $w \in [\eta_i]$ i.e. a probability of $\frac{1}{2}$. The probability that $w \in [\bigwedge_{i=1, i\neq j}^{k} \eta_i]$ is equal to $\frac{1}{2}^{k-1}$. For the case when $i = j$, since $w \in [\psi_j]$, one has $w \in [\eta_j]$ iff $q_j := p_j$. Once again, this can happen with a probability of $\frac{1}{2}$. Finally, using Equation 3.5.18, one has the probability that $w \in \beta$ is given by

$$prob(w \in \beta) = \frac{\sum_{l=1}^{2^{n-k}} J(n-k, l) l 2^{k-n}}{2^{2^{n-k}} - H(n-k)}$$

Combining these, one concludes that the probability of inheritance for the case when a logicatom is related to k logicatoms with $k < n$, is given by

$$prob_i(k) = \frac{\sum_{l=1}^{2^{n-k}} J(n-k, l) l 2^{-n}}{2^{2^{n-k}} - H(n-k)} \tag{3.5.20}$$

For the case when $k = n$, $\beta \leftrightarrow \top$ resulting in the probability being equal to $\frac{1}{2}^{n}$. To complete the analysis, one needs to calculate the probability that a logicatom is related to k other logicatoms. This is given by the probability that ϕ implies k affirmed or negated variables relative to the probability that ϕ implies at least 1 negated or affirmed variable. Using Equation 3.5.11 and 3.5.9 one obtains the final expression for the probability of inheritance as:

$$prob_i(w) = \left(\frac{1}{2}\right)^n \left[\sum_{k=1}^{n-1} \frac{I(n,k) \sum_{l=1}^{2^{n-k}} J(n-k, l) l}{(H(n)-1)\left(2^{2^{n-k}} - H(n-k)\right)} + \frac{I(n,n)}{(H(n)-1)} \right] \tag{3.5.21}$$

The table below shows how the value of this term converges rapidly to 0.25.

n	1	2	3	4	5	6	7	8	9	10
$prob_i(w)$	0.5	0.375	0.306	0.268	0.252	0.25	0.25	0.25	0.25	0.25

For (b), assume that $w \notin [\psi_i]$ for every i.
For the case $k < n$, $w \in B(\mu(t+1))$ iff $w \in \beta$ and $q_i := \neg p_i$ for every i.
The probability that $q_i := \neg p_i$ for every i is given by the number of propositions that imply exactly k negated variables relative to the total amount of propositions that imply k affirmed or negated variables. Using Equations 3.5.8 and 3.5.11 one obtains

$$prob(q_i := \neg p_i \text{ for every } i) = \frac{G(n,k)}{I(n,k)}$$

As before, the probability that $w \in [\beta]$ with $\beta \in Form(\Phi_{n-k})$ and β not implying any affirmed or negated variables given by Equation 3.5.20. Thus for a logicatom related to k other logicatoms with $k < n$, one has

$$prob_m(w,k) = \frac{G(n,k)}{I(n,k)} \frac{\sum_{l=1}^{2^{n-k}} J(n-k,l) l 2^{k-n}}{\left(2^{2^{n-k}} - H(n-k)\right)} \tag{3.5.22}$$

For the case $k = n$, one has $\eta \leftrightarrow \top$ resulting in $prob_m(w,n) = \frac{G(n,n)}{I(n,n)}$
As before, the probability that the logicatom is related to k other logicatoms is given $\frac{I(n,k)}{H(n)-1}$. Summing over k, one obtains

$$prob_m(w,k) = \frac{1}{H(n)-1} \left(\sum_{k=1}^{n-1} \left[G(n,k) \frac{\sum_{l=1}^{2^{n-k}} J(n-k,l) l 2^{k-n}}{\left(2^{2^{n-k}} - H(n-k)\right)} \right] + G(n,n) \right) \tag{3.5.23}$$

The table below shows how the value of this term converges rapidly to 0.25.

n	1	2	3	4	5	6	7	8	9	10
$prob_m(w)$	0.5	0.625	0.459	0.309	0.255	0.25	0.25	0.25	0.25	0.25

□

The above theorem shows that theory networks governed by the affirmed or negated implication relation and the constrained uniform substitution transition function do *not* satisfy the requirement of heritability since $prob_i(w) = prob_m(w)$ for large n.

Lemma 3.5.13. *Theory networks over $PropCal_n$ having the constrained uniform substitution transition function and the local relation determined by the affirmed implication relation have the following properties for any world $w \in W$:*

(a) $prob_i(w) \simeq \frac{1}{2}$ *for* $n \geq 6$

(b) $prob_m(w) = 0$

Proof. Consider a logicatom $\mu(t)$ in a universe at time t and let $\{\nu_i(t) : i \in \{1,\ldots,k\}\ ,\ k \leq n\}$ be a non-empty indexed set of logicatoms related to $\mu(t)$. For ease of notation, define

$$B(\mu)(t) := \phi$$
$$B(\nu_i)(t) := \psi_i$$
$$N(\nu_i)(t) := p_i$$

To prove (a), assume $\exists j \in \{1, 2, \ldots, k\}$ such that $w \in [\psi_j]$. Now since logicatoms are related using the affirmed implication relation, one can infer that

$$\vdash \phi \leftrightarrow \bigwedge_{j=1}^{k} p_j \wedge \beta \tag{3.5.24}$$

with $\beta \in Form(\Phi_n - \{p_1, \ldots, p_n\})$ not implying any affirmed variable in $\Phi_n - \{p_1, \ldots, p_n\})$ for any j. Applying the transition rule, one obtains

$$B(\mu(t+1)) = \bigwedge_{j=1}^{k} \psi_j \wedge \beta \tag{3.5.25}$$

Now by assumption, we have $w \in [\psi_j]$. Thus $w \in B[\mu(t+1)$ iff $w \in \beta$ and $w \in \psi_i$ for every $i \in \{1, 2, \ldots, k\}$. The probability that $w \in \beta$ is $\frac{1}{2}$. The probability that $w \in \psi_i$ for every i is $\frac{1}{2}^{k-1}$ (since by assumption, w is in ψ_j, this does not contribute to the term). Thus the probability that $w \in B[\mu(t+1)]$ is equal to $\frac{1}{2}^{k}$ if μ is related to exactly k logicatoms at time t. The probability that μ is related to k logicatoms is the proportion of propositions that imply exactly k propositional variables, relative to the number of propositions that imply at least one propositional variable. Using the previous lemma, I have

$$prob_i(w) = \frac{\sum_{k=1}^{n-1} G_n(k)\frac{1}{2}^{k} + G_n(n)\frac{1}{2}^{n-1}}{\sum_{k=1}^{n} G_n(k)} \tag{3.5.26}$$

The table below shows how the value of this term converges rapidly to 0.5.

n	1	2	3	4	5	6	7	8	9	10
$prob_i(w)$	1	0.5	0.453	0.480	0.498	0.5	0.5	0.5	0.5	0.5

To prove (b), assume $w \notin [\psi_i]$ for every i. By Equation 3.5.25, one has $w \notin B(\mu(t+1))$, concluding the result. □

In conclusion, theory networks governed by the affirmed implication relation and the constrained uniform substitution transition function satisfy the natural selection requirement of heritability.

Lemma 3.5.14. *Theory networks over $Propcal_n$ having the uniform substitution transition function and the local relation determined by the only affirmed relation($\mu R\nu$ iff only $N(\nu)$ is required in $B(\mu)$) have the following properties for any world $w \in W$:*

(a) $prob_i(w) = \frac{3}{4}$

(b) $prob_m(w) = \frac{1}{2}$

Proof. Consider a logicatom $\mu(t)$ in a universe at time t and let $\{\nu_i(t) : i \in \{1, \ldots, k\}\ , k \leq n\}$ be a non-empty indexed set of logicatoms related to $\mu(t)$. For ease of notation, define

$$B(\mu)(t) := \phi$$
$$B(\nu_i)(t) := \psi_i$$
$$N(\nu_i)(t) := p_i$$

To prove (a), assume $w \in [\psi_j]$ for some $j \in \{1, 2, \ldots, k\}$. I am thus assuming that at least 1 parent has the selecton. I need to calculate the probability that $w \in B(\mu(t+1))$. Now since $\mu R \nu_j$ we have

$$\vdash \phi \leftrightarrow (\alpha_j \wedge q_j) \vee \beta_j \tag{3.5.27}$$

for some propositions $\alpha_j, \beta_j \in \mathit{Form}(\Phi_n)$ not requiring q_j and $\alpha_j \nleftrightarrow \bot$. We deduce this since only $N(\nu_j(t))$ is required in $B(\mu(t))$. Now using the constrained uniform substitution transition rule we obtain

$$B(\mu(t+1)) := \alpha_j(\psi_1/q_1, \ldots, \psi_k/q_k) \wedge \psi_j \vee \beta_j(\psi_1/q_1, \ldots, \psi k/q_k) \tag{3.5.28}$$

Now by assumption we have $w \in [\psi_j]$. Now for any proposition $\chi \in \mathit{Form}(\Phi_n)$ we have $w \in [\chi]$ with probability $\frac{1}{2}$. Thus the probability that $w \notin [\alpha_j(\psi 1/q_1, \ldots, \psi_k/q_k)]$ and $w \notin [\beta_j(\psi 1/q_1, \ldots, \psi_k/q_k)]$ is equal to $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. The probability that w is in at least one of the propositions is $1 - \frac{1}{4} = \frac{3}{4}$. We can conclude that in a truly random sample (randomly select the beliefs of each logicatom), the probability that $w \in [B(\mu_{t+1})]$ is 0.75, thus concluding part (a) of the proof.

To prove (b), assume $\forall i \in \{1, 2, \ldots, k\}$ $w \notin [\psi_i]$. Now since $\mu R \nu_i$ for every i, we have

$$\vdash \phi \leftrightarrow \alpha_i \wedge q_i \vee \beta_i \ \forall i \in \{1, 2, \ldots, k\}$$

Applying the uniform substitution transition rule, one obtains

$$B(\mu(t+1)) := \alpha_i(\psi_1/q_1, \ldots, \psi_k/q_k) \wedge \psi_i \vee \beta_i(\psi_1/q_1, \ldots, \psi_k/q_k) \tag{3.5.29}$$

Since $w \notin [\psi]$, $w \in [B(\mu(t+1))]$ iff $w \in \beta_i(\psi_1/q_1, \ldots, \psi_k/q_k)$. This can happen with a probability of $\frac{1}{2}$. □

Thus, theory networks governed by the only affirmed relation and the constrained uniform substitution transition function satisfy the natural selection requirement of heritability.

3.5.2.3 The requirement of differential fitness

Section 3.3 quantifies the requirement of differential fitness using the correlation between a replicator having a selecton and the corresponding number of offspring. In theory networks, this would mean that one is required to show that there is a correlation between a logicatom believing something, and the amount of children the logicatom has. Now first observe that in the case of theory networks, it looks as if this property has been turned on its head! What a logicatom believes determines its parents, who in the next generation, will update its belief according to the transition function specified. One cannot immediately see a way that a logicatoms' belief can influence the amount of children it has i.e. the amount of logicatoms that will be related to it in future generations.

In order to solve this problem and prove that there exist theory networks that satisfy the requirement of differential fitness, consider the following argument:
Let $\mu(0)$ be a logicatom in some logicatom universe U_0 of a theory network at time $t = 0$, that satisfies the heritability requirement: $p_i(w) > p_m(w) \geq 0$. Further, assume the belief of $\mu(0)$ is such

that it is related to itself at time $t = 0$. Due to heritability, any logicatom $\nu(0)$ that is related to $\mu(0)$ at time $t = 0$ will have a small non-zero chance of believing at time $t = 1$ what $\mu(0)$ believes. This means $\nu(1)$ is more likely (compared to the case when $\mu(0)$ is not related to itself) to be related to $\mu(1)$ at time $t = 1$. Further, any logicatom $\eta(1)$ related to $\nu(1)$ will also have a chance of believing what $\nu(1)$ believes, implying that there is a non-zero chance that it will be related to μ in the future. In summary, if I (being a logicatom) am related to myself, anything that is related to me is more likely to be related to me in future generations. But if there is more chance of logicatoms being related to me, then there is more chance that I will be a parent, and thus there is a greater chance that I will have more children in future generations. So the belief *can* influence the amount of children in future generation. This argument is the essence of the proof for differential fitness below.

The concept that will prove central to the proof of differential fitness is that of inheriting a trait from a parent. Now a trait might be the phenotype of a single selecton, or might require a combination of selectons. The work so far has focused on the probability of inheriting individual selectons. The particular trait that I am interested in is the property of whether a logicatom μ is related to a specific logicatom (say the one named p). Towards this end, define the boolean function representing this trait as

$$Trait(\mu, p) = \begin{cases} 1 & \text{if } \mu R\nu \text{ with } N(\nu) = p \\ 0 & \text{otherwise} \end{cases} \tag{3.5.30}$$

What I need to analyse is the conditional probability that a logicatom μ will have this trait given that one of its parents has the trait i.e.

$$Prob\,(Trait(\mu, p) = 1 \mid \exists \nu \text{ with } \mu R\nu \text{ and } Trait(\nu, p) = 1)$$

This will naturally depend on the relation and transition function under consideration. The following two lemmas provide the fundamental results that I require to prove the differential fitness property in the theory networks that satisfied the heritability requirements in the previous section.

Lemma 3.5.15. *Consider a theory network over $PropCal_n$ governed by the affirmed implication relation and the constrained uniform substitution transition function. Define*

$$P_1(\mu, p, t) = Prob\,(Trait(\mu(t+1), p) = 1 \mid \exists \nu \text{ such that } \mu(t) R \nu(t) \text{ and } Trait(\nu(t), p) = 1)$$
$$P_0(\mu, p, t) = Prob\,(Trait(\mu(t+1), p) = 1 \mid \forall \nu \text{ such that } \mu(t) R \nu(t) \text{ one has } T(\nu(t), p) = 0)$$

$$P_1(\mu, p, t) > P_0(\mu, p, t) \tag{3.5.31}$$

for any propositional variable p, time t and logicatom μ.

Proof. Assume logicatom $\mu(t)$ is related to logicatoms $\nu_1(t), \ldots, \nu_k(t)$. Now since one is using the affirmed implication relation, one has

$$\vdash B(\mu(t)) \rightarrow \bigwedge_{i=1}^{k} N(\nu_i(t))$$

resulting in

$$\vdash B(\mu(t+1)) \leftrightarrow \left(\left(\bigwedge_{i=1}^{k} B(\nu_i(t)) \right) \wedge \eta \right)$$

for some proposition $\eta \in Form(\Phi_n)$ not requiring the variables $\{N(\nu^1), \ldots, N(\nu^k)\} \subset \Phi_n$ and not implying any affirmed variables in $\Phi_n - \{N(\nu^1), \ldots, N(\nu^k)\}$. Thus one can conclude that

$$\vdash B(\mu(t+1)) \rightarrow B(\nu_i(t))$$

for every i. Now if there exists a $j \in \{1, 2, \ldots, k\}$ such that $\nu_j(t) R \lambda^p(t)$, where $\lambda^p(t)$ is the unique logicatom with name p, one can conclude that $\vdash B(\nu_j(t)) \rightarrow p$. This implies that $\vdash B(\mu(t+1)) \rightarrow p$. Thus $B(\mu(t+1))$ is the contradiction $\perp$ or $\mu_{t+1} R \lambda^p_{t+1}$ i.e. $Trait(\mu_{t+1}, p) = 1$.
On the other hand, if no such j exists, $B(\mu(t+1))$ is the contradiction $\perp$ or $Trait(\mu(t+1), p) = 1$ (this is the case when $\eta \rightarrow p$) or $Trait(\mu(t+1), p) = 0$ (this is the case when η does not imply p). The chance that $B(\mu(t+1))$ is a contradiction as equal in both scenarios for a truly random sample, and one can thus eliminate these cases. For the case when $B(\mu(t+1))$ is not the contradiction, one has that $P_1(\mu, p, t) = 1$ and $P_0(\mu, p, t) = \frac{1}{2}$. Thus $P_1(\mu, p, t) > P_0(\mu, p, t)$ as required. □

Lemma 3.5.16. *Consider a theory network over $PropCal_n$ governed by the only affirmed relation and the constrained uniform substitution transition function. Define*

$$P_1(\mu, p, t) = Prob\left(Trait(\mu(t+1), p) = 1 \mid \exists \nu \text{ such that } \mu(t) R \nu(t) \text{ and } Trait(\nu(t), p) = 1\right)$$
$$P_0(\mu, p, t) = Prob\left(Trait(\mu(t+1), p) = 1 \mid \forall \nu \text{ such that } \mu(t) R \nu(t) \text{ one has } Trait(\nu(t), p) = 0\right)$$

Then

$$P_1(\mu, p, t) > P_0(\mu, p, t) \tag{3.5.32}$$

for any propositional variable p, time t and logicatom μ.

Proof. Consider an arbitrary proposition $\phi \in Form(\Phi_n)$. Let

$$\begin{aligned} \pi_1 &= Prob(\phi \text{ requires only } p) \\ \pi_2 &= Prob(\phi \text{ requires } p \text{ and } \neg p) \\ \pi_3 &= Prob(\phi \text{ requires only } \neg p) \\ \pi_4 &= Prob(\phi \text{ does not require } p) \end{aligned} \tag{3.5.33}$$

Clearly $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$.
Now assume logicatom $\mu(t)$ is related to logicatoms $\nu_1(t), \ldots, \nu_k(t)$. Consider the case where $\forall i \in \{1, 2, \ldots, k\}$, one has $Trait(\nu_i(t), p) = 0$. This implies that $\mu(t)$ is not related to any logicatom whose belief requires only p. Since we are using the only affirmed relation, we cannot say anything about the updated belief $B(\mu(t+1))$ except that $P_0(\mu, p, t) = \pi_1$.

On the other hand, assume there exists a $j \in \{1, 2, \ldots, k\}$ such that $\nu_j(t) R \lambda^p(t)$, where $\lambda^p(t)$ is the unique logicatom with name p. Now since one is using the only affirmed relation, one can conclude that

$$\vdash B(\mu(t)) \leftrightarrow (\alpha \wedge N(\nu_j(t)) \vee \beta)$$

for some propositions $\alpha, \beta \in Form(\Phi_n)$. Thus

$$\vdash B(\mu(t+1)) \leftrightarrow (\alpha' \wedge B(\nu_j(t)) \vee \beta')$$

Since $\nu_j(t) R \lambda^p(t)$, one can conclude that $\vdash B(\nu_j(t)) \leftrightarrow (\gamma \wedge p \vee \delta)$ for some propositions $\gamma, \delta \in Form(\Phi_n)$. Thus $\vdash B(\mu(t+1)) \leftrightarrow (\alpha'' \wedge p \vee \beta'')$ for some propositions $\alpha'', \beta'' \in Form(\Phi_n)$. Now note that in the case when α'' requires only p or does not require p or requires p and $\neg p$, and β'' requires only p or does not require p, one has that $B(\mu(t+1))$ requires only p. Also note that in the case when α'' requires only $\neg p$ and β'' requires only p, one has that $B(\mu(t+1))$ requires only p. Thus

$$\begin{aligned} P_1(\mu, p, t) &\geq (\pi_4 + \pi_3 + \pi_1)(\pi_1 + \pi_4) + \pi_2 \pi_1 \\ &= \pi_1(\pi_1 + \pi_2 + \pi_3 + \pi_4) + \pi_4(\pi_1 + \pi_3 + \pi_4) \\ &= \pi_1 + \pi_4(\pi_1 + \pi_3 + \pi_4) \\ &\geq \pi_1 \\ &= P_0(\mu, p, t) \end{aligned} \tag{3.5.34}$$

as required. □

The results of these lemmas are now used to prove that the theory networks satisfying the heritability requirement in Subsection 3.5.2.2 satisfy the differential fitness requirement of natural selection.

Theorem 3.5.17. *Consider theory networks $\mathcal{T}_1, \mathcal{T}_2$. Assume both are governed by the constrained uniform substitution transition function and the affirmed implication relation (the only affirmed relation respectively). Further assume that logicatom μ in $\mathcal{T}_1$ is self-referential at time t while logicatom μ in $\mathcal{T}_2$ is not. Let $S_1(\mu(t)) = \{\nu(t) \mid \nu(t) R \mu(t)\}$ be the successor set of μ at time t for theory network $\mathcal{T}_1$. Similarly, let $S_2(\mu(t))$ be the successor set of μ at time t for theory network $\mathcal{T}_2$ Define*

$$\begin{aligned} m_1 &= \# \bigcup_{\nu(t) \in S_1(\mu(t))} S_1(\nu(t)) \\ m_2 &= \# \bigcup_{\nu(t) \in S_2(\mu(t))} S_2(\nu(t)) \end{aligned}$$

If $m_1 = m_2$, then

$$Prob(\#S_1(\mu(t+k)) >= \#S_2(\mu(t+k))) > \frac{1}{2} \tag{3.5.35}$$

for $t > 0$ and $k > 1$ i.e the number of children that μ fathers in $\mathcal{T}_1$ in two or more generations is likely to be greater than the corresponding number of children that μ fathers in $\mathcal{T}_2$.

Proof. m_1 and m_2 represent the number of logicatoms that are related to logicatoms that have the trait I am interested in i.e. being related to μ. As shown in Lemma 3.5.15 (Lemma 3.5.16 respectively), the theory networks under consideration satisfy the property that these traits are

heritable. Assume the probability of inheriting the trait in $\mathcal{T}_1$ and $\mathcal{T}_2$ is p_R. Then there are likely to be $p_R m_1$ ($p_R m_2$) logicatoms related to μ in the next generation in $\mathcal{T}_1$ ($\mathcal{T}_2$ respectively). If $m_1 = m_2$, then the amount of logicatoms related to μ is likely to be equal in both theory networks. However, in $\mathcal{T}_1$, there is a more likely chance that $\mu(t+1)R\mu(t+1)$ than there is in $\mathcal{T}_2$. All things being equal, one will have more grandchildren in the following generation $t = 2$. It is this discrepancy that leads to the result sought after. □

3.5.2.4 The requirement of variation

In terms of the requirement of variation, Section 3.3.3 states that the $prob_m(w) > 0$ for variation to persist from one generation to the next. By Lemma 3.5.13, variation will not persist in theory networks governed by the affirmed implication relation. In summary, theory networks governed by the *only affirmed relation* and the *constrained uniform substitution transition function* satisfy the requirements of heritability, differential fitness and variation.

3.5.2.5 The requirement of non-deterministic replication

All theory networks clearly do not satisfy this requirement. They are deterministic systems. Given a configuration at time t, one can predict the exact configuration at time $t+1$. Thus even though I have shown that there exists a class of theory networks whose dynamics satisfy the three requirements of natural selection, they will still always evolve to a fixed point.

3.5.3 The answer in inverted theory networks

In order to overcome this hurdle, I turn to physics to try understand what more is required of the sought after structure. In particular, I consider the arrow of time.

3.5.3.1 The arrow of time

Looking closer at the problem, one notices that the unwanted determinism in theory networks actually acts in the wrong direction. To understand this, consider the following example.

Example 3.5.5. Let $\mathcal{T}_1 = (\xi_1, R_f, T)$ and $\mathcal{T}_2 = (\xi_2, R_f, T)$ be two theory networks, both governed by the affirmed implication relation and the constrained uniform substitution transition function. The substitution maps ξ_1 and ξ_2 define the logicatom universes $\mathcal{U}_1$ and $\mathcal{U}_2$ represented in Figure 3.6: The reader can easily confirm that both these logicatom universes will evolve, in the next time step, to the logicatom universe represented by $\mathcal{U}'$.

The above example shows how all future states are deterministic, while past states are not - a topsy turvy scenario that one expects to find in Carroll's wonderland. However, this ends up being my saving grace. Firstly, note that this asymmetry in time is a required component for a pregeometry in physics. Quantum theories in modern physics are symmetrical in time, allowing for two solutions -

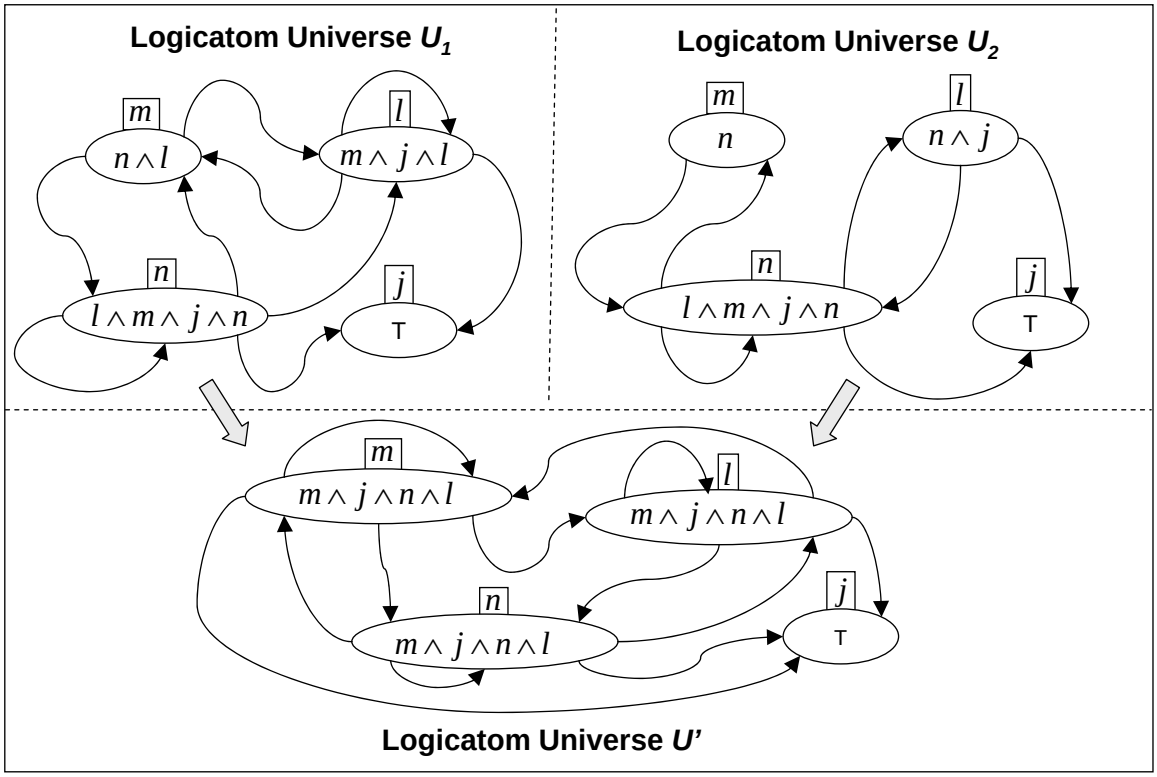

Figure 3.6: Non-isomorphic universes evolving to isomorphic universes

one the time reverse of the other [23].[3] But this is not embodied in the physics. From an observational perspective, this symmetry is clearly not realised in nature. However, this fundamental asymmetry in time is simply embodied in theory networks - the one direction is deterministic, the other not. Secondly, this property allows me to complete my construction objective. Just 'invert' the direction that the theory network evolves. In other words, run the example in the above diagram backwards. Starting at logicatom universe $\mathcal{U}'$ proceed to $\mathcal{U}_1$ or $\mathcal{U}_2$. The non-determinism is embodied in the fact that there is no property that specifies the choice of say $\mathcal{U}_1$ over $\mathcal{U}_2$. This also allows me to leave wonderland - the past is fully known; the future is unknown, yet constrained by the state I'm currently in.

[3]Only the solution that ensures the increase of entropy is viable.

3.5.3.2 Inverted theory networks

Formally, an inverted theory network represented as $\mathcal{T}^{-1}$ over PropCal_n is once again specified by the tuple (ξ, R_f, T). The difference comes in that for a given time t, it consists of the set of all substitution maps $\mathcal{T}^{-1}(t) = \{\zeta_1, \zeta_2, \ldots, \zeta_k\}$ satisfying the property that every theory network (ζ_i, R_f, T) specified by each substitution ζ_i will deterministically evolve to (ξ, R_f, T) after t time steps according to the usual definition of a theory network. In terms of notation, I say that each substitution map ζ_i specifies a 'possible future' to ξ. I will also decrement time to index the evolution of the individual logicatoms (and substitution maps). Thus if $\mu(t)$ is a logicatom in the logicatom universe specified by ξ, and $\zeta \in \mathcal{P}^{-1}(1)$ is a possible future one time step away, the corresponding logicatom in the universe represented by ζ is denoted as $\mu^{\zeta}(t-1)$. The superscript tells one which future is under consideration, and the time index $(t-1)$ tells one that I am one state away from the initial specified universe.[4] Other than this, the representation of inverted theory networks is identical to that of theory networks. I can represent the logicatom universe of each possible future using G-models and consequently, G-defining propositions. I have thus not lost the property that logicatoms have the 'ability' to describe their universe. I will now proceed to show that inverted theory networks are the sought after structures required to complete the construction objective. Firstly I need to clarify the notions of predecessors and successors, since everything has been turned on its head. In theory networks, if at time t one had $\mu(t)R\nu(t)$ then one regarded $\mu(t+1)$ as a successor of $\nu(t)$. This was clear since the belief of $\nu(t)$ influenced the belief of $\mu(t+1)$.
In inverted theory networks, given a possible future ζ one time step away from ξ, if $\mu^{\zeta}(t-1)R\nu^{\zeta}(t-1)$ then I say that $\nu^{\zeta}(t-1)$ is a (possible) successor of $\mu(t)$. Figure 3.7 below shows an example of a logicatom with its successors in two possible futures. The definitions for the probability of inheritance and mutation are now respectively:

(a) $prob_i^{\zeta}(w)$: the probability that $w \in [B(\nu^{\zeta}(t-1))]$ given that it is an element of at least 1 predecessors belief i.e. there exists a logicatom μ such that $\mu^{\zeta}(t-1)R\nu^{\zeta}(t-1)$ and $w \in B(\mu(t))$.

(b) $prob_m^{\zeta}(w)$: the probability that $w \in [B(\nu^{\zeta}(t-1))]$ given that it is not an element of any predecessors belief i.e. for every logicatom, μ such that $\mu^{\zeta}(t-1)R\nu^{\zeta}(t-1)$, one has $w \notin B(\mu(t))$

Example 3.5.6. Figure 3.7 below shows two possible futures of an inverted theory network, together with the successors of logicatom l in both futures By definition, logicatom l has m, l, j as successors in possible future ζ_1 and n, j in possible future ζ_2

[4] It must be emphasised that configurations can exists that have no possible futures. In cellular automaton systems, these are called 'Garden-of-Eden' states [9].

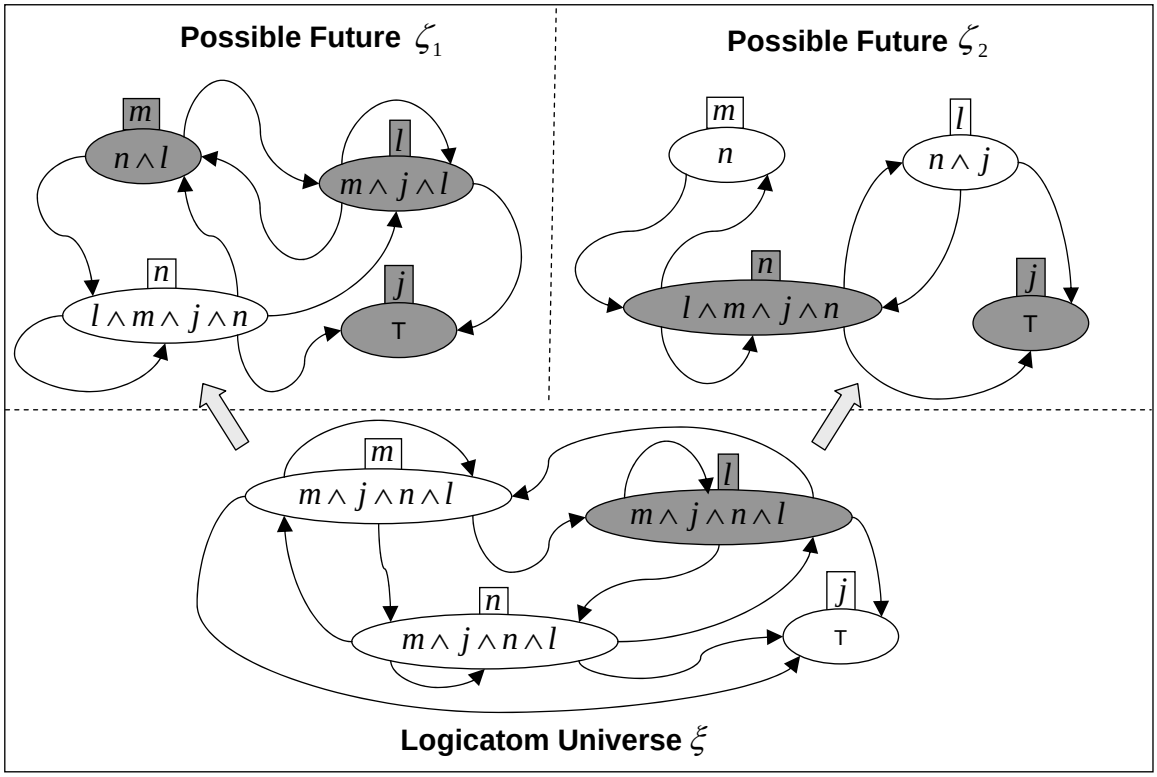

Figure 3.7: Non-deterministic replication in inverted theory networks

The following theorem proves that the requirements of heritability, variation, differential fitness and non-deterministic replication are satisfied in a particular class of inverted theory networks.

Theorem 3.5.18. *Let $\mathcal{T}^{-1} = (\xi, R_f, T)$ be an inverted theory network over PropCal$_n$, governed by the affirmed implication relation and the constrained uniform-substitution transition function. Let ζ be a possible future of ξ. Then the following properties hold:*

(a) $prob_i(w) = 1$

(b) $0 < prob_m(w) < 1$

(c) Any logicatom $\mu(t)$ in ξ that is self-referential (i.e. $\mu(t) R \mu(t)$) has at least 1 successor $(\nu^\zeta(t-1))$ in ζ

Proof. To prove (a), consider any logicatom $\nu^\zeta(t-1)$ in the possible future ζ, and let the set $\{\mu_1^\zeta(t), \ldots, \mu_k^\zeta(t)\}$ contain all its predecessors in ξ. Assume that there exists j such that $w \in [B(\mu_j^\zeta(t))]$. Since the local relation is the affirmed implication relation, the following property holds in the possible future ζ.

$$\vdash B\left[\mu_j^\zeta(t-1)\right] \rightarrow N\left[\nu^\zeta(t-1)\right] \tag{3.5.36}$$

Thus one may conclude that

$$\vdash B\left[\mu_j^{\zeta}(t-1)\right] \leftrightarrow \left(\phi \wedge N\left[\nu^{\zeta}(t-1)\right]\right) \tag{3.5.37}$$

for some ϕ not requiring $N[\nu^{\zeta}(t-1)]$. The constrained uniform substitution transition function implies that

$$B\left[\mu_j(t)\right] := \phi' \wedge B\left[\nu^{\zeta}(t-1)\right] \tag{3.5.38}$$

for some ϕ'. Now since $w \in [B[\mu_j(t)]]$, one has that $w \in \left[B\left[\nu^{\zeta}(t-1)\right]\right]$, thus proving the result.
For (b), one clearly has $prob_m < 1$ since if this were not the case, $prob_m = 1$ together with (a) would imply that the beliefs of all successors of a logicatom in an inverted theory network are a tautology, which is clearly not true. Further, if $prob_m = 0$, this together with (a) would imply that all successors of a logicatom μ in an inverted theory network will believe the same belief as μ, another clearly false statement. (The reader needs only look at Figure 3.7 to confirm these arguments)
For (c), assume there exists a self-referential logicatom $\mu(t)$ in ξ that has no successors. By definition, this implies that for every logicatom $\nu^{\zeta}(t-1)$ in ζ, it is not the case that $\nu^{\zeta}(t-1)R\mu^{\zeta}(t-1)$. Since no logicatom is related to $\mu^{\zeta}(t-1)$, the definition of the transition function implies that the belief of $\mu^{\zeta}(t-1)$ cannot change (when moving forward from a theory network perspective). Thus $B\left[\mu^{\zeta}(t-1)\right] = B\left[\mu_t\right]$. But since $\mu(t)$ is self-referential, this would imply that $\mu^{\zeta}(t-1)$ is self-referential, contradicting the fact that it is related to no other logicatom. □

The theorem above tells one that inverted theory networks regulated by the affirmed implication relation and the constrained uniform substitution transition function satisfy:

(a) The requirement of heritability, since $prob_i(w) = 1$ and $prob_m(w) < prob_i(w)$

(b) The requirement of variation, since $0 < prob_m(w) < 1$

(c) The requirement of differential fitness since a particular property of a logicatoms belief constrains the minimum amount of successors of the logicatom, showing that there can exist relationships between the belief of a logicatom and the number of successors i.e. the correlation is non-zero

(d) The requirement of non-deterministic replication is clearly satisfied (by construction) since the various possible futures imply that the number of successors of any logicatom is not set in stone. Figure 3.7 shows how logicatom l has 3 successors in one possible future and 2 successors in the other.

3.6 Summary of Chapter 3

The requirements of natural selection were formalised through the derivation of a generic equation. Inverted theory networks were defined and the construction objective was met with the construction of a particular example i.e. I have succeeded in creating a mathematical space defined using description logic, whose dynamics are regulated by natural selection.

Chapter 4

Simulating Pregeometries

In this chapter, I analyse whether inverted theory networks can serve as platforms for modelling Wheeler's pregeometries and creative thought. Section 4.1 argues how certain observables in physics can in fact be modelled using inverted theory networks. The fact that the physical dimension of space is three is shown to be predicted by inverted theory networks governed by the uniform substitution transition function and the affirmed implication relation i.e. the inverted theory networks governed by natural selection. In particular, I review work done by Nagels and show that these inverted theory networks will 'on the large scale' look like a 3+1 dimensional curved space. As mentioned in Chapter 3, the arrow of time is shown to be a consequence of the structure itself. Section 4.2 reviews the basics of quantum theory and argues how it is possible that it could arise naturally in the inverted theory network governed by natural selection. Section 4.3 concludes the chapter with a discussion analysing whether these structures could model thought and puts forward a proposal to achieve this objective.

4.1 Geometrodynamics

Einstein confused as all
as to why the apple did fall
said spacetime's not flat
and where we all sat
spacetime was curved like a ball!

And thus the story goes on
that Einstein thought very long
and after many derivations
wrote fifteen field equations
and so we continue our song!

Now as we come to the end
we must all remember to bend
as we pass a fat nun

or the bright heavy sun
or equations we'll all have to mend!

The geometric behaviour of physics is reviewed using the theory of geometrodynamics [55]. The motive herein is to show how purely geometric concepts can be used to explain certain classical and quantum physical observables. I conclude this section by discussing Wheeler's pregeometry within this context.

4.1.1 Classical geometrodynamics

Geometrodynamics is the theory that physics can be represented purely using geometry. This concept is best understood if one states the two highly contrasted views of the nature of physics [78]:

- Spacetime serves only as the 'arena' in which fields and particles interact. In this view, fields and particles, together with the laws that govern them, must be added to the spacetime geometry to permit any physics.
- All physics is a manifestation of the bending of space i.e. physics is geometry

Currently, accepted physics takes the middle ground in that both the curvature of spacetime (Einstein's theory of general relativity) and external fields (quantum field theories such as quantum electrodynamics) are used to explain observables. To review geometrodynamics as applied to classical physics (i.e. non-quantised matter or fields), I commence with the theory of general relativity and classical electrodynamics.
Einstein proposed the following fundamental principles to construct his general theory of relativity:

1. **Principle of General Relativity:** Otherwise known as the principle of general covariance: All laws in physics take the same form in any coordinate system.
2. **Principle of Equivalence:** There exists a coordinate system in which the effect of a gravitational field vanishes locally.

These postulates allowed Einstein to derive the field equations. These encompass tensor equations on a 4 dimensional Riemannian manifold:

$$R_{ab} - \frac{1}{2}Rg_{ab} \quad = \quad 8\pi T_{ab} \tag{4.1.1}$$

In keeping with the spirit of the topic, Equation 4.1.1 is expressed in geometrised units. (The gravitational constant G and speed of light c are set to 1. All quantities are as a result given a

dimension of a power of length.) General relativity can be summarised as follows [77]: "Spacetime is a manifold M on which there is defined a Lorentz metric g_{ab}. The curvature of g_{ab} is related to the matter distribution T_{ab} in spacetime by Einstein's equation". Intuitively, Equation 4.1.1 tells us that energy curves spacetime and curved spacetime determines the classical dynamics of energy. (Aside: It was shown in Chapter 2 that an analogous concept arises in theory networks: A logicatoms belief determines what logicatoms it is related too, and in turn these related logicatoms determine how the belief is updated.) The theory has been successful in predicting various observed phenomena (Mercury's precession rate, the cosmic microwave background, the expanding universe, the bending of light around massive objects to name but a few) and forms one of the foundations of modern physics.

Maxwell formulated his classical theory of electromagnetism after immersing himself in the accounts of Faraday's electrical researches. In tensor notation, Maxwell's equations of electromagnetism are [77]:

$$\begin{aligned} \nabla_{[a} F_{bc]} &= 0 \\ \nabla^a F_{ab} &= -4\pi j_b \end{aligned} \tag{4.1.2}$$

where F_{ab} is the spacetime tensor representing the electric and magnetic field and j^a is the current density 4-vector of electric charge. I have used the notation of square brackets around indices to denote the total antisymmetric part of the tensor. The central idea of (classical) geometrodynamics is to represent the electromagnetic field F_{ab} and electric current 4-density j^a purely in terms of curved spacetime. This means that any test charge should behave identically in the curved spacetime as it would according to Maxwell's electromagnetic equations. General relativity had done this exactly for Newton's theory of the gravitational force field. Classical geometrodynamics explored whether classical physics comprising gravitation, electromagnetism, non-quantised charge and non-quantised mass could be described in terms of empty curved space. Detail of the work reviewed below is available in [78].

The first breakthrough was made by Rainich, who showed under what conditions a curvature of spacetime can be regarded as due to an electromagnetic field. Further he also described how to find the field from the geometry. The Rainich conditions involved algebraic relations on the Ricci tensor R_{ab}, specifically

$$\begin{aligned} R_\mu{}^\alpha R_\alpha{}^\nu &= \delta_\mu{}^\nu \left(\frac{1}{4} R_{\alpha\beta} R^{\alpha\beta} \right) \\ R &= R^\alpha_\alpha = 0 \\ R_{00} &\geq 0 \end{aligned} \tag{4.1.3}$$

Now consider a field F_{ab} satisfying Maxwell's Equations 4.1.2. The electromagnetic energy-stress tensor is given by

$$T_{ab} = \frac{1}{4\pi}\left(F_{ac}F_b{}^c - \frac{1}{4}g_{ab}F_{de}F^{de}\right) \tag{4.1.4}$$

Solving the field equations 4.1.1 with T_{ab} specified by Equation 4.1.4 would give one a spacetime manifold with a Ricci tensor R_{ab} satisfying the Rainich conditions. Alternatively, given a solution to the field equations with the Ricci tensor satisfying the Rainich conditions, the electromagnetic tensor F_{ab} is specified up to a constant α, known as the 'complexion' of the electromagnetic field. Rainich (and Misner independently) showed that this constant is determined by the geometry using the equations

$$\begin{aligned} \alpha_\mu &= \frac{\sqrt{-g}\in_{\beta\alpha\mu\nu}\nabla^\beta R^{\alpha\rho}R_\rho{}^\nu}{R_{\sigma\tau}R^{\sigma\tau}} \\ \alpha &= \int \alpha_\mu dx^\mu \end{aligned} \tag{4.1.5}$$

Thus classical electromagnetism can be described in purely geometric terms. Einstein taught us that gravity could be described in purely geometric terms. Now electromagnetic waves are solutions to the Maxwell Equations 4.1.2 in empty space $j_a = 0$ [63]. Similarly, since gravitational waves are solutions to the source free field equations [77], one has gravitational and electromagnetic radiation described in purely geometric terms. The next step is to explain how mass and charge can be described using purely geometric terms. To solve this problem, Wheeler introduced the concept of a 'geon' (a gravitational electromagnetic entity): A geon is an object constructed out of electromagnetic radiation which holds itself together by its own gravitational attraction for a very long time. It owes its existence to curvature in spacetime. "Studied from a distance, such an object presents the same kind of gravitational attraction as any other mass. Moreover, it moves through space as a unit, and undergoes deflection by slowly varying fields of force just as does any other mass. In brief, the geon describes mass without mass" - Wheeler [78].

Wheeler attacked the problem of 'charge without charge' by considering a situation where lines of electric force thread through a 'wormhole' in spacetime, symbolically pictured below in Figure 4.1: "The two mouths appear to an observer with poor resolving power to be two equal and opposite electric charges" - Wheeler [78].

I conclude the review of classical geometrodynamics at this point. Firstly note that none of the entities described bear any resemblance to the elementary particles (governed by quantum theory) of observational science. They are all classical objects. However, the power of using geometry as a tool to model physics (at least in the classical realm) is evident. This is of singular importance in

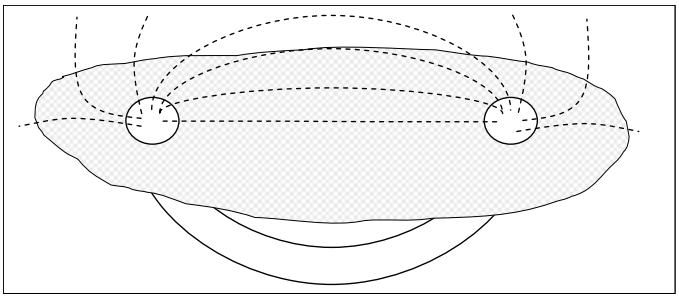

Figure 4.1: A wormhole with electric flux representing classical electric charge

the application analysis of inverted theory networks since I only have geometrical concepts to model physics.

It is interesting to note that geometry (although of a more generalised mathematical structure known as a fibre bundle [58]) also plays a role in quantum theory. This is experimentally captured in the Aharanov-Bohm (AB) effect. The AB experiment is schematically described in Figure 4.2 below. A beam of electrons comes in from the left and forms an interference pattern on the screen. A

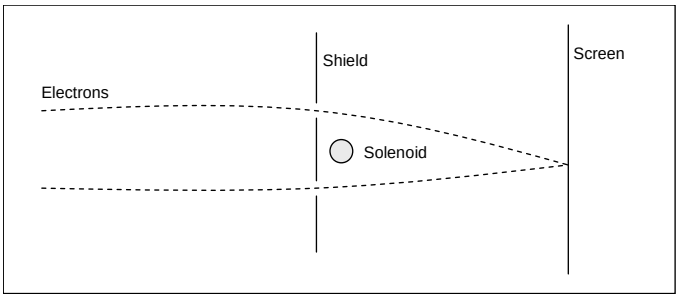

Figure 4.2: The Aharanov-Bohm experiment

solenoid is placed in the middle of the beam. A shield prevents the electrons from penetrating into the solenoid. Now Maxwell's electromagnetic field F_{ab} in equations 4.1.2 can be expressed in terms of a vector potential A_b by:

$$F_{ab} = \nabla_a A_b - \nabla_b A_a \tag{4.1.6}$$

In the AB experiment, a vector potential is chosen such that the field vanishes in the path of the electrons. Classically, the solenoid cannot have any influence on the electrons. The experiment shows however that even though the magnetic field is zero at the points in space through which the electrons travel, the resultant interference pattern depends on the potential A_b [70]. (The generalised

notion, of how the phase in a wavefunction influences the outcome, is known as Berry's phase [58].) This is theoretically explained using the geometric concept of holonomy which is described as the effect of parallel transporting a vector along a closed curve. In Riemannian geometry, this yields the Riemann curvature tensor R_{abcd}. The theoretical structure used in the AB effect is that of a fibre-bundle, but in both cases, the laws of physics is explained using the curvature of a space.

4.1.2 Pregeometries

Geometrodynamics fails to supply any "natural place for spin $\frac{1}{2}$ in general and for the neutrino in particular. There is no place in geometrodynamics for change in topology; therefore turn to pregeometry"[55]. A pregeometry is a basic structure that gives rise to the laws of physics. One therefore needs the pregeometry to look like a four dimensional Riemannian manifold on average. The first important point in this statement is the four dimensions. Einstein's field equations applied to a manifold with one, two or three dimensions demand flat space - not of any interest in this context. The second important point is the concept of continuity - spacetime in general relativity and particle physics, is a continuous manifold with continuous co-ordinates. Thus, the pregeometry must look continuous in some classical limit. This does not eliminate the option of discreteness. Einstein [21], in 1936 expressed the general feeling that "perhaps the success of the Heisenberg method points to a purely algebraic method of description of nature, that is, to the elimination of continuous functions from physics. Then however, we must give up, by principle, the spacetime continuum". Hawking on the other hand states that "although there have been suggestions that spacetime may have a discrete structure I see no reason to abandon the continuum theories that have been so successful." [36]. The pregeometry must incorporate dynamic topologies, a requirement emphasised by Wheeler [55] to model spin $\frac{1}{2}$ particles. It should explain the arrow of time. And finally, it should be fundamentally simple.

Since Wheeler mentioned the concept of a pregeometry, a vast amount of work has gone into formulating this structure, much of which is referenced in [27]. Some of this work that is relevant in my context will be detailed later. I do however need to mention one particular approach to the problem, namely Wheeler's original suggestion of "pregeometry as the calculus of propositions" [55]. Here, Wheeler hints at the fact that one might need to attack the problem from a completely different perspective, saying that it "may be hopeless to learn the basic operating principle of the universe, call it pregeometry or call it what one will, by any amount of work in general relativity and particle physics." Driven by the principle of simplicity, he asks: "What else can pregeometry be, one asks oneself, than the calculus of propositions?". However, in terms of a detailed description of the idea, he does say that "physics as a manifestation of logic ... is as yet not an idea, but an idea for an

idea." [55]

Is the pregeometry of physics, hypothesised by Wheeler [79] just one large inverted theory network? The objective of this section is to argue that it is in fact so. Towards this end, I will be using the geometrodynamics hypothesis [78]. I will thus define and analyse the discrete analogs to geometrical concepts defined on differentiable manifolds. This includes the 'dimension' of an inverted theory network, the 'metric' defining distance and volume measures for inverted theory networks, as well as the discrete analogs to concepts such as 'geodesics' and 'curvature'.

In terms of using inverted theory networks to model pregeometries, one must note that a similar structure was proposed by Cahill, Klinger and Kitto [11, 12]. They describe pregeometries in terms of 'process physics'. In particular, they hypothesise that the "solution to the end-game problem is to avoid the notion of things and rules; rather to use a bootstrapped self-referential system. Put simply, this models the universe as a self-organising and self-referential information system - 'information' denoting relationships as distinct from 'things' " [10]. No formal detail is given as to the exact structure of this space. Certain assumptions regarding how their structure evolves are made. In particular, an equation governing the relationships between the entities in their structure is hypothesised, resulting in a description of 'space and quantum physics'.

Goertzel et al [30] also start their pre-geometric modelling by assuming the universe is one large network. They define rules for the evolution of their network and show that they can generate clifford algebras from these 'event networks', giving one a basis in which to model quantum physics. It is interesting that this thesis has arrived at a similar end-point to the above examples, even though it originated from a knowledge modelling perspective.

To get a feel for what 'large' means in 'one large inverted theory network', consider the following: One would assume that the elementary length (say the distance between 2 related logicatoms) in a logicatom universes is one Planck length ($l_P = \sqrt{\frac{G\hbar}{c^3}} \approx 1.6 \times 10^{-33}$cm). This implies that 1cm^3 of space would consist of approximately 10^{100} logicatoms. So one is talking about a large (but finite) number of logicatoms interacting according to some local relation and transition rule. I hypothesise that these local interactions form the 'billions and billions of elementary quantum phenomena' and 'the laws and initial conditions of physics arise out of this chaos by the action of' natural selection.

4.1.3 The dimension of spacetime

The first problem I attack is in describing and analysing the 'dimension' of a logicatom universe. Towards this end, I need to define various metrics on inverted theory networks. I therefore exploit the similarity between logicatom universes and graph theory and reuse all the definitions available in this field.

Definition 4.1.1. A *graph* X consists of a vertex set $V(X)$ (the nodes) and an edge set $E(X)$ (the links), where an edge is an unordered pair of vertices of X. One says a graph is *directed* if an edge is an ordered pair of distinct vertices. Further, the graph is called *simple* if the edges comprise distinct pairs of vertices of X (i.e. an edge cannot start and terminate at the same vertex).

One can clearly see that a logicatom universe is a non-simple directed graph, where the logicatoms U are the vertices and the relation R specifies the edge set. This mapping just ignores the fact that there is more structure to a logicatom than the corresponding vertex. Graph theory is a large area of active research in mathematics and there are many books detailing all the concepts (subgraphs, automorphisms, homomorphisms) and properties regarding these structures [28]. In order to encapsulate the dynamic behaviour of inverted theory networks one would need to consider evolving graphs [19]. For my purposes, I borrow the following concepts:

Definition 4.1.2. Consider an inverted theory network $\mathcal{T}^{-1} = (\xi, R_f, T)$ over PropCal_n. A path between a logicatom $\mu(t)$ at time t and logicatom $\nu(t-k)$ at time $t-k$ (time is decremented as I move along possible futures) is defined as an ordered sequence of $k-1$ logicatoms together with their corresponding futures, $\{\eta_1^{\zeta_1}, \eta_2^{\zeta_2}, \ldots, \eta_{k-1}^{\zeta_{k-1}}\}$ satisfying the following properties:

(i) $\mu^{\zeta_1}(t-1) R \eta_1^{\zeta_1}(t-1)$

(ii) $\eta_i^{\zeta_{i+1}}(t-i-1) R \eta_{i+1}^{\zeta_{i+1}}(t-i-1)$ for every $i \in [1, k-2]$

(iii) $\eta_{k-1}^{\zeta_k}(t-k) R \nu^{\zeta_k}(t-k)$

Here R is the local relation induced by the relation generating function R_f. I say that the path $P_{\mathcal{F}}[\mu(t), \nu(t-k)]$ originates at logicatom μ at time t and terminates at ν at time $(t-k)$ along the possible future $\mathcal{F} = \{\zeta_1, \zeta_2, \ldots, \zeta_k\}$. The length of the path $P_{\mathcal{F}}[\mu(t), \nu(t-k)]$ is k and is denoted by $L(P_{\mathcal{F}}[\mu(t), \nu(t-k)])$.

The definition of a path allows me to inherit the most natural distance measure used in graph theory:

Definition 4.1.3. The distance between 2 logicatoms at time t for possible future $\mathcal{F}$ is defined by $d(\mathcal{F}, \mu, \nu, t) = \text{Min}\{L(P_{\mathcal{F}}[\mu_t, \nu_{t'}]) | \exists t' < t : P = P_{\mathcal{F}}[\mu_t, \nu_{t'}]$ is a path with future $\mathcal{F}\}$. In other words, it is the length of the shortest possible path or 'geodesic' between the 2 logicatoms along the specified possible future. If no such path exists I set $d(\mu, \nu, t) = \infty$.

Note that this is not a topological metric, since $d(\mu, \mu, t)$ is not necessarily 0. Once again, borrowing concepts from graph theory, I define the neighbourhood of a logicatom as [59]:

Definition 4.1.4. Let $\mu(t)$ be a logicatom in an inverted theory network $\mathcal{T}^{-1}$ at time t. A *k-neighbourhood* of $\mu(t)$ for possible future $\mathcal{F}$ is the set $N_k(\mathcal{F}, \mu, t) = \{\eta \mid d(\mathcal{F}, \mu, \eta, t) \leq k\}$. I define the *surface* of a neighbourhood for possible future $\mathcal{F}$ as the set $S_k(\mathcal{F}, \mu, t) = \{\eta \mid d(\mathcal{F}, \mu, \eta, t) = k\}$. The number of elements of a neighbourhood $\#N_k(\mathcal{F}, \mu, t)$ for possible future $\mathcal{F}$ is denoted by $V_k(\mathcal{F}, \mu, t)$. Similarly $D_k(\mathcal{F}, \mu, t) = \#S_k(\mathcal{F}, \mu, t)$ denotes the number of elements of the surface for possible future $\mathcal{F}$.

$D_k(\mathcal{F}, \mu, t)$ counts the total amount of logicatoms at a distance k from μ at time t for possible future $\mathcal{F}$. As will be seen, this will be used to get a measure of the 'dimension' of an inverted theory network along a possible future. I will motivate the definition of *surface sequences* using some simple examples.

Example 4.1.1. Consider the inverted theory network displayed in Figure 4.3. The diagram shows the relationships between the 11 logicatoms of an inverted theory network that has reached a fixed point (i.e. the relations between the logicatoms do not change. The time parameter t can thus be ignored in this example.) The surface of the neighbourhood at a distance of 1 from μ_5 is given by $S_1(\mu_5) = \{\mu_4, \mu_6\}$ resulting in $D_1(\mu_5) = 2$. Similarly, the surface of a neighbourhood at a distance of 2 from μ_5 is given by $S_2(\mu_5) = \{\mu_3, \mu_7\}$ resulting in $D_2(\mu_5) = 2$

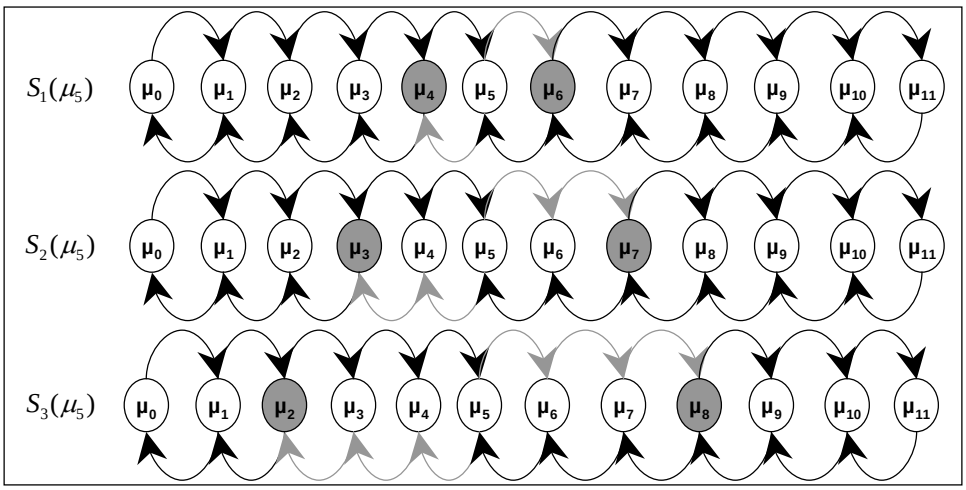

Figure 4.3: An inverted theory network having a dimension of 1

Figure 4.3 should convince the reader that the cardinality of the surfaces of every neighbourhood around μ_5 is 2. The fact that the cardinality of surfaces of all neighbourhoods around μ_5 is a constant, suggests that this inverted theory network 'approximates' a one dimensional space. This motivates the definition of the *surface sequence* of a logicatom at time t.

Definition 4.1.5. The *surface sequence* of a logicatom μ in an inverted theory network $\mathcal{T}^{-1}$ at time t for possible future $\mathcal{F}$ is defined by

$$\mathbf{\Sigma}(\mathcal{F}, \mu, t) = \{D_0(\mathcal{F}, \mu, t), D_1(\mathcal{F}, \mu, t), D_2(\mathcal{F}, \mu, t), \ldots, D_{L_{\max}}(\mathcal{F}, \mu, t)\}$$

where

$$L_{\max} = \max\{d(\mathcal{F}, \mu, \nu, t) \mid \forall \nu \in \mathcal{T} : d(\mathcal{F}, \mu, \nu, t) \neq \infty\}$$

The volume sequence is analogously defined by

$$\mathbf{\Omega}(\mathcal{F}, \mu, t) = \{V_0(\mathcal{F}, \mu, t), V_1(\mathcal{F}, \mu, t), V_2(\mathcal{F}, \mu, t), \ldots, V_{L_{\max}}(\mathcal{F}, \mu, t)\}$$

These sequences are powerful tools with which to analyse the geometry of discrete structures such as inverted theory networks. They give one the ability to define the analogous concepts available in differentiable manifolds. For example, in general relativity, homogeneity implies that no point in the universe has a privileged position. In a discrete structure, one could define this by saying that the surface sequence of every logicatom in the universe along a possible future $\mathcal{F}$ is identical. The example below shows the properties of the surface sequence of a 2-dimensional structure.

Example 4.1.2. The surface sequence of logicatom $\mu_{(2,2)}$ in the inverted theory network (that has reached a fixed point) represented in Figure 4.4 is $\mathbf{\Sigma}(\mathcal{F}, \mu_{(2,2)}, t) = \{4, 8, 12...\}$. The linear relationship of this sequence indicates an approximation to a two dimensional geometry, in the sense that it approximates a subset of $\mathbb{Z}^2$.

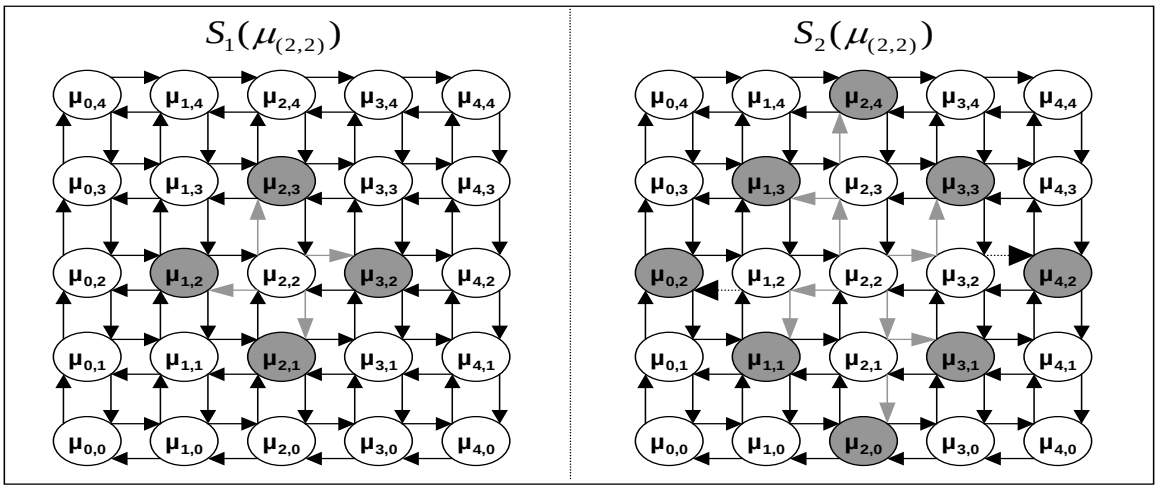

Figure 4.4: A '2 dimensional' inverted theory network

Intuitively, one expects the surface sequence of a three (four) dimensional structure to adhere to some quadratic (respectively cubic) relation. Alternative definitions of the dimension of discrete structures are available in [59], although I personally find the use of surface sequences intuitive. I now have the tools with which to analyse the dimension of an inverted theory network modelling a pregeometry. I am interested in answering the question posed by Wheeler: "How does the world manage to give the impression that it has the dimension of three?". Nagels' ingenuity in a great paper headed 'Space as a "Bucket of Dust" ' [57] provides one with the complete argument as to why certain inverted theory networks (viewed as a pregeometry) would give the impression that space has a dimension of three. Nagels answered the question: What is the most likely surface sequence for a given inverted theory

network along a fixed possible future? This problem was completely solved for the case where the probability that any two logicatoms being related is very small. I will review his work and detail his calculations - the main reason being that I believe his mathematical methodology to be the most relevant in understanding why pregeometries give rise to the 'classical physics' observed - a sought of 'correspondence principle' for pregeometries. Before I delve into Nagels' calculations, I would like to spend some time discussing his crucial constraint: *the probability that any 2 logicatoms are connected is small.* Consider the inverted theory networks governed by the affirmed implication relation and the constrained uniform substitution transition function i.e. the inverted theory network governed by natural selection. Equation 3.5.5 tells one the number of propositions in Φ_n that implies at least one affirmed variable. Thus for an inverted theory network over PropCal_n governed by this relation, the probability that any 2 logicatoms are related is given by $p = \frac{F(n)}{2^{2^n}}$. The table below shows the value of this ratio for the first few values of n.

n	2	3	4	5	6	7
p	0.375	0.148	0.014	7.57×10^{-5}	1.40×10^{-9}	3.80×10^{-19}

One sees that this ratio rapidly tends to 0. Thus for the number of logicatoms required in a pregeometry (the order of 10^{100} in 1cm^3 of space), one sees that the probability that any 2 logicatoms will be related is infinitesimally small. Consequently, the class of inverted theory networks that are regulated by natural selection as specified in Theorem 3.5.18 satisfy Nagels' constraint.

Returning to Nagels' work, I need to calculate the most likely surface sequence of a logicatom μ in an inverted theory network $\mathcal{T}^{-1}$ at some time t along a possible future $\mathcal{F}$:

$$\mathbf{\Sigma}(\mathcal{F}, \mu, t) = \{D_1(\mathcal{F}, \mu, t), D_2(\mathcal{F}, \mu, t), \ldots, D_{L_{\max}}(\mathcal{F}, \mu, t)\} \tag{4.1.7}$$

The elements of the surface sequence satisfy the following constraints:

$$\sum_{\lambda=0}^{L_{\max}} D_\lambda(\mathcal{F}, \mu, t) = V_{L_{Max}}(\mathcal{F}, \mu, t) \tag{4.1.8}$$

$$D_\lambda(\mathcal{F}, \mu, t) = 0 \text{ for } \lambda > L_{max} \tag{4.1.9}$$

$$D_\lambda(\mathcal{F}, \mu, t) > 0 \text{ for } 0 < \lambda \leq L_{max} \tag{4.1.10}$$

$L_{\max}$ is defined in Definition 4.1.5 and corresponds to the largest finite distance from logicatom μ at time t along possible future $\mathcal{F}$. $V_{L_{Max}}(\mathcal{F}, \mu, t)$ is the total number of logicatoms connected (in the sense that a path exists) to μ at time t along possible future $\mathcal{F}$. Equation 4.1.10 specifies this

connectivity by stating that there must exist at least one logicatom a distance λ from μ for λ less than the largest finite distance $L_{\max}$.
One is required to calculate the probability distribution of the spatial surface series i.e.

$$\mathcal{P}(\mu, t) = Prob[\{\Sigma(\mathcal{F}, \mu, t)\}] \tag{4.1.11}$$

All the subsequent derivations assume that one starts at logicatom μ at time t_0 for a fixed possible future $\mathcal{F}$. To simplify the notation, I omit the time and possible future parameter. To calculate the probability distribution 4.1.11, one notes that the following conditions hold:

(a) $\forall \eta \in S_{i+1}(\mu), \exists \nu \in S_i(\mu) : \nu R \eta$

This expresses the notion that every logicatom a distance $(i+1)$ from μ must be related to at least one logicatom a distance i from μ (at time $t_0 + i$ when one considers the definition of distance).

Now at any time, 2 logicatoms are related with probability p. Select $\eta \in S_{i+1}(\mu)$. It is not related to a logicatom in $S_i(\mu)$ with probability $1 - p = q$. It is not related to any logicatom in $S_i(\mu)$ with probability $q^{D_i(\mu)}$. It is related to at least 1 logicatom in $S_i(\mu)$ with probability $1 - q^{D_i(\mu)}$. This is true for every logicatom in $S_{i+1}(\mu)$. Thus the probability for a given surface sequence $\{D_1(\mu), D_2(\mu), \ldots, D_{L_{(max)}}(\mu)\}$ includes the factor $[1 - q^{D_i(\mu)}]^{D_{i+1}(\mu)}$

(b) $\forall \eta \in S_{i+1}(\mu), \forall \nu \in \bigcup_{k=0}^{k=i-1} S_k(\mu)$: It is not the case that $\mu R \eta$
This expresses the notion that each logicatom a distance $i+1$ from μ cannot be related to any logicatom at a distance $0, 1, 2, 3, ..., i-1$ from μ at time $t_0, t_0 + 1, \ldots, t_0 + i - 1$ respectively. Select an logicatom $\eta \in S_{i+1}(\mu)$. This is not related to a logicatom $\nu \in \bigcup_{k=0}^{k=i-1} S_k(\mu)$ (at any time) with probability q. This is not related to any logicatom $\nu \in \bigcup_{k=0}^{k=i-1} S_k(\mu)$ with probability $q^{\sum_{j=0}^{i-1} D_j(\mu)}$. This is true for every logicatom in $S_{i+1}(\mu)$. Thus the probability for a given surface sequence $\{D_1(\mu), D_2(\mu), \ldots, D_{L_{(max)}}(\mu)\}$ includes the factor $(q^{\sum_{j=0}^{i-1} D_j(\mu)})^{D_{i+1}(\mu)}$

(c) The logicatoms are all distinguishable and may be permuted between the sets of logicatoms at the same distances. This introduces a factor of $\frac{(N-1)!}{D_1! D_2! \ldots D_L!}$ in the probability for a given surface sequence $\{D_1(\mu), D_2(\mu), \ldots, D_{L_{(max)}}(\mu)\}$. Here $N = V_{L_{Max}(\mu)}$ is the total amount of logicatoms under consideration.

For notational simplicity, I denote $D_k = D_k(\mu)$ The probability for a specific surface sequence is

therefore the product of these three terms:

$$\begin{aligned}
\mathcal{P}(\Sigma(\mu)) &= \frac{(N-1)!}{D_1!D_2!...D_L!}\prod_{k=0}^{L-1}(1-q^{D_k})^{D_{k+1}}\prod_{k=1}^{L-1}(q^{\sum_{j=0}^{k-1}D_j})^{D_{k+1}} \\
&= \frac{(N-1)!}{D_1!D_2!...D_L!}p^{D_1}\prod_{k=1}^{L-1}\left[(1-q^{D_k})q^{\sum_{j=0}^{k-1}D_j}\right]^{D_{k+1}} \\
&= \frac{(N-1)!}{D_1!D_2!...D_L!}p^{D_1}\prod_{k=1}^{L-1}\left[q^{\sum_{j=0}^{k-1}D_j}-q^{\sum_{j=0}^{k}D_j}\right]^{D_{k+1}} \\
&= (N-1)!\frac{p^{D_1}}{D_1!}\prod_{i=2}^{L}\frac{1}{D_i!}\left[q^{\sum_{j=0}^{i-2}D_j}-q^{\sum_{j=0}^{i-1}D_j}\right]^{D_i} \qquad (4.1.12)
\end{aligned}$$

One is interested in the most likely probability distribution. Towards this end, one needs to derive the distribution of the surface sequence that maximises 4.1.12 and satisfies the constraints 4.1.8, 4.1.9 and 4.1.10. The calculation is made more tractable by taking the natural logarithm of 4.1.12 and using the Lagrange multiplier method to implement the constraints. One obtains

$$\begin{aligned}
F &= \ln\mathcal{P}+\lambda(\sum_{i=0}^{L}D_i-N) \\
&= \ln(N-1)!+D_1\ln(p)-\ln(D_1!) \\
&+ \sum_{i=2}^{L}D_i\ln(q^{\sum_{j=0}^{i-2}D_j}-q^{\sum_{j=0}^{i-1}D_j})-\sum_{i=2}^{L}\ln(D_i!)+\rho(\sum_{i=0}^{L}D_i-N) \qquad (4.1.13)
\end{aligned}$$

One proceeds to take the partial derivatives $\frac{\partial F}{\partial D_k}$. For $k=1$:

$$\frac{\partial F}{\partial D_1} = -\frac{D_2\ln q}{1-q^{D_1}}-\psi(D_1+1)+\ln(p)+\ln q\sum_{i=2}^{L}D_i \qquad (4.1.14)$$

For $k\in[2,L]$:

$$\frac{\partial F}{\partial D_k} = \ln q\sum_{j=0}^{k-2}D_j+\ln(1-q^{D_{k-1}})-\frac{D_{k+1}q^{D_k}\ln q}{1-q^{Dk}}+\ln q\sum_{i=k+2}^{L}D_i-\psi(D_k+1) \qquad (4.1.15)$$

This is obtained by splitting the term $\left[\sum_{i=2}^{L}D_i\ln(q^{\sum_{j=0}^{i-2}D_j}-q^{\sum_{j=0}^{i-1}D_j})\right]$ into the sum over the ranges $[2,k]\bigcup[k+1]\bigcup[k+2,l]$ and observing that

$$\frac{\partial(\sum_{i=2}^{L}\ln D_i!)}{\partial D_k} = \frac{\partial(\ln D_k!)}{\partial D_k} = \psi(D_k+1) \qquad (4.1.16)$$

is the digamma function, defined as the logarithmic derivative of the gamma function $\Gamma(z)$.

$$\psi(z) = \frac{d}{dz}\Gamma(z) = \frac{\Gamma^{'}(z)}{\Gamma(z)} \tag{4.1.17}$$

The gamma function (6.1.1 in [2])

$$\Gamma(z) = \int_0^\infty t^{z-1}e^{-t}dt \tag{4.1.18}$$

is the extension of the factorial to real and complex number arguments and satisfies the properties that

$$\Gamma(n) = (n-1)! \text{ for } n \in [1, 2, 3, \ldots] \tag{4.1.19}$$

Simplifying Equations 4.1.14, 4.1.15 and setting $\frac{\partial F}{\partial D_k} = 0$, results in

$$\ln(1 - q^{D_{k-1}}) - \psi(D_k + 1)(N - D_k - D_{k-1}) \ln q - D_{k+1}\frac{\ln q}{1 - q^{D_k}} + \rho = 0 \tag{4.1.20}$$

for all $k \in [1, L]$. Solving for D_k in 4.1.20 and 4.1.8 would give one the most likely surface sequence for any theory network.

One can solve this in the case of $p \ll 1$ and $D_k \gg 1$. The asymptotic expansion for the digamma function is given by

$$\psi(z + 1) = \ln z + \frac{1}{2z} - \sum_{n=1}^{\infty} \frac{B_{2n}}{2nz^{2n}} \tag{4.1.21}$$

where B_n are the Bernoulli numbers defined by the identity

$$\frac{x}{e^x - 1} \equiv \sum_{n=0}^{\infty} \frac{B_n x^n}{n!} \tag{4.1.22}$$

One approximates the $\psi(D_k + 1)$ term by

$$\psi(D_k + 1) \approx \ln D_k + \frac{1}{2D_k} + \mathcal{O}\left(\frac{1}{D_k^2}\right) \approx \ln D_k \tag{4.1.23}$$

The $\ln(q^{D_{k-1}})$ term is approximated using the Taylor expansion

$$q^{D_{k-1}} = (1 - p)^{D_{k-1}} \approx 1 - D_{k-1}p + \frac{1}{2}D_{k-1}(D_{k-1} - 1)p^2 + \mathcal{O}(p^3) \tag{4.1.24}$$

$$\begin{aligned} \ln(1 - q^{D_{k-1}}) &\approx \ln\left[D_{k-1}p\left(1 - \frac{(D_{k-1} - 1)}{2}p\right)\right] \\ &\approx \ln D_{k-1} + \ln p + \frac{D_{k-1}p}{2} - \frac{p}{2} + (O)(p^2) \end{aligned} \tag{4.1.25}$$

Similarly, the $D_{k+1}\frac{\ln q}{1-q^{D_k}}$ term is approximated to

$$\begin{aligned} -D_{k+1}\frac{\ln q}{1-q^{D_k}} &\approx \frac{D_{k+1}}{D_k}\left[\frac{1}{1-\frac{(D_k-1)p}{2}}\right] \\ &\approx \frac{D_{k+1}}{D_k}+\frac{D_{k+1}p}{2}+\mathcal{O}(p^2) \end{aligned} \tag{4.1.26}$$

using $\ln q = \ln(1-p) \approx -p + (O)\left(p^2\right)$. Substituting these approximations into 4.1.20 and grouping all the D_k terms, one obtains

$$\ln\left(\frac{D_{k-1}}{D_k}\right)+\frac{D_{k+1}}{D_k}+\frac{p}{2}\left(D_{k-1}+2D_k+Dk+1\right)=\text{constant} \tag{4.1.27}$$

The final approximation is that $pD_k \approx 0$. Incorporating this allows one to arrive at the equation

$$\frac{D_{k+1}}{D_k}=\ln\left(\frac{D_{k-1}}{D_k}\right)+1-\eta \tag{4.1.28}$$

where the constant is rewritten as $1-\eta$. Consider the case where $\eta \leq 0$. One has $D_0 = 1$ (This assumes that the logicatom is not self referential. The approximation will however still hold if the logicatom was self referential.) and $D_1 \geq 1$. Then $\ln\left(\frac{D_0}{D_1}+1-\eta\right) \geq 1$. One can thus deduce $D_2 \geq D_1$. Using this inductive argument, one deduces that D_k is a monotonically increasing function, violating the constraint specified by Equation 4.1.9. If $\eta > 0$, one has a function with a single maximum. Assume this maximum occurs at $k = m$. Define $n = k - m$. Then D_0 is the maximum element in the sought after surface sequence. One now has

$$D_{n+1}=D_n\left(\ln(D_n)-\ln(D_{n-1})+1-\eta\right) \tag{4.1.29}$$

Further defining $y = \beta n$ with $\beta = \sqrt{\frac{\eta}{2}}$ and

$$w(y)=w(\beta n)=D_n \tag{4.1.30}$$

so that

$$w(y\pm\beta)=w(\beta n\pm\beta)=w(\beta(n\pm 1))=D_{n\pm 1} \tag{4.1.31}$$

and substituting this into 4.1.28 results in

$$w(y+\beta)=w(y)\left[\ln w(y)-\ln(w(y-\beta))+1-2\beta^2\right] \tag{4.1.32}$$

In order to solve this equation, substitute the Taylor expansion for $w(y+\beta)$ and $\ln(w(y))$ and thereafter retain all terms up to the first order of β. This gives one the differential equation:

$$w''(y)+2w(y)-\frac{w'(y)^2}{2w(y)}+\frac{w'(y)w''(y)}{2w(y)}\beta-\frac{w'(y)^3}{3w(y)^2}\beta+\mathcal{O}(\beta^2)=0 \tag{4.1.33}$$

subject to the initial conditions

$$w(0) = W_0 \tag{4.1.34}$$

$$w'(0) = 0 \tag{4.1.35}$$

To solve this differential equation, one considers a perturbation expansion around β i.e.

$$w(y) = \sum_{k=0}^{\infty} w_k(y)\beta^k \tag{4.1.36}$$

Substituting this expansion into 4.1.33 and again retaining all terms up to the first order of β, one obtains the constant and first order coefficient in the expansion as

$$\begin{aligned} & \left[w_0'' + 2w_0 - \frac{w_0^2}{2w_0}\right] \\ + & \left[w_1'' - \frac{w_0'w_1'}{w_0} - \frac{(w_0')^2 w_1}{2(w_0)^2} + 6w_1 + 2\frac{w_0''w_1}{w_0} + \frac{w_0'w_0''}{2w_0} - \frac{(w_0')^3}{3(w_0)^2}\right]\beta \\ = & \ 0 \end{aligned} \tag{4.1.37}$$

This is obtained by first multiplying 4.1.33 by $w(y)^2$ and then substituting the expansion 4.1.36. Now since this is true for all β, one has

$$w_0'' + 2w_0 - \frac{w_0^2}{2w_0} = 0 \tag{4.1.38}$$

$$w_1'' - \frac{w_0'w_1'}{w_0} - \frac{(w_0')^2 w_1}{2(w_0)^2} + 6w_1 + 2\frac{w_0''w_1}{w_0} + \frac{w_0'w_0''}{2w_0} - \frac{(w_0')^3}{3(w_0)^2} = 0 \tag{4.1.39}$$

subject to the initial condition

$$\begin{aligned} w_0 = W_0 \quad & w_1(0) = 0 \\ w_0'(0) = 0 \quad & w_1'(0) = 0 \end{aligned} \tag{4.1.40}$$

The unique solution of 4.1.38 is given by

$$w_0 = W_0 \cos^2(y) \tag{4.1.41}$$

Equation 4.1.39 is equivalent to

$$w_1'' - \frac{w_0'w_1'}{w_0} + \frac{(w_0')^2 w_1}{2(w_0)^2} + 2w_1 + \frac{w_0'w_0''}{2w_0} - \frac{(w_0')^3}{3(w_0)^2} = 0 \tag{4.1.42}$$

which can be confirmed by substituting the following identity, derivable from Equation 4.1.41

$$\frac{2w_0''w_1}{w_0} = \frac{(w_0')^2 w_1}{w_0^2} - 4w_1 \tag{4.1.43}$$

for the $2\frac{w_0'' w_1}{w_0}$ term in Equation 4.1.39.
The unique solution is given by

$$w_1(y) = \frac{1}{3} W_0 \sin(2y) \ln(\cos y) \tag{4.1.44}$$

resulting in

$$w(y) = W_0 \left(\cos^2 y + \frac{1}{3} \beta \sin(2y) \ln(\cos y) \right) \tag{4.1.45}$$

One returns to the variable n by substituting $y = \beta n$ and obtaining

$$D_n = D_{\max} \left[\cos^2(\beta n) + \frac{\beta \sin(2\beta n) \ln(\cos \beta n)}{3} \right] \tag{4.1.46}$$

One returns to the original variables by noting that if one sets $n_\epsilon = -\frac{\pi}{2\beta} + \epsilon$, then for small positive ϵ

$$\lim_{\epsilon \to 0^+} D_{n_\epsilon} = \lim_{\epsilon \to 0^+} D_{Max} \left[\epsilon^2 \beta^2 - \frac{2\beta^2 \epsilon \ln(\beta \epsilon)}{3} \right] = 0 \tag{4.1.47}$$

where I have used $\lim_{x \to 0^+} x \ln x = 0$. Thus $y = \beta n = \beta k - \frac{\pi}{2}$ which on substitution into 4.1.46 gives one

$$D_k \simeq D_{\max} \left[\sin^2(\beta k) - \frac{\beta \sin(2\beta k) \ln(\sin \beta)}{3} \right] \tag{4.1.48}$$

Further, one notes that for $\beta < \exp\left(-\frac{3}{2}\right)$, the second zero occurs between $k = \frac{\pi}{\beta}$ and $k = \frac{\pi}{\beta} - 1$, so one can define $L = \frac{\pi}{\beta}$, the furthest point (and thus the total amount of elements in the surface sequence), resulting in

$$D_k \simeq D_{\max} \left[\sin^2(\frac{\pi k}{L}) - \frac{\beta \sin(\frac{2\pi k}{L}) \ln(\sin \frac{\pi}{L})}{3} \right] \tag{4.1.49}$$

Finally note that $\sum_{k=0}^{L} \sin^2\left(\frac{\pi k}{L}\right) = \frac{L}{2}$. Further $\sum_{k=0}^{L} \sin\left(\frac{2\pi k}{L} \ln(\sin(\frac{k\pi}{L}))\right) = 0$. This result is obtained by noting that the k^{th} term cancels the $(L-k)^{th}$ since
$\sin\left(\frac{2\pi(L-k)}{L} \ln \sin\left(\frac{(L-k)\pi}{L}\right)\right) = -\sin\left(\frac{2\pi k}{L} \ln \sin\left(\frac{k\pi}{L}\right)\right)$. One thus obtains $D_{\max} = \frac{2N}{L}$ on satisfying the constraint 4.1.8, resulting in

$$D_k \simeq \frac{2N}{L} \left[\sin^2(\frac{\pi k}{L}) - \frac{\pi}{3L} \sin(\frac{2\pi k}{L}) \ln(\sin(\frac{\pi k}{L})) \right] \tag{4.1.50}$$

Note that the first term dominates for $\frac{k}{L} \geq \frac{\ln(L)}{L}$.
The surface area of a sphere for a three dimensional manifold of constant positive curvature is given by

$$S(k) = \frac{4L^2}{\pi} \sin^2\left(\frac{\pi k}{L}\right) \tag{4.1.51}$$

In summary, what Nagels' showed using the above derivation was that inverted theory networks over PropCal_n ($n \gg 1$) governed by the constrained uniform substitution function and the affirmed implication relation give one the impression of a 3-dimensional curved space when viewed on the large scale as a pregeometry. Together with the single time dimension, he has explained why we live in a 4-dimensional curved spacetime.

4.2 Quantum theory

'Surely you're joking Mr Feynman!' - R.P. Feynman

This section reviews work done by Feynman [24], [25] showing how the Schrödinger and Dirac equations can be derived using path integrals. The reason I have included this section is to provide the non-physicist with the preliminaries required to understand the arguments put forward in Section 4.2.2.

The beginning of the 20^{th} century saw experimental physics lead the way in breaking down the notion that classical mechanics was an adequate theoretical framework for describing atomic structures. Various experiments showed the conflicts that arose using the classical interpretations of describing for example, electrons as particles and light using waves. The photo-electric effect showed light (viewed as waves modelled using Maxwell's Electromagnetic theory) of a particular frequency aimed at a charged plate would effectively knock off electrons from the plate's surface. On the other hand, electrons (viewed as classical particles and modelled using classical mechanics) displayed characteristic wave-behaviour in the double slit experiment. The interference patterns obtained could be explained by a wave theory but not a particle theory [35]. Quantum Physics (born out of quantum mechanics) was developed to explain these and other phenomena.

Quantum Mechanics changed our perception of physical processes. The theory showed that experiments exist in which the observables can only be represented as a probability distribution of various outcomes. The fact that the exact outcome of the experiment was fundamentally unpredictable was not due to any unknowns in the experiment (as would be the case in statistical mechanics), but was an actual property of the physics [25]. Furthermore, the mathematical laws governing these probabilities were different from those of the classical probability theory of Laplace. For example, in the double slit experiment performed with particles, the probability that one observes a particle at a point x on the wall with both slits open is not the sum of the probabilities of observing the particle at that same point x with only slit 1 open and only slit 2 open. A *probability amplitude* was associated with every observable. When there was more than one alternative in an experiment (go through slit one *or* slit two and arrive at point x), the probability amplitudes interfered. These

concepts also elevated the observer to a new position in physics. Unlike classical mechanics, the observer in the experiment affected the outcome of the experiment. This was elegantly stated by Heisenberg as the *Uncertainty Principle*. 'Any determination of the alternative taken by a process capable of following more than one alternative destroys the interference between alternatives' [25]. Thus closing slit 2 (and determining that the particles should all go through slit 1) results in no interference pattern.

From the quantum mechanical perspective, all information regarding possible observations in an experiment resided in the *wave function* ϕ that satisfied the Schrödinger equation:

$$-\frac{\hbar}{i}\frac{\partial\phi}{\partial t} = (-\frac{\hbar^2}{2m}\nabla^2 + V)\phi \tag{4.2.1}$$

Any observation resulted in localising the extent of the wave function in space. The consequence was that only a discrete set of wavelengths - and correspondingly, only a discrete set of frequencies could occur - "Localisation leads to quantisation" [35]. The *correspondence principle* further showed that classical mechanics was an approximation to quantum mechanics. The expectation values of the observables in quantum mechanics behaved in the same way as the observables themselves did in classical mechanics, as the size of the particles involved increased.

4.2.1 Path Integrals

Path integrals were introduced by Dirac in the 1930's, but their mathematical impreciseness initially discouraged their serious application to quantum theory. The integrals could not be rigorously defined using a measure on the space of all paths. Cameron's theorem [42] stated that no finite Lebesque measure existed for these path integrals. Feynman reintroduced path integrals [24] in a seminal paper showing how non-relativistic quantum mechanics could be reformulated using this approach. Despite the mathematical impreciseness, the manipulations suggested by the approach provided valuable insights and useful approximation schemes, such as the perturbative expansions represented by Feynman diagrams in quantum field theory. Another great feature was the simple explanation the formalism provided in describing how the classical limit arose from the quantum theory i.e. the correspondence principle. The formalism also portrayed a relatively simple process for quantising classical theories. The concept of summing over all possible alternatives, as opposed to promoting elements in the classical phase space to quantum operators on a Hilbert space, is conceptually easier to understand, although it may be argued that this is a matter of taste.

The path-integral approach to quantum theory [25, 24] incorporates as its starting point the basic quantum mechanical amplitude for a complete history of the system under consideration. In the simple example of a single particle in one dimension, a history is a path $x(t)$, and the amplitude

for the particle to travel on this particular path is given by $e^{\frac{i}{\hbar}\mathcal{S}[x(t)]}$ where $\mathcal{S}[x(t)]$ is the action functional describing the classical dynamics of the system [69]. From this basic amplitude and the principle that one sums over all *unobserved* histories of the system, one can compute the quantum mechanics of the system.

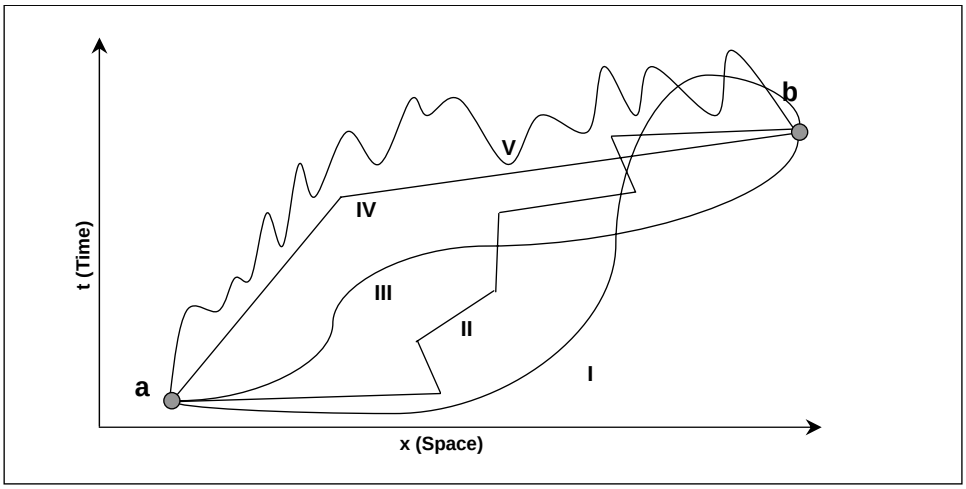

Figure 4.5: Various possible paths for a particle in one spatial dimension

For example, if the particle was observed at x_a at time t_a, the probability amplitude to observe the particle at x_b at a later time t_b is obtained by summing over all possible paths starting at (x_a, t_a) and ending at (x_b, t_b):

$$K(x_b, t_b|x_a, t_a) = \mathcal{N} \int_{\mathcal{C}} \mathcal{D}[x(t)] e^{\frac{i}{\hbar}\mathcal{S}[x(t)]} \tag{4.2.2}$$

Here $\mathcal{N}$ is a normalisation factor (needed due to the fact that no well defined measure exists for this integral) and $\mathcal{C}$ represents the set of all paths joining (x_a, t_a) to (x_b, t_b).
The quantity $K(x_b, t_b|x_{1a}, t_a)$ called the propagator, summarises the entire quantum mechanics of the system. Given a wave function $\psi(x, t_1)$ specifying the state of a system at time t_1, the wave function at a later time t_2 will be given by

$$\psi(y, t_2) = \int K(y, t_2|x, t_1)\psi(x, t_1)dx \tag{4.2.3}$$

Conceptually, this principle is easily extended to deal with systems with an infinite number of degrees of freedom, such as the electromagnetic field in Minkowski spacetime. The essence of the idea remains the same in that, the probability amplitude of observing an electromagnetic field configuration $B''(x, y, z)$ on a space-like hypersurface Σ'', given that one observed the configuration

$B'(x, y, z)$ on a hypersurface Σ', is given by a sum over all histories which are compatible with the initial and final field configurations on the respective surfaces.

$$K(B'', \Sigma''|B', \Sigma') = \mathcal{N} \int_{\mathcal{C}} \mathcal{D}[A_\mu] e^{\frac{i}{\hbar}\mathcal{S}[A_\mu]} \tag{4.2.4}$$

Here $\mathcal{S}$ is the classical action for the electromagnetic field. I will now proceed to show how Feynman [25] calculated the propagator $K(x_b, t_b|x_a, t_a)$ for the free particle, showing the equivalence to the Schrödinger equation.

The classical action functional for a free particle is given by

$$S[x(t)] = \int_{t_a}^{t_b} L(\dot{x}, x, t)dt \tag{4.2.5}$$

where the integral is along the path $x(t)$ from $x(t_a)$ to $x(t_b)$. L, the Lagrangian of a free particle is given by

$$L(\dot{x}, x, t) = \frac{1}{2}m\dot{x}^2 \tag{4.2.6}$$

In order to sum over all paths joining (x_a, t_a) to (x_b, t_b), divide the time into steps of length ϵ. This gives one the series of values $(t_a, t_a + \epsilon, \ldots, t_a + (n-1)\epsilon, t_b)$. At each of these time points t_i, choose a value x_i and construct a path by linearly joining these points as shown in Figure 4.6. Integrating over all paths is now equivalent to integrating over all values of x_i for $i \in \{1, \ldots, n-1\}$

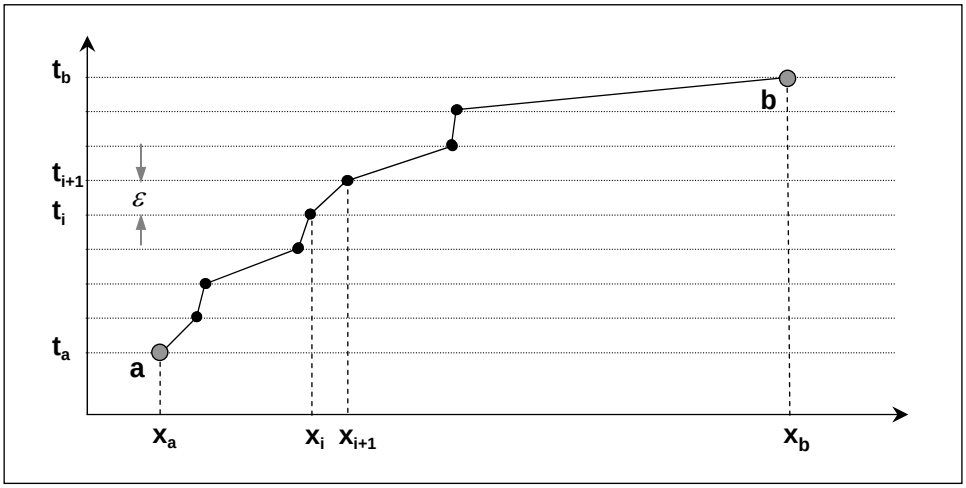

Figure 4.6: Constructing the sum over all paths

(specifically not integrating over x_a and x_b since these are fixed points) and taking the limit as

$\epsilon \to 0$. A normalisation constant A is defined to allow the limit to exist. Feynman shows how the value of A can be calculated using Equation 4.2.3 and choosing $t_2 = t_1 + \epsilon$ with $\epsilon \ll 1$
The path integral is then just

$$K(x_a, t_a; x_b, t_b) = \lim_{\epsilon \to 0} \frac{1}{A} \int \int \dots \int e^{i/\hbar S(a,b)} \frac{dx_1}{A} \dots \frac{dx_{N-1}}{A} \tag{4.2.7}$$

with

$$A = \frac{2\pi i\hbar\epsilon}{m}^{1/2} \tag{4.2.8}$$

For $L = \frac{1}{2}m\dot{x}^2$, the action S for the path P defined by the series
$((x_a, t_a), (x_1, t_1), \dots, (x_{n-1}, t_{n-1}), (x_b, t_b))$ where $t_k = t_a + k\epsilon$ is given by

$$\begin{aligned} S(P) &= \int_{t_a}^{t_b} \frac{1}{2} m\dot{x}^2 dt \\ &= \frac{1}{2} m \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \left(\frac{x_k - x_{k-1}}{\epsilon} \right)^2 dt \\ &= \frac{m}{2\epsilon} \sum_{k=1}^{n} (x_k - x_{k-1})^2 \end{aligned} \tag{4.2.9}$$

resulting in

$$K(x_a, t_a; x_b, t_b) = \left(\frac{2\pi\ i\hbar\epsilon}{m} \right) \lim_{\epsilon \to 0} \int \int \dots \int \exp \left[\frac{im}{2\hbar\epsilon} \sum_{k=1}^{n} (x_k - x_{k-1})^2 \right] dx_1 \dots dx_{n-1} \tag{4.2.10}$$

Integrating the gaussians and taking the limit results in

$$K(x_b, t_b; x_a, t_a) = \left[\frac{2\pi i\hbar(t_b - t_a)}{m} \right]^{-\frac{1}{2}} \exp \frac{im(x_b - x_a)^2}{2\hbar(t_b - t_a)} \tag{4.2.11}$$

Equivalence with the Schrödinger equation is established on noticing that this is just the Green's function for Equation 4.2.1.

For the case of the Dirac equation, Feynman states [25] that the propagator for a relativistic particle starting at spacetime point (x_a, t_a) and ending at (x_b, t_b) is given by

$$K(b, a) = \sum_R N(R)(i\varepsilon)^R \tag{4.2.12}$$

where $N(R)$ is the number of paths joining a and b with R corners / reversals and ε is the length of a single step. Figure 4.7 shows three of these possible paths.
In order to calculate the propagator one specifies a move to the right (respectively left) in Figure 4.7 as a move in the positive (respectively negative) direction. The set of all paths that start at a

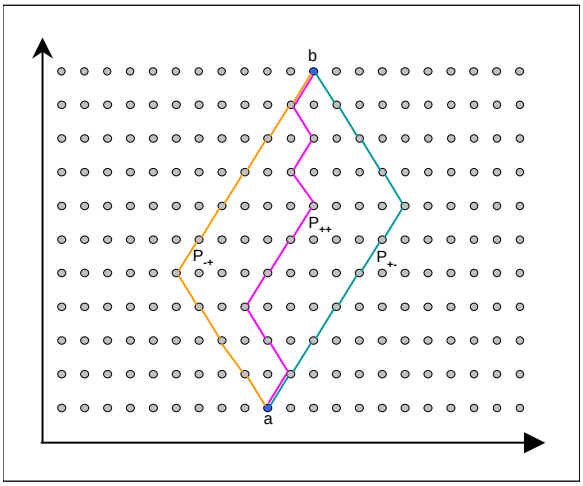

Figure 4.7: Possible paths in the path integral formulation of the Dirac Equation

moving in the positive direction, and end at b with a move in the positive direction will be denoted by P_{++}. Similarly P_{+-} denotes the paths starting with a move in the positive direction and ending with a move in the negative direction. In order to count the total amount of paths with R reversals between (x_a, t_a) and (x_b, t_b), one splits the paths into the 4 disjoint sets P_{++}, P_{+-}, P_{-+}, and P_{--}. Let ε be the length of the a single move. Define $T = (t_b - t_a)$ and $X = d(x_a, x_b)$. Further, let

$$n_X = \frac{X}{\varepsilon} \tag{4.2.13}$$

$$n_T = \frac{T}{\varepsilon} \tag{4.2.14}$$

$$\alpha = \frac{n_T + n_X}{2} \tag{4.2.15}$$

$$\beta = \frac{n_T - n_X}{2} \tag{4.2.16}$$

One clearly has $n_X \leq n_T$, otherwise no paths between the points would exist. In the calculations, one assumes $n_X < n_T$ i.e. there are no paths with 0 reversals.

Now using combinatorial counting arguments, one has

$$N_{++}(2m) = \#P_{++} = \begin{pmatrix} \alpha - 1 \\ m \end{pmatrix} \begin{pmatrix} \beta - 1 \\ m - 1 \end{pmatrix} \tag{4.2.17}$$

$$N_{--}(2m) = \#P_{--} = \begin{pmatrix} \alpha - 1 \\ m - 1 \end{pmatrix} \begin{pmatrix} \beta - 1 \\ m \end{pmatrix} \tag{4.2.18}$$

$$N_{-+} = N_{+-}(2m-1) = \#P_{+-} = \begin{pmatrix} \alpha - 1 \\ m - 1 \end{pmatrix} \begin{pmatrix} \beta - 1 \\ m - 1 \end{pmatrix} \tag{4.2.19}$$

with $m \in [1, 2, 3, ...]$

Splitting the propagator $K(a,b)$ into $K_{++}(a,b)$,$K_{+-}(a,b)$,$K_{-+}(a,b)$ and $K_{--}(a,b)$ one obtains:

$$\begin{aligned} K_{-+}(a,b) = K_{+-}(a,b) &= \sum_{m=1}^{\infty} N_{+-}(2m-1)\,(i\varepsilon)^{2m-1} \\ &= \sum_{m=1}^{\infty} \left(\begin{array}{c} \alpha-1 \\ m-1 \end{array} \right) \left(\begin{array}{c} \beta-1 \\ m-1 \end{array} \right) (i\varepsilon)^{2m-1} \\ &= i\varepsilon \sum_{k=0}^{\infty} \left(\begin{array}{c} \alpha-1 \\ k \end{array} \right) \left(\begin{array}{c} \beta-1 \\ k \end{array} \right) \left(-\varepsilon^2\right)^k \end{aligned} \tag{4.2.20}$$

Similarly

$$K_{++}(a,b) = -\varepsilon^2 \sum_{k=0}^{\infty} \left(\begin{array}{c} \alpha-1 \\ k+1 \end{array} \right) \left(\begin{array}{c} \beta-1 \\ k \end{array} \right) \left(-\varepsilon^2\right)^k \tag{4.2.21}$$

and

$$K_{--}(a,b) = -\varepsilon^2 \sum_{k=0}^{\infty} \left(\begin{array}{c} \alpha-1 \\ k \end{array} \right) \left(\begin{array}{c} \beta-1 \\ k+1 \end{array} \right) \left(-\varepsilon^2\right)^k \tag{4.2.22}$$

Expressing the combinatorial in terms of the gamma function 4.1.18

$$\left(\begin{array}{c} j \\ k \end{array} \right) = \frac{\Gamma(j+1)}{\Gamma(k+1)\Gamma(j-k+1)} \tag{4.2.23}$$

allows one to use the combinatorial relation

$$\begin{aligned} \left(\begin{array}{c} j \\ k \end{array} \right) &= (-1)^k \left(\begin{array}{c} k-j-1 \\ k \end{array} \right) \\ &= (-1)^k \frac{\Gamma\left(k-j\right)}{\Gamma\left(k+1\right)\Gamma\left(-j\right)} \end{aligned} \tag{4.2.24}$$

to obtain

$$\begin{aligned} & K_{-+}(a,b) = K_{+-}(a,b) \\ = \;& i\varepsilon \frac{1}{\Gamma(1-\alpha)\Gamma(1-\beta)} \sum_{k=0}^{\infty} \frac{\Gamma(k+1-\alpha)\Gamma(k+1-\beta)}{\Gamma(k+1)} \frac{\left(-\varepsilon^2\right)^k}{k!} \\ = \;& i\varepsilon \; {}_2F_1(1-\alpha, 1-\beta, 1, -\varepsilon^2) \end{aligned} \tag{4.2.25}$$

where ${}_2F_1$ is the hypergeometric function defined by (15.1.1 in [2])

$$ {}_2F_1(a,b;c;z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{n=0}^{\infty} \frac{\Gamma(a+n)\Gamma(b+n)}{\Gamma(c+n)} \frac{z^n}{n!} \tag{4.2.26} $$

Similarly one obtains

$$ K_{++}(a,b) = -\frac{(\alpha-1)\varepsilon^2}{2}\ {}_2F_1(2-\alpha, 1-\beta, 2, -\varepsilon^2) \tag{4.2.27} $$

and

$$ K_{--}(a,b) = -\frac{(\beta-1)\varepsilon^2}{2}\ {}_2F_1(1-\alpha, 2-\beta, 2, -\varepsilon^2) \tag{4.2.28} $$

To show that this is in fact the discrete version of the propagator of the Dirac Equation in one dimension,

$$ -i\sigma_z \frac{\partial \psi}{\partial x} - m\sigma_x \psi = i\frac{\partial \psi}{\partial t} \tag{4.2.29} $$

one proceeds to take the limit as $\varepsilon \to 0$. Firstly, since the propagators $K_{\mp\mp}$ vanish at every other lattice point, they must be divided by 2ε in order to obtain their continuum form [51]. Further using the hypergeometric relation (15.3.3 in [2])

$$ {}_2F_1(a,b;c;z) = (1-z)^{c-a-b}\ {}_2F_1(c-a, c-b; c; z) \tag{4.2.30} $$

one obtains

$$ K_{-+}(a,b) = K_{+-}(a,b) = \frac{i}{2}\left(1+\varepsilon^2\right)^{\alpha+\beta-1}\ {}_2F_1(\alpha,\beta,1,-\varepsilon^2) \tag{4.2.31} $$

$$ K_{++}(a,b) = -\frac{1}{4}(\alpha-1)\varepsilon\left(1+\varepsilon^2\right)^{\alpha+\beta-1}\ {}_2F_1(\alpha,\beta+1,2,-\varepsilon^2) \tag{4.2.32} $$

$$ K_{--}(a,b) = -\frac{1}{4}(\beta-1)\varepsilon\left(1+\varepsilon^2\right)^{\alpha+\beta-1}\ {}_2F_1(\alpha+1,\beta,2,-\varepsilon^2) \tag{4.2.33} $$

Substituting the original variables

$$ \alpha = \frac{n_t + n_x}{2} = \frac{1}{\varepsilon}\frac{(X+T)}{2} \tag{4.2.34} $$

$$ \beta = \frac{n_t - n_x}{2} = \frac{1}{\varepsilon}\frac{(T-X)}{2} \tag{4.2.35} $$

$$ \varepsilon^2 = \frac{T^2 - X^2}{4\alpha\beta} \tag{4.2.36} $$

in the K_{-+} and K_{+-} propagators, one obtains

$$\begin{aligned} K_{-+}(a,b) &= K_{+-}(a,b) \\ &= i\left(1+\frac{(T^2-X^2}{4\alpha\beta})\right)^{\alpha+\beta-1}\ {}_2F_1(\alpha,\beta,1,-\frac{(T^2-X^2}{4\alpha\beta}) \end{aligned} \tag{4.2.37}$$

Finally, using the hypergeometric relation with the Jacobian (9.1.70 in [2])

$$J_\nu(z) = \frac{(\frac{1}{2}z)^\nu}{\Gamma(\nu+1)} \lim_{\lambda,\mu\to\infty}\ {}_2F_1(\lambda,\mu;\nu+1;-\frac{z^2}{4\lambda\mu}) \tag{4.2.38}$$

and the fact that

$$\lim_{\lambda,\mu\to\infty}\left(1+\frac{x^2}{4\lambda\mu}\right)^{\lambda+\mu+y} = 1 \tag{4.2.39}$$

for any constants x, y, one obtains the continuum form of the propagator as $(\varepsilon\to 0)\equiv(\alpha,\beta\to\infty)$ to be

$$\begin{aligned} &\lim_{\varepsilon\to 0} K_{-+}(a,b) \\ =\ &\lim_{\varepsilon\to 0} K_{+-}(a,b) \\ =\ &\lim_{\alpha,\beta\to\infty}\frac{i}{2}\left(1+\frac{(T^2-X^2)}{4\alpha\beta}\right)^{\alpha+\beta-1}\ {}_2F_1(\alpha,\beta,1,-\frac{(T^2-X^2}{4\alpha\beta}) \\ =\ &\frac{i}{2}J_0(\sqrt{T^2-X^2}) \end{aligned} \tag{4.2.40}$$

In the case of K_{--}, one uses the hypergeometric relation (15.2.17 in [2])

$${}_2F_1(a+1,b;c,z) = \frac{(1+a-c)}{a}\ {}_2F_1(a,b;c,z) + \frac{(c-1)}{a}\ {}_2F_1(a,b;c-1,z) \tag{4.2.41}$$

to obtain

$$\begin{aligned} K_{--} = &- \frac{(\beta-1)}{4\alpha}\varepsilon(\alpha-1)\left(1+\varepsilon^2\right)^{\alpha+\beta-1}\ {}_2F_1(\alpha,\beta,2,-\varepsilon^2) \\ &- \frac{(\beta-1)}{4\alpha}\varepsilon\left(1+\varepsilon^2\right)^{\alpha+\beta-1}\ {}_2F_1(\alpha,\beta,1,-\varepsilon^2) \end{aligned} \tag{4.2.42}$$

Once again, taking the limit as $\varepsilon\to 0$ and noting that

$$\begin{aligned} &\lim_{\varepsilon\to 0}\frac{(\beta-1)(\alpha-1)\varepsilon}{2\alpha} \\ =\ &\lim_{\varepsilon\to 0}\left(\frac{T-X}{8}-\frac{\varepsilon}{4}-\frac{T-X}{4(T+X)}\varepsilon+\frac{\varepsilon^2}{2(T+X)})\right) \\ =\ &\frac{T-X}{8} \end{aligned} \tag{4.2.43}$$

and

$$\lim_{\varepsilon\to 0}\left(\frac{\beta-1}{4\alpha}\right)\varepsilon=\lim_{\varepsilon\to 0}\varepsilon\left(\frac{(T-X)-2\varepsilon}{4(T+X)}\right)=0 \tag{4.2.44}$$

one uses 4.2.38 to obtain

$$\begin{aligned} K^c_{--}(a,b) &= -\frac{T-X}{8}\lim_{\alpha,\beta\to\infty}{}_2F_1(\alpha,\beta,2,-\frac{T^2-X^2}{4\alpha\beta}) \\ &= -\frac{\Gamma(2)}{\frac{1}{2}\sqrt{T^2-X^2}}\frac{(T-X)}{8}J_1(\sqrt{T^2-X^2}) \\ &= -\frac{(T-X)}{2\sqrt{T^2-X^2}}J_1(\sqrt{T^2-X^2}) \end{aligned} \tag{4.2.45}$$

Finally, an identical argument gives

$$K^c_{++}(a,b)=-\frac{(T+X)}{2\sqrt{T^2-X^2}}J_1(\sqrt{T^2-X^2}) \tag{4.2.46}$$

Equations 4.2.40,4.2.45 and 4.2.46 give the continuum propagator for the Dirac equation in one dimension [51].

Feynman failed in his efforts to generalise to higher dimensions his successful $2=(1+1)$ dimensional derivation of the Dirac equation. Tony Smith [71] generalised Feynman's argument and arrived at the hyperdiamond Feynman Checkerboard model based on the 4-dim hyperdiamond lattice. I use his work to hypothesise in the subsequent section, how quantum theory could arise within inverted theory networks governed by natural selection.

4.2.2 Inverted theory networks and quantum theory

The possible futures of inverted theory networks clearly lend themselves to the same interpretation as the sum over all possible paths methodology in the path integral approach to quantum theory. If one is to accept this, the following question needs to be answered: What is the probability amplitude associated with each of these possible futures. I will spend the remainder of this section detailing a hypothesis that I will call the $(TEA)^{-1}$ hypothesis in Chapter 5. Towards this end, I define a projection $\sigma : Form(\Phi_n)\to\mathbb{H}$ mapping propositional formulae to quaternion numbers. Any quaternion $H\in\mathbb{H}$ can be represented as

$$H=a+bI+cJ+dK \tag{4.2.47}$$

where $a, b, c, d \in \mathbb{R}$ are real numbers and I, J, K are the complex 2×2 matrices

$$I = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$$

$$J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$K = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}$$

Given $\phi \in Form(\Phi_n)$, define the following subsets of Φ_n:

$$\Pi^{\emptyset}(\phi) = \{p \in \Phi_n \mid \phi \text{ does not require } p\}$$

$$\Pi^{+}(\phi) = \{p \in \Phi_n \mid \phi \text{ requires only } p\}$$

$$\Pi^{-}(\phi) = \{p \in \Phi_n \mid \phi \text{ requires only } \neg p\}$$

$$\Pi^{+-}(\phi) = \{p \in \Phi_n \mid \phi \text{ requires } p \text{ and } \neg p\}$$

Further let

$$\pi^{\emptyset}(\phi) = \#\Pi^{\emptyset}(\phi)$$

$$\pi^{+}(\phi) = \#\Pi^{+}(\phi)$$

$$\pi^{-}(\phi) = \#\Pi^{-}(\phi)$$

$$\pi^{+-}(\phi) = \#\Pi^{+-}(\phi)$$

Then

$$\sigma(\phi) = \frac{1}{n}(\pi^{\emptyset}(\phi) + \pi^{+}(\phi)I + \pi^{-}(\phi)J + \pi^{+-}(\phi)K) \tag{4.2.48}$$

Now consider the inverted theory network $\mathcal{T}^{-1} = (\xi, R_f, T)$ over PropCal_n governed by the affirmed implication relation and the constrained uniform substitution transition function i.e. the inverted theory network governed by natural selection. Let $\mu(t)$ be a logicatom at time t, ζ_1 a possible future and $\mu^{\zeta_1}(t-1)$ the corresponding logicatom 1 time step away along possible future ζ. Then the probability amplitude for logicatom $\mu(t)$ to evolve along possible future ζ to logicatom $\mu^{\zeta_1}(t-1)$ is given by

$$K(\mu(t), \zeta_1) = \frac{\sigma(B[\mu^{\zeta_1}(t-1)])}{\sigma(B[\mu(t)])}$$

One iteratively continues this argument to calculate the probability amplitude for the next possible future ζ_2. Thus

$$K(\mu^{\zeta_1}(t-1), \zeta_2) = \frac{\sigma(B[\mu^{\zeta_2}(t-2)])}{\sigma(B[\mu^{\zeta_1}(t-1)])}$$

The product of these probability amplitudes defines the amplitude for a particular future path $\mathcal{F}_k = (\zeta_1, \zeta_2, \ldots, \zeta_k)$ i.e. the probability amplitude for a logicatom $\mu(t)$ at time t to evolve to logicatom $\mu(t-k)$ at time $t-k$ along the future path $\mathcal{F}_k = (\zeta_1, \zeta_2, \ldots, \zeta_k)$ is given by:

$$K(\mu(t), \mu(t-k), \mathcal{F}_k) = \frac{\sigma(B[\mu^{\zeta_k}(t-k)])}{\sigma(B[\mu(t-1)])}$$

Then the propagator for logicatom $\mu(t)$ to believe $B(\mu(t-k))$ at $t-k$ is given by

$$K(\mu(t), \mu(t-k)) = \sum_{\mathcal{F}_k} K(\mu(t), \mu(t-k), \mathcal{F}_k) \tag{4.2.49}$$

where the sum is over all possible future paths. Equation 4.2.49 will be used in Chapter 5 to detail the $(TEA)^{-1}$ hypothesis. This basically states that Equation 4.2.49 is actually the (3+1) dimensional Dirac equation.

4.3 Modelling knowledge

Inverted theory networks comprise logicatoms that are defined using description logic. Further, certain inverted theory networks are regulated by natural selection. In summary, I have a space comprising entities that are defined using the formal mathematical language for knowledge description and are regulated by natural selection i.e. Dawkins' memes. Thus if one accepts Dawkins' hypothesis that knowledge is modelled using memes, then inverted theory networks are a possible platform in which to model the dynamics and evolution of knowledge. Further, using the argument that theory networks can simulate cellular automata, and noting that the cellular automata model specified by [66] can be viewed as a possible future in an inverted theory network shows that emergent structures can arise that have the ability to self-replicate i.e. evangelical (self-propagating) beliefs can arise.

The above arguments are merely qualitative, and do not provide a way forward to research whether inverted theory networks can actually model the way say knowledge evolves in our brains. Towards this end, consider the inverted theory network over PropCal_n governed by the affirmed implication relation and the constrained uniform substitution transition function i.e. the inverted theory network governed by natural selection. In Section 2.4.2, I showed that the evolution of theory networks of this type have an interesting interpretation in terms of rule based reasoning. In particular, I proved that the updated belief of any logicatom could be viewed as the consequence (using rule based reasoning) of its previous belief and monotonic rules specified by all logicatoms it is related to. Now consider what is happening in the inverted theory networks of this category. A logicatom has a particular belief at time t. Every possible future at time $t-1$ represents a belief that the logicatom could

have, together with a set of monotonic rules (defined by the substitution map) that would make the belief at time t a consequent. In layman's terms, a possible future gives one a possible answer to the question: 'How did I get to believe what I believe?'. Now I will argue why I think this is the way we think!. Einstein conceptualised his theory of general relativity long before he learnt the mathematics of differential geometry, that provided him with a formal language in which to explain his theory to the rest of the world. I believe the process of thought involves having a belief/idea and considering all the possible monotonic rule sets and beliefs that would infer the current belief using rule-based reasoning. These possible sets of monotonic rules and corresponding beliefs are constrained by the current beliefs. This is exactly what the above class of inverted theory networks achieve. The reader can clearly question why I believe this? The only way I can motivate this argument is by analysing the way I think. Ten years ago, I barged into my then supervisor's office claiming that the universe is just one giant brain. He in response did what any reasonable person would do - he asked how I deduced this i.e. he asked for a current accepted belief (i.e. a physical theory) together with a set of monotonic rules that would lead one to affirm this statement. My thought process over the last ten years encompassed acquiring these rules and beliefs, resulting in the current thesis.

Even if one accepts the above argument, the question still remains - how does one conclude the one possible future that contains the correct belief and monotonic rules that affirm my current belief as a consequent. Since there are many possible futures, quantum theory tells us rather to assign a probability amplitude with each possible future. Summing over all possible futures provides one with a distribution that tells one the probability that we will be in one particular state (i.e. a particular future). Statistics then dictates that the 'answer' is actually a probability distribution of possible 'answers'. Our 'choice' in choosing the answer is then governed by these associated probabilities i.e. I am most likely to choose the possible future with the highest associated probability. I believe these probability amplitudes are specified by Equation 4.2.49 i.e. the mathematics used to show how quantum theory can arise is identical to the mathematics used to model thought.

Chapter 5

The $(TEA)^{-1}$ Hypothesis

I believe $(TEA)^{-1}$ Theory.
This is because $(TEA)^{-1} = A^{-1}E^{-1}T^{-1} = \forall\exists\ \perp$
That is, everything exists in some inverted theory network

I would really have wanted to say that I have proved $(\text{TEA})^{-1}$ Theory, but in reality, I have not even touched the surface - this, in my opinion is just a bit more detail regarding an "idea for an idea" [55]. I personally feel that I would need another 5 years of focused research to complete a comprehensive mathematical theory of inverted theory networks, in order to convince the sceptical scientist that this is a path worth researching. I will now detail what has (and has not) been achieved.

I have come some way in formalising natural selection. This formalisation resulted in interesting results: Firstly, the requirement of non-deterministic replication that is always taken for granted comes to the fore when formalising the theory. Secondly, there is a specific equation that holds in any system regulated by natural selection. Further, the power of using the equation of natural selection in biological arguments was evidenced in the argument as to why sex evolved. Thirdly, the mathematical arguments delivered is in effect a formal proof that we as humans are currently (to some degree) regulated by this principle.

I have succeeded in constructing a space, known as an inverted theory network, that is built up using propositional calculus and is regulated by the principle of natural selection i.e. the inverted theory network over PropCal_n governed by the affirmed implication relation and constrained uniform substitution function. This space can thus be seen as the formal analog to the space of Dawkins' memes. I have argued that this inverted theory network can be viewed as the pregeometry proposed by Wheeler, implying that his sought after regulating principle is in fact natural selection. This is evidenced in showing that this structure predicts the arrow of time (by construction) and the dimension of space to be three.

I have shown that projected inverted theory networks can model non-deterministic cellular automata. This is a consequence of the fact that projected theory networks can simulate cellular automata systems. Inverted theory networks thus inherit many characteristics of cellular automata, including self-replication, a requirement to modelling memes. Further, by construction, logicatoms in inverted theory networks over $ML(\Diamond, \Phi_n)$, adhere to the philosophy of Russell's logical atomism - they can be composed wholly of constituents with which they are acquainted i.e. logicatoms.

I have not shown what I ideally would have liked to do i.e. the $(TEA)^{-1}$ hypothesis specified below:

Proposition 5.0.1. *The $(TEA)^{-1}$ hypothesis:*
Consider the projected inverted theory network $\mathcal{T}^{-1} = (\xi, R_f, T, \sigma)$ over $PropCal_n$ where

$$R_f(\phi) = \{p \in \Phi_n \mid \vdash \phi \to p\}$$

is the affirmed implication relation generating function,

$$T(\xi)(p) = [\xi(p)]^{\xi|_{R_f(p)}}$$

is the constrained uniform substitution transition map and

$$\sigma(\phi) = \frac{1}{n}(\pi^{\emptyset}(\phi) + \pi^{+}(\phi)I + \pi^{-}(\phi)J + \pi^{+-}(\phi)K)$$

is the projection map defined in Section 4.2.2, mapping propositions to quaternion numbers. Then the propagator

$$K(\mu(t), \mu(t-k)) = \sum_{\mathcal{F}_k} K(\mu(t), \mu(t-k), \mathcal{F}_k) \tag{5.0.1}$$

generates the Green's function of the $(3+1)$ dimensional Dirac equation. Further, this propagator will model creative thought as specified in Section 4.3.

I believe that proving the above hypothesis (or alternatively, fine-tuning the $(TEA)^{-1}$ hypothesis so as to achieve the sought after objective in my synthesis) provides opportunities for good future research in this field. This is not to say that research in this field, from an applied mathematical perspective is not warranted. Section 2.4.1 shows how these structures provide powerful modelling platforms, and I believe further research is warranted in terms of the application of these structures to the applied sciences. From a purely mathematical perspective, the exact relationship between the regulating principle of natural selection and pregeometric concepts also needs further analysis.

I will end this journey by giving the reader one philosophical point to ponder: Assume I have a mathematical platform with which to model human thought. Now the environment together with some regulating process in my brain resulted in me having certain unalterable beliefs e.g. God exists. But I can argue that the same platform and regulating process will create unalterable laws of nature e.g. quantum electromagnetism. All I have done is argue that this regulating principle is natural selection.

Appendix A

Numerical Simulations

A substantial part of the research done involved the numerical simulation of the various models and structures that have been discussed in this thesis. These numerical simulations provided a powerful laboratory and steered me in the correct direction, allowing me to numerically confirm that various aspects of the thesis were (numerically) true before attempting to formally prove them. This appendix details the mathematics and algorithms involved in these numerical simulations.

Section A.1 details the numerical simulations of the equation of natural selection. (These simulations were actually used to derive the equation.) I use these simulations to show the reader how the various definitions make sense from a numerical perspective.

Section A.2 details the work done in numerically simulating theory networks. These simulations helped me identify which local relations would satisfy the requirements of natural selection, prior to attempting the formal proofs. Further, the simulations enabled me to confirm all the counting arguments proven in Chapter 3.

A.1 Simulating the equation of natural selection

In the derivation of the equation of natural selection, there are two definitions that the reader might query. The first was Equation 3.3.14

$$p_i(w)\left[1-p_m(w)\right]+p_m(w)=\mathcal{E}\left[\frac{\sum\limits_{\phi\in H}q(\phi,w)z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H}q(\phi,w)z(\phi)}\right]$$

that stated that the expectation value of the ratio of summation terms shown on the right hand side is equal to the probability that a selecton was mutated or that the selecton was inherited and no mutation occurred.

The second involved the definition of the probability of mutation. For the asexual case, this was specified in Equation 3.3.13 as

$$p_m(w)=\mathcal{E}\left[\frac{\sum\limits_{\phi\in H}\left[1-q(\phi,w)\right]z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H}\left[1-q(\phi,w)\right]z(\phi)}\right]$$

For the case of 'polysexual' reproduction (i.e. more than 1 parent), this was generalised to Equation 3.3.20, specified as

$$\begin{aligned}p_m(w)+p_i(w)(1 \quad &- \quad p_m(w))(1-\left[1-Q(H,w)-\rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}}\right]^{C-1})\\ &= \quad \mathcal{E}\left[\frac{\sum\limits_{\phi\in H}\left[1-q(\phi,w)\right]z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H}\left[1-q(\phi,w)\right]z(\phi)}\right]\end{aligned}$$

As specified, it has been assumed that $p_i(w)$ and $p_m(w)$ are the means of normally distributed samples. The numerical simulations detailed below should convince the reader as to the validity of these equations.

A.1.1 Algorithms implemented

I wrote a program using VBA (Visual Basic for Applications) in Excel. This particular platform was chosen to facilitate graphical representation of the results. The program numerically simulated the dynamics of one selecton in a population of replicators over one generation. Multiple simulations allowed me to come up with the statistical results shown graphically in the various charts.

The following parameters were defined in the code. The notation used below is identical to that used in Section 3.3.2.

Parameter	Type	Description
$n_{\text{Predecessors}}$	Integer	This parameter specifies the number of replicators in the predecessor set H.
$n_{\text{Successors}}$	Integer	This parameter specifies the number of replicators in the successor set H'.
$Q(H)$	$[0, 1]$	This parameter specifies the frequency of the replicators in H that have the selection w.
$C_{\max}$	Integer	This parameter specifies the maximum number of predecessors (parents) of each successor in H'.
$\text{Parent}_{\text{Constant}}$	Boolean	This parameter determines whether the successors all have the same number of predecessors (True) or can have between 1 and C parents (False).
p_m	$[0, 1]$	This parameter specifies the probability that a successor has the selection w given that none of its parents have w.
p_i	$[0, 1]$	This parameter specifies the probability that a successor has the selection w, given that at least one of its parents has w.
$p_{\rho(q,z)}$	$[-1, 1]$	This parameter is a measure of how correlated with fitness we want the selecton w to be. The exact use is clarified in the algorithm below.

Table A.1: Parameters used in the numerical simulation of the equation of natural selection

The algorithm implemented to simulate the selecton frequency from one generation to the next is described below.

(1) Create all the data structures required.

(1.1) Instantiate a boolean array of size $n_{\text{Predecessors}}$ that represents the names of the predecessors (defined as the index of the element) and whether they have the selecton w (Boolean Value = True) or not (Boolean Value = False).

(1.2) Similarly, instantiate a boolean array of size $n_{\text{Successors}}$ for the successors.

(1.3) Instantiate an integer array of size $(n_{\text{Successors}} \times C)$. This data structure will demarcate the names of every parent for any given successor.

(2) Distribute the selecton w among the predecessor population. For each predecessor in the array, generate a random number $r \in [0, 1]$. If $r \leq Q(H)$ then give the predecessor the selecton i.e assign value of True in the respective boolean array. Otherwise assign a value of False.

(3) For each successor, select its parents from the predecessor set.

(3.1) If $\text{Parent}_{\text{Constant}} = \text{True}$, then repeat the Step 3.2 C times. Otherwise generate a random number $r \in [0, 1]$. Assign $C' = \text{Min}(1,\text{Round}(r \times C))$ parents to the successor, and repeat Step 3.2 C' times. Note that in this case, I have set $p_C(H) = \frac{1}{C_{\max}}$ in Equation 3.3.21 i.e. Every replicator is equally likely to have anywhere between 1 and $C_{\max}$ predecessors.

(3.2) Generate a random number $r \in [0, 1]$.

(3.2.1) For the case $0 \leq p_{\rho(q,z)} \leq 1$: If $r \leq p_{\rho(q,z)}$ then randomly select a unique parent (i.e. it must not exist in the successors parent set already) that has the selecton w. Otherwise randomly select a unique parent from the whole set H.

(3.2.1) For the case $-1 \leq p_{\rho(q,z)} < 0$: If $r \leq |p_{\rho(q,z)}|$ then randomly select a unique parent that does not have the selecton w. Otherwise randomly select a unique parent from the whole set H. This process verifies how we use $p_{\rho(q,z)}$ to simulate various values of $\rho(q, z)$

(4) Distribute the selecton w among the successor population. For every successor in H' repeat the following:

(4.1) Generate a random number $r \in [0, 1]$. If $r \leq p_m$ then give the successor the selecton and goto Step 4 otherwise implement Step 4.2

(4.2) Iterate through all the predecessors of the successor and confirm whether at least one predecessor has the selecton. If this is not the case, then don't assign the selecton to the successor and goto Step 4. Otherwise generate a random number $r \in [0, 1]$. If $r \leq p_i$ then give the successor the selecton, otherwise not.

The process described above generates all the data required to calculate the various terms in Equations 3.3.14 and 3.3.20.

A.1.2 Simulation Results

I first show the results pertaining to the equation

$$p_i(w)\left[1-p_m(w)\right]+p_m(w)=\mathcal{E}\left[\frac{\sum\limits_{\phi\in H} q(\phi,w)z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H} q(\phi,w)z(\phi)}\right]$$

I calculate the value of $\frac{\sum\limits_{\phi\in H} q(\phi,w)z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H} q(\phi,w)z(\phi)}$ in a single simulation, and compare it with the expected value of $p_i(w)\left[1-p_m(w)\right]+p_m(w)$. I plot the relationships between the various variables for fixed $n_{\text{Predecessors}}=1000$, $n_{\text{Successors}}=1000$ and $Q(H,w)=0.6$. The graph below shows the relationship for 3 values of p_i ($p_i\in\{\frac{1}{3},\frac{1}{2},\frac{4}{5}\}$) and all values of p_m with $C=2$.

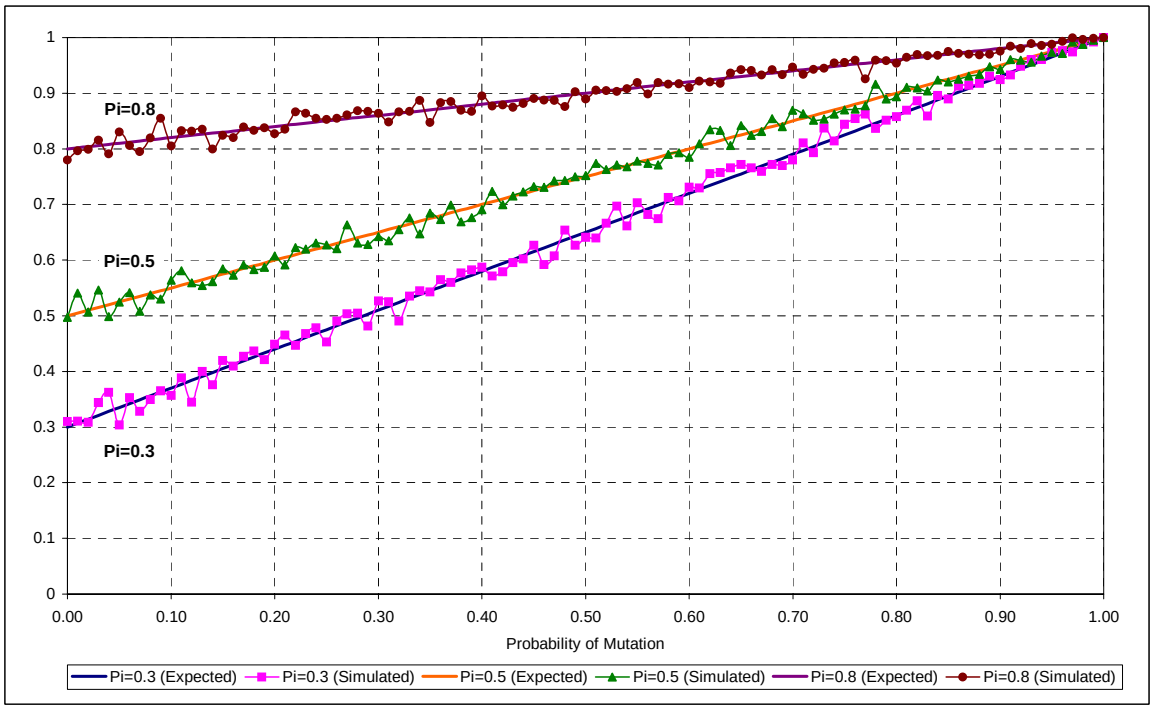

Figure A.1: The simulated value of $\frac{\sum\limits_{\phi\in H} q(\phi,w)z(\phi)q'(\phi,w)}{\sum\limits_{\phi\in H} q(\phi,w)z(\phi)}$ compared to the expected value of $p_i(w)\left[1-p_m(w)\right]+p_m(w)$ for fixed p_i and C.

The projection below shows the relationship for 3 values of p_m ($p_m \in \{\frac{1}{3}, \frac{1}{2}, \frac{4}{5}\}$) and all values of p_i, with $C = 2$.

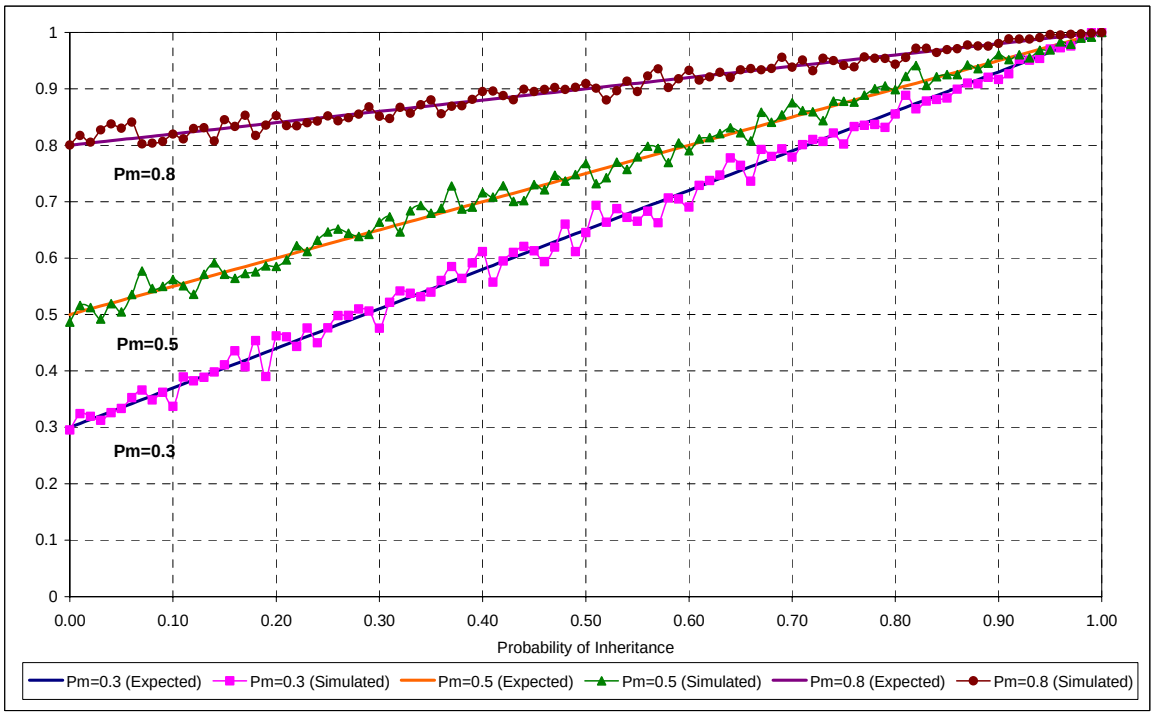

Figure A.2: The simulated value of $\frac{\sum_{\phi \in H} q(\phi,w) z(\phi) q'(\phi,w)}{\sum_{\phi \in H} q(\phi,w) z(\phi)}$ compared to the expected value of $p_i(w)\left[1 - p_m(w)\right] + p_m(w)$ for fixed p_m and C.

Figures A.1 and A.2 show that the theoretical equations fit the numerical simulations very well, evidencing a correct definition of the expectation value in Equation 3.3.14.

I now show the results pertaining to the equation

$$p_m(w) + p_i(w)(1 \quad - \quad p_m(w))(1 - \left[1 - Q(H,w) - \rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}}\right]^{C-1}) \\ = \quad \mathcal{E}\left[\frac{\sum_{\phi \in H}\left[1 - q(\phi,w)\right] z(\phi) q'(\phi,w)}{\sum_{\phi \in H}\left[1 - q(\phi,w)\right] z(\phi)}\right] \tag{A.1.3}$$

Once again, the simulated values are plotted against the expected values. As before, the parameters take on the following values: $n_{\text{Predecessors}} = 1000$, $n_{\text{Successors}} = 1000$, $Q(H,w) = 0.6$.
Figure A.3 shows the results plotted over p_m for 3 values of p_i ($p_i \in \{\frac{1}{3}, \frac{1}{2}, \frac{4}{5}\}$) and $C = 2$.

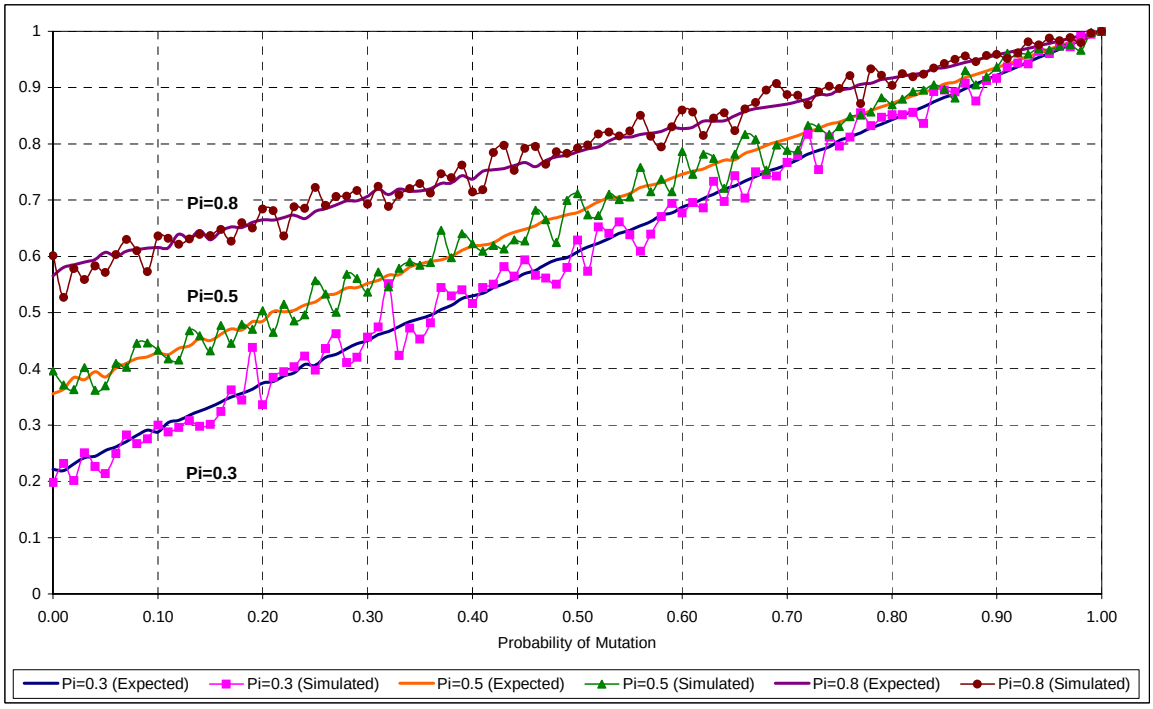

Figure A.3: The simulated value of $\frac{\sum_{\phi \in H}[1-q(\phi,w)]z(\phi)q'(\phi,w)}{\sum_{\phi \in H}[1-q(\phi,w)]z(\phi)}$ compared to the expected value of $p_m(w) + p_i(w)(1 - p_m(w))(1 - \left[1 - Q(H,w) - \rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}}\right]^{C-1})$ for fixed p_i and C.

Figure A.4 shows the results plotted over p_i for 3 values of p_m ($p_m \in \{\frac{1}{3}, \frac{1}{2}, \frac{4}{5}\}$) and $C = 2$

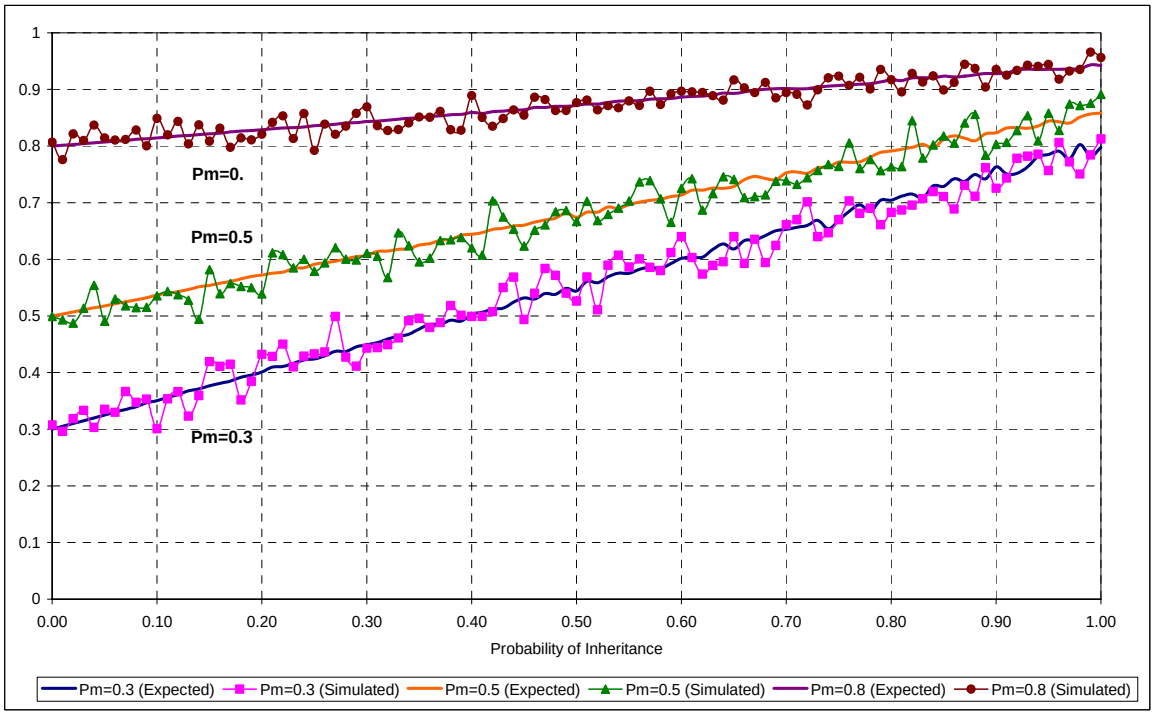

Figure A.4: The simulated value of $\frac{\sum_{\phi \in H}[1-q(\phi,w)]z(\phi)q'(\phi,w)}{\sum_{\phi \in H}[1-q(\phi,w)]z(\phi)}$ compared to the expected value of $p_m(w)+p_i(w)(1-p_m(w))(1-\left[1-Q(H,w)-\rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}}\right]^{C-1})$ for fixed p_m and C

Figure A.5 shows the results plotted over p_m for 3 values of C ($C \in \{1, 2, 5\}$) and $p_i = 0.3$

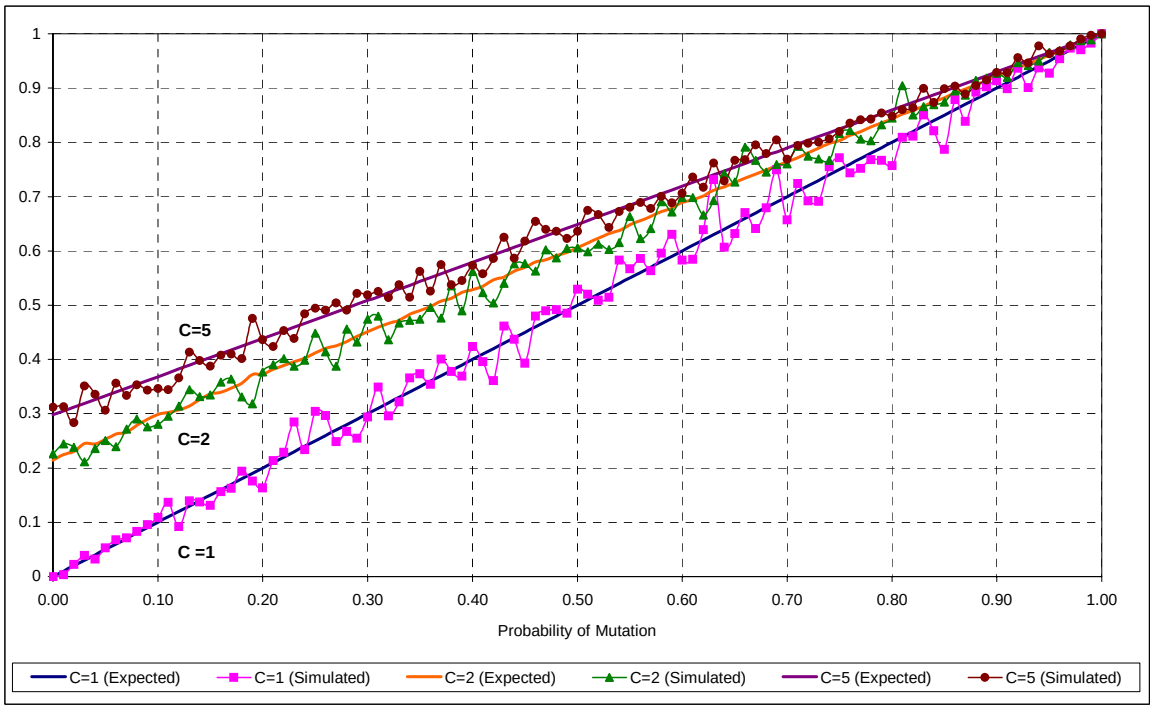

Figure A.5: The simulated value of $\frac{\sum_{\phi \in H}[1-q(\phi,w)]z(\phi)q'(\phi,w)}{\sum_{\phi \in H}[1-q(\phi,w)]z(\phi)}$ compared to the expected value of $p_m(w)+p_i(w)(1-p_m(w))(1-\left[1-Q(H,w)-\rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}}\right]^{C-1})$ for fixed p_i and C

Finally Figure A.6 shows the results plotted over p_i for 3 values of C ($C \in \{1, 2, 5\}$) and $p_m = 0.3$

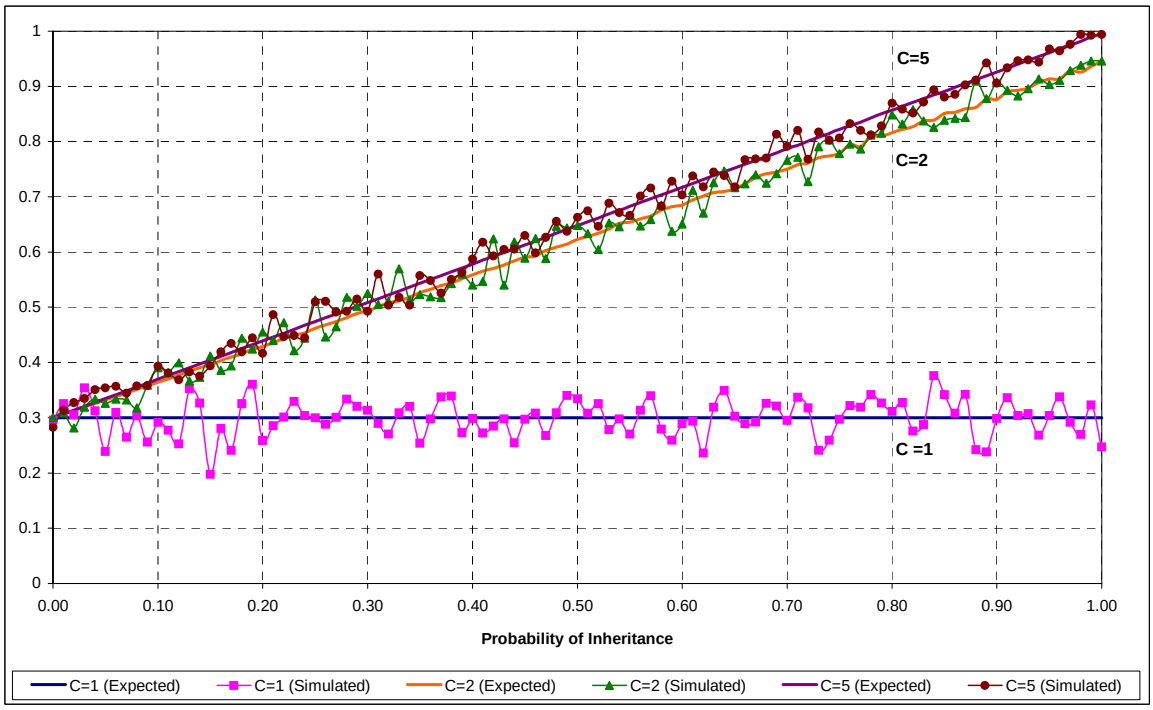

Figure A.6: The simulated value of $\frac{\sum_{\phi \in H} [1-q(\phi,w)]z(\phi)q'(\phi,w)}{\sum_{\phi \in H} [1-q(\phi,w)]z(\phi)}$ compared to the expected value of $p_m(w) + p_i(w)(1-p_m(w))(1-\left[1-Q(H,w)-\rho(q,z)\sigma(q)\frac{\sigma(z)}{\bar{z}}\right]^{C-1})$ for fixed p_m and C

Figures A.3, A.4, A.5, A.6 evidence the good fit between the theoretical and numerical simulations, validating the definition implied in Equation 3.3.20.

A.2 Simulating theory networks

In order to numerically simulate theory networks, I would be required to write algorithms to implement the classes of local relations defined in Definitions 3.5.1 and 3.5.4. Further, I would need to be able to numerically implement the constrained uniform substitution transition function. Section A.2.1 shows the theorems used to implement these algorithms using the many world formalism of propositional calculus.

A.2.1 Mathematics used for numerical simulations

In the many world formalism of propositional calculus, the three logical connectives $\vee$ (or), $\wedge$ (and) and $\neg$ (not) are semantically equivalent to the set operations $\cup$ (union), $\cap$ (intersection) and (complement) respectively.

$$\begin{aligned} [\psi \wedge \phi] &= [\psi] \cap [\phi] \\ [\psi \vee \phi] &= [\psi] \cup [\phi] \\ [\neg\psi] &= \overline{[\psi]} \end{aligned} \tag{A.2.1}$$

I will first prove the theorems that are used to simulate the local relations of the various theory networks under consideration. In Definition 2.4.5, I defined the concept of a proposition *requiring* a variable iff the proposition cannot be written without explicitly referring to the variable. One can calculate which variables are required to represent any proposition in $Form(\Phi_n)$ using the following lemma:

Lemma A.2.1. *p is not required in ϕ iff $\vdash \phi \leftrightarrow \phi(\neg p/p)$.*

Proof. The necessary case is trivial since if p is not required in ϕ, I can express ϕ without explicitly referring to p, resulting in the uniform substitution $\phi\,(\neg p/p)$ having no effect on ϕ.
For sufficiency, assume p is required in ϕ. Then $\exists \phi_-, \phi_+$ not requiring p with

$$\vdash \phi \leftrightarrow [(\phi_- \wedge \neg p) \vee (\phi_+ \wedge p)] \tag{A.2.2}$$

and

$$\nvdash \phi_- \leftrightarrow \phi_+ \tag{A.2.3}$$

Formula A.2.2 is derived by writing ϕ in full disjunctive normal form and grouping the disjuncts into those that imply p and those that imply $\neg p$. Now Formula A.2.2 allows one to deduce

$$\phi\,(\neg p/p) := [(\phi_+ \wedge \neg p) \vee (\phi_- \wedge p)] \tag{A.2.4}$$

To prove the result, assume $\vdash \phi \leftrightarrow \phi\,(\neg p/p)$. Then

$$\vdash (\phi \wedge p) \leftrightarrow (\phi\,(\neg p/p) \wedge p) \tag{A.2.5}$$

Using A.2.2 and A.2.4 results in

$$\vdash (\phi_- \wedge p) \leftrightarrow (\phi_+ \wedge p) \tag{A.2.6}$$

Since p is a variable and ϕ_-, ϕ_+ do not require p, one can conclude $\vdash \phi_- \leftrightarrow \phi_+$ contradicting Formula A.2.3. □

The following lemma determines whether only p or $\neg p$ is required in a proposition:

Lemma A.2.2. *Given a proposition ϕ and a variable p, define*

$$\gamma_1 := (p \wedge \phi) \vee (\neg p \wedge \phi(\neg p/p)) \tag{A.2.7}$$
$$\gamma_2 := (\neg p \wedge \phi) \vee (p \wedge \phi(\neg p/p)) \tag{A.2.8}$$

Then the following hold:

(a): $\vdash (\gamma_1 \leftrightarrow \gamma_2)$ iff p is not required in ϕ

(b): $\vdash (\gamma_1 \rightarrow \gamma_2)$ and (a) does not hold iff only $\neg p$ is required in ϕ.

(c): $\vdash (\gamma_2 \rightarrow \gamma_1)$ and (a) does not hold iff only p is required in ϕ

*(d): (a), (b) and (c) do not hold iff p **and** $\neg p$ are required in ϕ*

Proof. Now ϕ can be written in a disjunctive normal form, such that

$$\vdash \phi \leftrightarrow [(\psi_1 \wedge p) \vee (\psi_2 \wedge \neg p) \vee \psi_3] \tag{A.2.9}$$

with ψ_1, ψ_2, ψ_3 not requiring the variable p. (To achieve this, write ϕ in full disjunctive normal form and group the conjuncts requiring p, $\neg p$ and not requiring p or $\neg p$ respectively.) Now by definition of γ_1, one has $\gamma_1 := ((\psi_1 \vee \psi_3) \wedge p) \vee ((\psi_1 \vee \psi_3) \wedge \neg p)$ resulting in

$$\vdash \gamma_1 \quad \leftrightarrow \quad (\psi_1 \vee \psi_3) \tag{A.2.10}$$

Similarly

$$\vdash \gamma_2 \quad \leftrightarrow \quad (\ \psi_2 \vee \psi_3) \tag{A.2.11}$$

Expression A.2.9 can be written in the form

$$\vdash \phi \leftrightarrow [((\psi_1 \vee \psi_3) \wedge p) \vee ((\psi_2 \vee \psi_3) \wedge \neg p)] \tag{A.2.12}$$

resulting in

$$\vdash \phi \leftrightarrow (\gamma_1 \wedge p) \vee (\gamma_2 \wedge \neg p) \tag{A.2.13}$$

(a) Assume $\vdash \gamma_1 \leftrightarrow \gamma_2$. Then using A.2.13 and A.2.10 one has

$$\vdash \phi \leftrightarrow \gamma_1$$
$$\vdash \phi \leftrightarrow \psi_1 \vee \psi_3 \tag{A.2.14}$$

with ψ_1 and ψ_3 not requiring p. Thus ϕ does not require p.
Conversely assume ϕ does not require p. Then $\exists\psi_3$ such that $\vdash \psi_1 \leftrightarrow \bot$ and $\vdash \psi_2 \leftrightarrow \bot$ in A.2.9. Using A.2.10 and A.2.10 one has

$$\vdash \gamma_1 \leftrightarrow \psi_3$$
$$\vdash \psi_3 \leftrightarrow \gamma_2$$

resulting in $\gamma_1 \leftrightarrow \gamma_2$ as required.
(b) Assume $\vdash \gamma_1 \rightarrow \gamma_2$ and $\nvdash \gamma_2 \rightarrow \gamma_1$.
Thus

$$\vdash \gamma_2 \leftrightarrow (\gamma_1 \vee \eta) \tag{A.2.15}$$

for some η such that

- $\nvdash \eta \leftrightarrow \top$: If η was a tautology, then γ_2 and γ_1 would both be tautologies, contradicting the assumption that they aren't equivalent.
- $\nvdash \eta \leftrightarrow \bot$: If η was a contradiction, one would have $\vdash \gamma_1 \leftrightarrow \gamma_2$, once again contradicting the assumption.
- η does not require p since by A.2.11 and A.2.10, γ_1 and γ_2 do not require p

Substituting the equivalent formula $(\gamma_1 \vee \eta)$ for γ_2 (as specified by A.2.17) into A.2.13 results in

$$\vdash \phi \leftrightarrow (\gamma_1 \vee (\neg p \wedge \eta)) \tag{A.2.16}$$

concluding that ϕ can be expressed using only the affirmed variable p.
Conversely assume ϕ only requires p. Then $\exists\psi_1, \psi_2, \psi_3$ with
$\nvdash \psi_1 \leftrightarrow \bot$ and $\vdash \psi_2 \leftrightarrow \bot$ in A.2.9. Using A.2.10 and A.2.11 one has

$$\vdash \gamma_1 \leftrightarrow \psi_3$$
$$\vdash \psi_3 \rightarrow (\psi_2 \vee \psi_3)$$
$$\vdash (\psi_2 \vee \psi_3) \leftrightarrow \gamma_2$$

proving $\vdash \gamma_1 \rightarrow \gamma_2$. Now if $\gamma_2 \rightarrow \gamma_1$ then $\gamma_1 \leftrightarrow \gamma_2$, which by (a) implies that ϕ does not require p, contradicting the assumption. Thus $\nvdash (\gamma_2 \rightarrow \gamma_1)$ as required.
(c) The proof of (c) is completely symmetrical to that of (b); Assume $\vdash \gamma_2 \rightarrow \gamma_1$ and $\nvdash \gamma_1 \rightarrow \gamma_2$.
Thus

$$\vdash \gamma_1 \leftrightarrow (\gamma_2 \vee \eta) \tag{A.2.17}$$

for some η such that

- $\nvdash \eta \leftrightarrow \top$: If η was a tautology, then γ_2 and γ_1 would both be tautologies, contradicting the assumption that they aren't equivalent.
- $\nvdash \eta \leftrightarrow \bot$: If η was a contradiction, one would have $\vdash \gamma_2 \leftrightarrow \gamma_1$, once again contradicting the assumption.
- η does not require p since by A.2.11 and A.2.10, γ_1 and γ_2 do not require p

Substituting the equivalent formula $(\gamma_2 \vee \eta)$ for γ_1 (as specified by A.2.17) into A.2.13 results in

$$\vdash \phi \leftrightarrow (\gamma_2 \vee (p \wedge \eta)) \tag{A.2.18}$$

concluding that ϕ can be expressed using only the negated variable $\neg p$.
Conversely assume ϕ only requires $\neg p$. Then $\exists \psi_1, \psi_2, \psi_3$ with
$\nvdash \psi_2 \leftrightarrow \bot$ and $\vdash \psi_1 \leftrightarrow \bot$ in A.2.9. Using A.2.10 and A.2.11 one has

$$\begin{aligned} \vdash \gamma_2 &\leftrightarrow \psi_3 \\ \vdash \psi_3 &\rightarrow (\psi_1 \vee \psi_3) \\ \vdash (\psi_1 \vee \psi_3) &\leftrightarrow \gamma_1 \end{aligned}$$

proving $\vdash \gamma_2 \rightarrow \gamma_1$. Now if $\gamma_1 \rightarrow \gamma_2$ then $\gamma_1 \leftrightarrow \gamma_2$, which by (a) implies that ϕ does not require p, contradicting the assumption. Thus $\nvdash (\gamma_1 \rightarrow \gamma_2)$ as required. (d) Since all the other options have been eliminated, one can conclude (d). □

Finally, the following lemma shows whether a proposition implies an affirmed (respectively negated) variable.

Lemma A.2.3. *Given a proposition ϕ and a variable p, one has $\vdash \phi \rightarrow p$ iff $\vdash (\neg\phi \vee p) \leftrightarrow \top$*

Proof. Assume $\vdash \phi \rightarrow p$. Then ϕ can be expressed in the form $\phi := \alpha \wedge p$ for some proposition α not requiring p. The result follows. □

The above three lemmas would allow me to semantically (and therefore numerically) calculate whether a relation holds. However they do not suffice by themselves since I have not specified how uniform substitution (required in Lemmas A.2.1 and A.2.2) would be implemented using the semantic representation of propositional calculus. This is also required in order to simulate the constrained uniform substitution transition function. Towards this end, I need to derive the following algorithms:

- **Algorithm I**: Consider any propositional variable $p_i \in \Phi_n$ in PropCal_n. Determine $[p_i]$, the set of worlds representing this atomic statement.
- **Algorithm II**: If ψ is a proposition, uniformly substitute $\xi : \Phi_n \rightarrow \mathit{Form}(\Phi_n)$ to evaluate the proposition η. Denote $\xi(p_i) = \phi_i$. Then

$$\eta = \psi\left(p_1/\phi_1, p_2/\phi_2, \ldots, p_n/\phi_n\right)$$

In order to proceed, I define an indexing methodology for any world in PropCal_n. One can represent a world $w : \mathit{Form}(\Phi_n) \rightarrow \{0,1\}$ as a binary series $w \equiv (b_1, b_2, \ldots, b_n)$ of size n by defining $w(p_k) = b_k \in \{0,1\}$. On the other hand, I uniquely index all worlds in PropCal_n from 0 to $2^n - 1$ by implementing the following rule:

$$w \equiv (b_1, b_2, \ldots, b_n) \Leftrightarrow w = w_i \text{ where } i = \sum_{k=0}^{n-1} b_{k+1} 2^k. \tag{A.2.19}$$

I am basically using the decimal representation of a binary number to calculate the index. Since every integer between $[0, 2^n - 1]$ has a unique binary representation (using n digits), one has a method of moving between the 2 representations. I will refer to this as the *binary indexing methodology.*

In order to create '**Algorithm I**', I generate the world representation for every atomic statement $p_k \in \Phi_n$ by using the following lemma:

Lemma A.2.4. *Using the binary indexing methodology for worlds in PropCal$_n$, one has* $[p_k] = \{w_i \mid i \in [(2j-1)2^{k-1}, j2^k - 1], j \in [1, 2^{n-k}]\}$ *for* $1 \leq k \leq n$.

Proof. Let $w_i \in [p_k] = \{b_1, b_2, \ldots, b_n\}$. By definition, we require $b_k = 1$. Consider the binary representation $B_{j-1} = (B_{j-1}(1), B_{j-1}(2), \ldots, B_{j-1}(n-k)$ of the integer $j - 1 \in [0, 2^{n-k} - 1]$. The $B_{j-1}(l)$ are defined uniquely by the fact that

$$j - 1 = \sum_{m=1}^{n-k} B_{j-1}(m) 2^{m-1} \tag{A.2.20}$$

Let $b_{k+l} = B_{j-1}(l)$. As one iterates over all possible values of $j \in [1, 2^{n-k}]$, all the possible binary sequences for $\{b_{k+1}, b_{k+2}, \ldots, b_n\}$ are obtained. Thus we only have to consider all the permutations of the binary series $\{b_1, b_2, \ldots, b_{k-1}\}$. For a given j, the smallest possible decimal value of the binary series is where $\forall l < k$, $b_l = 0$. This has the decimal representation of

$$\begin{aligned} i &= \sum_{m=1}^{n} b_m 2^{m-1} = 2^{k-1} + \sum_{m=k+1}^{n} b_m 2^{m-1} \\ &= 2^{k-1} + \sum_{m=1}^{n-k} b_{m+k} 2^{m+k-1} = 2^{k-1} + 2^k \sum_{m=1}^{n-k} B_{j-1}(m) 2^{m-1} \\ &= 2^{k-1} + 2^k (j-1) = 2^{k-1}(2j-1) \end{aligned} \tag{A.2.21}$$

Similarly, the largest possible decimal value of the binary series is where $\forall l < k$, $b_l = 1$. This has the decimal representation of

$$\begin{aligned} i &= \sum_{m=1}^{n} b_m 2^{m-1} = (2^k - 1) + \sum_{m=k+1}^{n} b_m 2^{m-1} \\ &= (2^k - 1) + \sum_{m=1}^{n-k} b_{m+k} 2^{m+k-1} = (2^k - 1) + 2^k \sum_{m=1}^{n-k} B_{j-1}(m) 2^{m-1} \\ &= (2^k - 1) + 2^k (j-1) = j2^k - 1 \end{aligned} \tag{A.2.22}$$

□

Using A.2.1 together with A.2.4 will allow one to generate the world set for any proposition ψ.

Algorithm II requires encapsulating 'uniform substitution' using the many world formalism. I build up to the solution by initially considering simple examples. In particular, consider the following simple case of substitution where

- $\eta = \psi\left(\phi/p_i\right)$
- the propositions ψ and ϕ are semantically represented by single worlds i.e. $[\psi] = \{w\}$ and $[\phi] = \{w'\}$ for some $w, w' \in \mathbf{V}$
- If $[\psi] = \{w\}$, then $w(p_i) = 1$ i.e. only the affirmed variable p is required in ϕ

The following examples satisfy these criteria in PropCal_3.

Example A.2.1. Set

$$\begin{aligned} \psi &= p_1 \wedge p_2 \wedge \neg p_3 \\ \phi_1 &= \neg p_1 \wedge \neg p_2 \wedge p_3 \\ \phi_2 &= p_1 \wedge \neg p_2 \wedge \neg p_3 \end{aligned}$$

and evaluate

$$\begin{aligned} \eta_1 &= \psi\left(\phi_1/p_2\right) \\ \eta_2 &= \psi\left(\phi_2/p_2\right) \end{aligned}$$

Using the binary indexing methodology, I define the following worlds in PropCal_3.

p_1	p_2	p_3	**Valuation**
0	0	0	w_0
1	0	0	w_1
0	1	0	w_2
1	1	0	w_3
0	0	1	w_4
1	0	1	w_5
0	1	1	w_6
1	1	1	w_7

I thus have $[\psi] = \{w_3\}$, $[\phi_1] = \{w_4\}$ and $[\phi_2] = \{w_1\}$. Now

$$\begin{aligned} \eta_1 &:= p_1 \wedge (\neg p_1 \wedge \neg p_2 \wedge p_3) \wedge \neg p_3 = \bot \\ \eta_2 &:= p_1 \wedge (p_1 \wedge \neg p_2 \wedge \neg p_3) \wedge \neg p_3 = \phi_2 \end{aligned}$$

Notice that when substituting, the 'value' of p_2 in ψ i.e. $w_3(p_2)$, takes on the 'value' of p_2 in the proposition being substituted i.e. $w_4(p_2)$ and $w_1(p_2)$ respectively. If the other variable values concur, then the result is always the proposition being substituted, else one obtains a contradiction. In other words for $[\psi] = \{w\}$, $[\phi] = \{w'\}$ and $w(p_k) = 1$, I have

$$[\psi\left(\phi/p_k\right)] := \begin{cases} [\phi] = \{w'\} & \text{if } w'(p_i) = w(p_i)\ \forall i \neq k \\ \emptyset & \text{otherwise} \end{cases} \tag{A.2.23}$$

Define the function $f : W \times \Phi_n \times \mathbf{2} \to W$ by

$$\begin{aligned} f(w, p_k, b) &= w' \text{ where} \\ w'(p_j) &= \begin{cases} w(p_j) & \text{if } j \neq k \\ b & \text{if } j = k \end{cases} \end{aligned} \tag{A.2.24}$$

Define $g : W \times \Phi_n \times W \to W$ by

$$g(w, p_k, w') = f\left(w, p_k, w'(p_k)\right) \tag{A.2.25}$$

And finally, define $h_0 : W \times \Phi_n \times W \to \mathcal{P}(W)$ by

$$h_0(w, p_k, w') = \begin{cases} w' & \text{if } g(w, p_k, w') = w' \\ \emptyset & \text{otherwise} \end{cases} \tag{A.2.26}$$

In summary, for the case where the propositions ϕ, ψ are semantically represented by a single world, with substitution occurring in only one variable $p \in \Phi_n$ with the property that ϕ only requires the affirmed variable p A.2.23 is equivalent to

$$[\psi\left(\phi/p_k\right)] = h_0([\psi], p_k, [\phi]) \tag{A.2.27}$$

Extending this argument to the case where the proposition ϕ being substituted has a semantic representation $[\phi] = \{w_1, w_2, \ldots, w_l\}$ of more than one world, one observes that it is possible to express ϕ in its full disjunctive normal form $\phi := \phi_1 \vee \phi_2 \vee \ldots \vee \phi_l$ where $\forall i \in [1, l]$, $[\phi_i] = \{w_j\}$ for some $w_j \in W$.
The substitution $\psi\left(\phi/p_k\right)$ is then just the disjunction of the propositions $\bigvee_{i=1}^{l} \psi\left(\phi_i/p_k\right)$.
Define $h_1 : W \times \Phi_n \times \mathcal{P}(W) \to \mathcal{P}(W)$ by

$$h_1(w, p_k, U') = \bigcup_{w' \in U'} h_0(w, p_k, w') \tag{A.2.28}$$

giving one

$$[\psi\left(\phi/p_k\right)] = h_1(w, p_k, [\phi]) \tag{A.2.29}$$

where the restrictions $[\psi] = \{w\}$ and $w(p_k) = 1$ still apply.

One now has the tools to handle the case when $w(p_k) = 0$ for $[\psi] = \{w\}$. Observe that substituting ϕ for a variable p_k in ψ with $w(p_k) = 0$ is equivalent to substituting $\neg\phi$ for the variable p_k in the proposition ψ' where $\psi' = \psi(\neg p_k/p_k)$. Towards this end, define $c : W \times \Phi_n \times \mathcal{P}(W) \to \mathcal{P}(W)$ by

$$c(w, p_k, U) = \quad = \quad \begin{cases} U & \text{if } w(p_k) = 1 \\ \overline{U} & \text{if } w(p_k) = 0 \end{cases} \tag{A.2.30}$$

and $h_2 : W \times \Phi_n \times \mathcal{P}(W) \to \mathcal{P}(W)$ by

$$h_2(w, p_k, U') = h_1(w, p_k, c(w, p_k, U')) \tag{A.2.31}$$

Now one has

$$[\psi(\phi/p_k)] = h_2(w, p_k, [\phi]) \tag{A.2.32}$$

where the only restriction is that $[\psi] = \{w\}$ i.e. a singleton.
The final case where $[\psi] = \{w_1, w_2, \ldots, w_p\}$ is argued in exactly the same way as when one extends ϕ to multiple worlds.
Defining $h_3 : \mathcal{P}(W) \times \Phi_n \times \mathcal{P}(W) \to \mathcal{P}(W)$ by

$$h_3(U, p_k, U') = \bigcup_{w \in U} h_2(w, p_k, U') \tag{A.2.33}$$

one has the result with no restrictions.

$$[\psi(\phi/p_k)] = h_3([\psi], p_k, [\phi]) \tag{A.2.34}$$

In order to handle the case of uniform substitution over multiple variables,

$$\eta \quad = \quad \psi\left(\phi_1/p_{k_1}; \phi_2/p_{k_2} \ldots \phi_j/p_{k_j}\right)$$

I return to the simple case once again:

- All propositions ψ and ϕ_i are semantically represented by single worlds
- If $[\psi] = \{w\}$, then $\forall i \; w(p_{k_i}) = 1$

Now unless $\phi_i = \phi_j \; \forall i, k \in [1, j]$, the substitution will result in the contradiction $\perp$. In the case where all ϕ_i are equal, an extension of the above argument gives one for

$$[\psi] = \{w\}$$
$$[\phi] = \{w'\}$$
$$\forall i \in [1, j] \, w'(p_{k_i}) = 1$$

the expression

$$\left[\psi\left(\phi/p_{k_1}\ldots\phi/p_{k_j}\right)\right] = \begin{cases} [\phi] = \{w'\} & \text{if } w'(p_i) = w(p_i)\ \forall i \notin [k_1, k_2, \ldots, k_j] \\ \emptyset & \text{otherwise} \end{cases} \tag{A.2.35}$$

Thus one proceeds to generalise all the functions defined for the single substitution. I always assume that $p_{k_i} \neq p_{k_l}$ for $i \neq l$ (as implied by uniform substitution) in the definitions below.
Define the function $f' : W \times \mathcal{P}(\Phi_n \times \mathbf{2}) \to W$ by

$$\begin{aligned} &f'[w, \{(p_{k_1}, b_1), (p_{k_2}, b_2), \ldots, (p_{k_j}, b_j)\}] = w' \text{ where} \\ &w'(p_j) = \begin{cases} w(p_j) & \text{if } j \notin [k_1, k_2, \ldots, k_j] \\ b_i & \text{if } j = k_i \text{ for some } i \in [1, 2, \ldots, j] \end{cases} \end{aligned} \tag{A.2.36}$$

The rest of the functions get similarly extended.
In particular $g' : W \times \mathcal{P}(\Phi_n \times W) \to W$ is defined by

$$g'[w, \{(p_{k_1}, w_{k_1}), \ldots, (p_{k_j}, w_{k_j})\}] = f'\left[w, \{(p_{k_1}, w(p_{k_1})), \ldots, (p_{k_j}, w(p_{k_j}))\}\right] \tag{A.2.37}$$

$h'_0 : W \times \mathcal{P}(\Phi_n \times W) \to \mathcal{P}(W)$ by

$$h'_0[w, \{(p_{k_1}, w_{k_1}), \ldots, (p_{k_j}, w_{k_j})\}] = \begin{cases} w' & \text{if } g'[w, \{(p_{k_1}, w_{k_1}), \ldots, (p_{k_j}, w_{k_j})\}] = w' \\ & \text{and } w_{k_i} = w_{k_m}\ \forall i, m \in [1, j] \\ \emptyset & \text{otherwise} \end{cases} \tag{A.2.38}$$

Define $h'_1 : W \times \mathcal{P}(\Phi_n \times \mathcal{P}(W)) \to \mathcal{P}(W)$ by

$$\begin{aligned} h'_1[w, \{(p_{k_1}, U'_1), \ldots, (p_{k_j}, U'_j)\}] &= \bigcup_{v_1 \in U'_1} \ldots \bigcup_{v_j \in U'_j} h'_0[w, \{(p_{k_1}, v_1), \ldots, (p_{k_j}, v_j)\}] \\ &= \bigcup_{w' \in \bigcap_{i=1}^{j} U'_i} h'_0[w, \{(p_{k_1}, w'), \ldots, (p_{k_j}, w')\}] \end{aligned} \tag{A.2.39}$$

Define $h'_2 : W \times \mathcal{P}(\Phi_n \times \mathcal{P}(W)) \to \mathcal{P}(W)$ by

$$\begin{aligned} &h'_2[w, \{(p_{k_1}, U_1), \ldots, (p_{k_j}, U_j)\}] \\ = \ &h'_1[w, \{(p_{k_1}, c(w, p_{k_1}, U_1)), \ldots, (p_{k_j}, c(w, p_{k_j}, U_j))\}] \end{aligned} \tag{A.2.40}$$

Finally $h'_3 : \mathcal{P}(W) \times \mathcal{P}(\Phi_n \times \mathcal{P}(W)) \to \mathcal{P}(W)$ by

$$h'_3(U, \{(p_{k_1}, U'_1), \ldots, (p_{k_j}, U'_j)\}) = \bigcup_{w \in U} h'_2[w, \{(p_{k_1}, U'_1), \ldots, (p_{k_j}, U'_j)\}] \tag{A.2.41}$$

that provides us with the general expression required to calculate any uniform substitution using the many world formalism in $\mathrm{PropCal}_n$

$$\begin{aligned} [\eta] &= \left[\psi\left(\phi_1/p_{k_1}; \phi_2/p_{k_2} \ldots \phi_j/p_{k_j}\right)\right] \\ &= h'_3([\psi], \{(p_{k_1}, [\phi_1]), \ldots, (p_{k_j}, [\phi_j])\}) \end{aligned} \tag{A.2.42}$$

This completes the section on substitution using the semantic representation. One now has the tools to numerically simulate the results proven in this thesis.

Bibliography

[1] M. Abercrombie, C.J. Hickman, and M.L. Johnson, *A dictionary of biology*, Penguin Books, 1975.

[2] M. Abramowitz and I.A. Stegum, *Handbook of mathematical functions*, Dover Publications, inc., New York, 1965.

[3] C.E. Alchourrón, P. Gärdenfors, and D. Makinson, *On the logic of theory change: Partial meet contraction and revision functions*, The Journal of Symbolic Logic **50** (1985), 510–530.

[4] J.J. Alferes, J.A. Leite, L.M. Pereira, H. Przymusinska, and T.C. Przymusinska, *Dynamic logic programming*, Linköping electronic articles in computer and information science, http://www.ep.liu.se/ea/cis/1997/018/.

[5] J.J. Alferes and L.M. Pereira, *Logic programming updating - a guided tour*, citeseer.ist.psu.edu/383707.html.

[6] E. Berlekamp, J. Conway, and R. Guy, *Winning ways for your mathematical plays*, vol. 2, AcademicPress, New York, 1982.

[7] P. Blackburn, M. de Rijke, and Y. Venema, *Modal logic*, Cambridge University Press, Cambridge, 2001.

[8] C. Brink and I. Rewitzky, *A paradigm for program semantics: Power structures and duality*, CSLI Publications, Stanford, 2001.

[9] A.W. Burks (ed.), *Essays on cellular automata*, University of Illinois Press, Urbana, Chicago, London, 1970.

[10] R.T. Cahill and C.M. Klinger, *Bootstrap universe from self-referential noise*, General Relativity and Quantum Cosmology **arXiv:gr-qc/9708013** (1997).

[11] R.T. Cahill, C.M. Klinger, and K. Kitto, *Pregeometric modelling of the spacetime henomenology*, General Relativity and Quantum Cosmology **ArXiv:gr-qc/9605018** (1996).

[12] ———, *Process physics: Modelling reality as self-organising information*, General Relativity and Quantum Cosmology **ArXiv:gr-qc/0009023** (2000).

[13] A.G. Cairns-Smith and H. Hartman (eds.), *Clay minerals and the origin of life*, Cambridge University Press, Cambridge, 1986.

[14] C. Darwin, *The origin of the species*, 1852.

[15] R. Dawkins, *The selfish gene*, Oxford University Press, Oxford, 1976.

[16] ———, *The extended phenotype*, Oxford University Press, Oxford, 1982.

[17] ———, *The blind watchmaker*, Penguin Books Ltd, England, 1986.

[18] M. di Gregorio, *Darwin: Evolution by natural selection*, Lecture, UCT Summer School, 2004.

[19] S.N. Dorogovtsev and J.F.F. Mendes, *Evolution of networks*, arXiv:Cond-mat/0106144 (2001).

[20] G.M. Edelman, *Neural Darwinism*, Basic Books, New York, 1987.

[21] A. Einstein, *Physics and reality*, Journal of the Franklin Institute **221** (1936), 378.

[22] A. Einstein, H.A. Lorentz, H. Minkowski, and H. Weyl, *The principle of relativity*, Dover Publications, Inc., New York, 1952.

[23] G.F.R. Ellis, *Science, complexity and the natures of existence*, University of Cape Town, Seminar, 2004.

[24] R.P. Feynman, *Space-time approach to non-relativistic quantum mechanics*, Reviews of Modern Physics **20** (1948), 367–387.

[25] R.P. Feynman and A.R. Hibbs, *Quantum mechanics and path integrals*, McGraw-Hill Book Company, New York, 1965.

[26] D.B. Fogel, *Evolutionary computation: Toward a new philosophy of machine intelligence*, IEEE Press, New Jersey, 1995.

[27] P. Gibbs, *The small scale structure of spacetime: A bibliographical review*, arXiv:hep-th/9506171 (1995).

[28] C. Godsil and G. Royle, *Algebraic graph theory*, Springer-Verlag, New York, 2001.

[29] B. Goertzel, *The evolving mind*, Gordon and Breach, 1993.

[30] B. Goertzel, T. Smith, O. Aam, and K. Palmer, *Universe as network*, (1998), http://www.goertzel.org/papers/eventnet.html.

[31] M. Goldszmidt and J. Pearl, *Qualitative probabilities for default reasoning, belief revision and causal modelling*, Artificial Intelligence **84** (1996), 57–112.

[32] A. Grafen, *Developments of the price equation and natural selection under uncertainty*, Proceedings of the Royal Society **267** (2000), 1223–1227.

[33] D. Gries and F.B. Schneider, *A logical approach to discrete math*, Springer-Verlag, New York, 1993.

[34] P.E. Griffiths and R.D. Gray, *The developmental systems perspective: Organism - environment systems as units of development and evolution*, The Evolutionary Biology of Complex Phenotypes.

[35] D. Halliday and R. Resnick, *Fundamentals of physics*, Wiley, New York, 1974.

[36] S.W. Hawking, *Nature of space and time*, arXiv:hep-th/9409195 (1994).

[37] E.S. Herman and N. Chomsky, *Manufacturing consent: The political economy of the mass media*, Vintage, London, 1994.

[38] J.H. Holland, *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control and artificial intelligence*, MIT Press, Cambridge, London, 1992.

[39] ———, *Emergence: From chaos to order*, Oxford University Press, Oxford, 1998.

[40] K. Culik II, L.P. Hurd, and S. Yu, *Computation theoretic aspects of cellular automata*, Journal of Physics D **45** (1990), 357–378.

[41] A. Ilachinski and P. Halpern, *Structurally dynamic cellular automata*, Complex Systems **1** (1987), 503–527.

[42] A. Korzeniowski, J.L. Fry, D.E. Orr, and N.G. Fazleev, *Feynman-kac path-integral calculation of the ground-state energies of atoms*, Physical Review Letters **69** (1992), 893–896, See references herein.

[43] C.G. Langton, *Self-reproduction in cellular automata*, Physica D **10** (1984), 135–144.

[44] E. Lawrence (ed.), *Henderson's dictionary of biological terms*, Longman Scientific and Technical, London, 1995.

[45] J.A. Leite, J.J. Alferes, and L.M. Pereira, *Multi-dimensional dynamic logic programming*, CL-2000 Workshop on Computational Logic in Multi-Agent Systems, 2000, citeseer.ist.psu.edu/article/leite00multidimensional.html, pp. 17–26.

[46] J.A. Leite, J.J.Alferes, and L.M. Pereira, *Multidimensional dynamic knowledge representation*, http://centria.di.fct.unl.pt/∼jleite/papers/lpnmr01.pdf.

[47] R.C. Lewontin, *The units of selection*, Annual Review of Ecology and Systematics **1** (1970), 1–14.

[48] V. Marek and M. Truszczyński, *Jelia'94: Revision specifications by means of programs*, vol. 838, Springer-Verlag, 1994.

[49] J. McCarthy, *Nonmonotinicity*, http://www.aaai.org/AItopics/html/nonmon.html.

[50] D. McDermott and J. Doyle, *Nonmonotonic logic 1*, Artificial Intelligence **13** (1980), 41–72.

[51] D.A. Meyer, *Quantum stochastics: The passage from relativistic to a non-relativistic path integral*, Journal Physics A: Math. Gen. **17** (1984), 375–383.

[52] ———, *From quantum cellular automata to quantum lattice gases*, Journal Statistical Physics **85** (1996), 551–574, quant-ph/9604003.

[53] P. Meyer, *Lattice geometries*, (2001), http://hermetic.nofadz.com/compsci/lattgeom.htm.

[54] J. Minker, *An overview of nonmonotonic reasoning and logic programming*, University of Maryland Technical Report **UMIACS TR-91-112, CS-TR-2736** (1991), 31pp.

[55] C.W. Misner, K.S. Thorne, and J.A. Wheeler, *Gravitation*, W.H. Freeman, San Francisco, 1973, Not geometry, but pregeometry as the magic building material; Pregeometry as the calculus of propositions; The Black Box: The reprocessing of the Universe.

[56] R. Moore, *Semantical considerations on nonmonotonic logic*, Artificial Intelligence **25(1)** (1985), 75–94.

[57] G. Nagels, *Space as a "bucket of dust"*, General Relativity and Gravitation **17(6)** (1985), 545–557.

[58] M. Nakahara, *Geometry, topology and physics*, IOP Publishing Ltd, UK, 1990.

[59] T. Nowotny and M. Requardt, *Pregeometric concepts on graphs and cellular networks as possible models of space-time at the planck-scale*, Journal of Chaos, Solitons and Fractals (1998), hep-th/9801199.

[60] P. Peppas, A. Nayak, M. Pagnucco, N.Y. Foo, R. Kwok, and M. Prokopenko, *Revision versus update*, John Wiley and Sons Ltd, 1996, ECAI 96. 12th European Conference on Artificial Intelligence.

[61] G.R. Price, *Selection and covariance*, Nature **227** (1970), 520–521.

[62] ———, *Extension of covariance selection mathematics*, Annals of Human Genetics **35** (1972), 485–490.

[63] E.M. Purcell, *Electricity and magnetism*, McGraw-Hill Inc., 1985.

[64] A. Ramsay, *Formal methods in artificial intelligence*, Cambridge University Press, Cambridge, 1988.

[65] J.A. Reggia, S.L. Armentrout, H.-H. Chou, and Y. Peng, *Simple systems that exhibit self-directed replication*, Science **259** (1993), 1282–1287.

[66] J.A. Reggia and H.-H. Chou, *Emergence of self-replicating structures in cellular automata space*, Physica D **110** (1997), 252–276.

[67] J.A. Reggia, H.-H. Chou, and J.D. Lohn, *Cellular automata models of self-relicating systems*, Advances in Computers **47** (1998), 141–183.

[68] R. Reiter, *A logic for default reasoning*, Artificial Intelligence **13** (1980), 81–132.

[69] H. Rund, *The Hamilton-Jacobi theory in the calculus of variations*, R.E. Krieger Publishing Company, 1973.

[70] D. Shapiro and W.C. Henneberger, *The Aharonov-Bohm effect in double and single-slit diffraction*, Journal Physics A **22** (1989), 3605–3611.

[71] T. Smith, *From sets to quarks: Deriving the standard model plus gravitation from simple operations on finite sets*, arXiv:hep-ph/9708379 (1998).

[72] L. Smolin, *Did the universe evolve?*, Classical and Quantum Gravity **9** (1992), 173–191.

[73] ———, *The fate of black hole singularities and the parameters of the standard models of particle physics and cosmology*, General Relativity and Quantum Cosmology **arXiv:gr-qc/9404011** (1994).

[74] ———, *Using neutron stars and primordial black holes to test theories of quantum gravity*, General Relativity and Quantum Cosmology **ArXiv:gr-qc/9712189** (1998).

[75] L. Smolin, R. Dawkins, N. Humphrey, B. Goodwin, J. Lanier, G. Johnson, and M. Gleiser, *Implications of natural selection and the laws of physics*, www.edge.org (2004), http://www.edge.org/discourse/smolin_natselection.html.

[76] J. Urias, *Arithmetic representations of cellular automata*, Journal of Physics D **68** (1993), 437–446.

[77] R.M. Wald, *General relativity*, The University of Chicago Press, Chicago, 1984.

[78] J.A Wheeler, *Geometrodynamics*, Academic Press, New York and London, 1962, Topics in Modern Physics: Volume 1.

[79] J.A. Wheeler, *On recognising "law without law"*, American Journal of Physics **51** (1983), 398–404.

[80] L.H. Yang and R.E. Page, *The evolution of sex*, (2004), http://entomology.ucdavis.edu/courses/ent104/docs/11_The_Evolution_of_Sex.pdf.