

THE TRANSFORMATIONAL COMPLEXITY OF ACYCLIC NETWORKS
OF NEURONS

By

VENKATAKRISHNAN RAMASWAMY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2011

© 2011 Venkatakrisnan Ramaswamy

To Appa and Amma

ACKNOWLEDGMENTS

I feel a certain tinge of sadness as I drive this final nail in the dissertation, for I have had a wonderful time in grad school. A large measure of credit for this goes to my advisor Arunava Banerjee. I owe him a debt of gratitude for his kindness, his collegiality and his appetite for the big questions and ambitious attacks that made life interesting and worthwhile. I don't know how or why, but I got a virtual carte blanche, right from the beginning, to pursue directions that interested me, even during times when progress was slow or nonexistent. I used to joke that his office is the opposite of IHOP[®]¹; you come happy and leave hungry. Few ideas survived the crucible of his intellect, but those that did were worth pursuing. I am also grateful for his amazing ability to be uplifting. When the chips are down and things seem hopeless, a visit to his office was sufficient to get your mojo back. He also has a curious method of accomplishing this; by relating anecdotes of times when he was in even more dire straits.

I thank Arunava Banerjee, Mingzou Ding, Alireza Entezari, Jeffrey Ho, Meera Sitharam and Don Stehouwer for serving on my supervisory committee. I am especially thankful to Jeffrey Ho for being very generous with his time a couple of years ago when I was pursuing an attack that did not bear fruit (yet).

This work was supported, in part, by a grant (IIS-0902230) from the National Science Foundation to Arunava Banerjee, which I gratefully acknowledge. I am also thankful to the CISE Department for supporting me through Teaching Assistantships over the years. I also benefited from an opportunity to teach a few undergraduate courses, for which I thank Steve Thebaut. I am thankful to John Bowers and Joan Crisman for their kindness and all their help with matters administrative.

¹ International House of Pancakes; a restaurant with the slogan, "Come hungry, leave happy.[®]"

I wish to thank MIT Press for permission to reproduce Figure 2-1, Figure 2-2 and Figure 2-3 that appear in (Dayan & Abbott, 2005). Pamela Quick made the process unbelievably easy.

I would like to thank Arunava Banerjee, Jeffrey Ho, Anand Rangarajan, Meera Sitharam and Baba Vemuri for their excellent and motivating courses in Computer Science and for being patient with my questions. I am also thankful to Mike Jury, Alexandre Turull and Peter Sin for their excellent math courses.

I wish to thank my housemates over the years, Aashish Parikh, Shivkumar Swaminathan, Vikas Aggarwal, Karthik Nagarajan, Swami Sankaran, Ganesh Viswanathan, Sanketh Bhat, Avinash Deshpande, Harish Bharadwaj, Jignesh Soni and Abhijeet Lawande for tolerating me and for interesting conversations, food and fun. Special thanks to Vikas Aggarwal for putting up with me (pun intended) the longest.

I am also grateful to all my friends. Ajit Rajwade, being a co-sufferer right through, helped blunt some of the despondency that goes with the territory. Mohsen Ali, Jason Chi, Amit Dhurandhar, Neko Fisher, Karthik Gurumoorthy, Ritwik Kumar, Shahed Nejhum, Subhajit Sengupta, Manu Sethi and Nathan Vanderkraats were people I bumped into either in the hallway or in the lab on a daily basis. Thank you for all the banter and the bull sessions. And, a big thank you to the rest of my friends, too numerous to list here. You know who you are !

I will always remain grateful to N.S. Narayanaswamy, for offering me my first taste of research. That year working on Satisfiability algorithms with Nari was very formative and introduced me to the subtle joys and unsubtle disappointments that come packaged with hard problems. On a recent meeting, I realized how much Nari has influenced the way I think about problems and the questions I ask.

I have used free software almost exclusively for a number of years now. I thank the free software/open source community for preserving our freedom.

But most of all, I am thankful to my parents, C.V. Ramaswamy and Saraswathi for all that I am. Their love, affection and pursuit of high ideals have made life worth living. I am grateful for all the sacrifices they have made over the years for the sake of my education and well-being. I am especially thankful to my father for instilling in me a love for Mathematics from an early age and my mother for being my moral compass. I thank my sister Usha for her affection, curiosity and teaching me about computers and programming, back when I was in high school. Paradoxically, if not for that early exposure (and subsequent satiation), I would probably still be doing core Computer Science and not the present type of work. And, finally, I thank you the reader !

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	4
LIST OF FIGURES	9
ABSTRACT	11
CHAPTER	
1 INTRODUCTION	13
2 BACKGROUND: NEURONAL BIOPHYSICS	17
2.1 Introduction	17
2.2 Neuron: Basic Structure and Function	17
2.3 The Cell Membrane and the Membrane Potential	20
2.4 Passive Conductance in Dendrites - Cable Theory	22
2.5 Active Conductance in the Axon - Hodgkin Huxley Equations	24
2.6 Other Varieties of Neuronal Behavior	26
3 MODEL	28
3.1 Notation and Preliminaries	28
3.2 Assumptions Underlying the Model	30
3.3 Formal Definition of the Model	32
4 ACYCLIC NETWORKS AS SPIKE-TRAIN TO SPIKE-TRAIN TRANSFORMATIONS	34
4.1 On the Consistency of Defining Input/Output Transformations on Neurons	34
4.2 Gap Criteria	38
4.2.1 Gap Lemma	38
4.2.2 Gap Criterion for a Neuron	41
4.2.3 Practical Issues with the Gap Criteria	45
4.3 Flush Criterion	46
5 TRANSFORMATIONAL COMPLEXITY: DEFINITIONS AND RESULTS	49
5.1 Motivation and Definition	49
5.2 Gap-Flush Equivalence Lemma	49
5.3 Complexity Results for the Abstract Model	51
6 COMPLEXITY RESULTS AND NETWORK DEPTH: A BARRIER AT DEPTH TWO	60
6.1 Technical Structure of the Proof	61
6.2 Causal, Time-Invariant and Resettable Transformations	61

6.3	Construction of a Depth Two Acyclic Network for every Causal, Time-Invariant and Resettable Transformation	65
6.4	Directions for Further Constraining the Abstract Model	76
6.5	Some Auxiliary Propositions	76
7	COMPLEXITY CLASSES, TRANSFORMATION HIERARCHIES AND HIERARCHY CLASSES	79
7.1	Definitions and Preliminaries	80
7.2	Lower Bounds on the Hierarchy Classes of some Architectures in a certain Transformation Hierarchy	85
8	CONCLUSION AND FUTURE WORK	90
8.1	Contributions	90
8.2	Directions for Future Work	91
8.3	Epilogue	92
	REFERENCES	93
	BIOGRAPHICAL SKETCH	95

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1 Diagrams showing morphology of three neurons.	18
2-2 Action potentials and the synapse.	19
2-3 A schematic diagram of a section of the lipid bilayer that forms the cell membrane of the neuron with two ion channels embedded in it.	21
2-4 Schematic diagram of the equivalent circuit for a passive membrane.	23
4-1 Counterexample to the claim that the membrane potential of a neuron can be made a function of a bounded window of input spikes alone.	35
4-2 Counterexample to the claim that the membrane potential of a neuron can be made a function of input spikes alone in the infinite past.	36
4-3 The idea behind the Gap Lemma.	38
4-4 Illustration accompanying the Gap Lemma.	39
4-5 Illustration demonstrating that for an input spike configuration χ that satisfies a T -Gap criterion, the membrane potential at any point in time is dependent on atmost T seconds of input spikes in χ before it.	42
4-6 Schematic diagram illustrating how the Gap criterion works for a simple two-neuron network.	44
4-7 Illustration showing that an input spike configuration satisfying a Flush Criterion also satisfies a Gap Criterion.	47
5-1 Example of a transformation that no acyclic network can effect.	52
5-2 A transformation that no single neuron can effect, that a network with two neurons can.	53
5-3 A two-neuron network effecting the transformation prescribed in Figure 5-2. . .	55
5-4 A transformation that no network with a path-disjoint architecture can effect. .	56
5-5 A two-neuron network effecting the transformation prescribed in Figure 5-4. . .	58
6-1 Network architecture for $m = 2$	66
6-2 Example illustrating the operation of l	70
7-1 Venn Diagram illustrating the basic notions of a transformation hierarchy, complexity class and hierarchy class.	80
7-2 Venn Diagram showing complexity classes when one set of networks is more	

	complex than another.	81
7-3	Venn Diagram illustrating a transformation hierarchy.	83
7-4	Venn Diagram demonstrating how upper bounds and lower bounds on hierarchy classes in a transformation hierarchy can be used to establish complexity results.	84
7-5	Diagram depicting architecture of networks in Σ_2	87
7-6	Diagram depicting architecture of networks in Σ_3	88

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

THE TRANSFORMATIONAL COMPLEXITY OF ACYCLIC NETWORKS
OF NEURONS

By

Venkatakrishnan Ramaswamy

May 2011

Chair: Arunava Banerjee
Major: Computer Engineering

Neurons are cells in the brain thought to be primarily responsible for information processing and cognition. Neurons communicate with each other via stereotypical electric pulses called action potentials or spikes. The exact computational mechanisms used by the brain for information processing are not yet well understood. The current work seeks to reduce this gap in our understanding, by theoretical analysis.

Our thesis is that specific properties of neurons and the network's architecture constrain the computations it might be able to perform. This dissertation substantiates the thesis for the case of acyclic networks of neurons.

In particular, the idea is to make precise the intuitive view that, acyclic networks are transformations that map input spike trains to output spike trains, to ask what transformations acyclic networks of specific architectures cannot accomplish.

Our neurons are abstract mathematical objects that satisfy a small number of axioms, which correspond to basic properties of biological neurons. To begin with, we find that even a single neuron cannot be consistently viewed as a spike-train to spike-train transformation, in general (in a sense that we will make precise). However, under conditions consistent with spiking regimes observed in-vivo, we prove that the aforementioned notions of transformations are indeed well-defined and correspond to mapping finite-length input spike trains to finite-length output spike trains. Armed with this framework, we then ask what transformations acyclic networks of specific

architectures cannot accomplish. We show such results for certain classes of architectures. While attempting to ask how increase in depth of the network constrains the transformations it can effect, we were surprised to discover that, with the current abstract model, every acyclic network has an equivalent acyclic network of depth two. What this suggests is that more axioms need to be added to the abstract model in order to obtain results in this direction. Finally, we study the space of spike-train to spike-train transformations and develop some more theoretical tools to facilitate this line of investigation.

CHAPTER 1 INTRODUCTION

Neurons and their networks, fundamentally, are machines that transform spike trains into spike trains. It is these transformations that form the basis of information processing, indeed even cognition, in the brain.

This work is broadly motivated by the following question: What constraints, if any, do local properties of neurons impose on the computational abilities of various networks? At a coarse level, we would like to say something about what specific networks “cannot” do, by virtue of their architecture alone. Now the answer may depend on the specific neuron model in use. To mitigate this possibility, we instead assume that the neuron is an abstract object that satisfies certain axioms, which correspond well-known properties of biological neurons. This admits a variety of models as long as they satisfy the axioms (including ones that may be more powerful than what is biologically possible). We then ask, if (in spite of this possibly additional power), there exist specific transformations that cannot be effected by networks of certain architectures that are equipped with these abstract neurons. In order to then rule out the prospect that the transformation in question is so hard that no network can do it, we also provide a network (of a different architecture) comprising simple neurons that can in fact effect this transformation. It is such considerations that motivate our definition of transformational complexity. These notions are made precise in the remainder of the dissertation.

Since the functional role of single neurons and small networks in the brain is not yet well understood, we do not make any assumptions about particular “tasks” that the network is trying to perform; we are just interested in physical spike-train to spike-train transformations. Furthermore, since the kinds of neural code employed is still unclear, we make no overarching assumptions about the neural code either.

We restrict our study to acyclic¹ networks of neurons, i.e. networks that do not have a directed cycle. While even single neurons are functionally recurrent², acyclic networks quickly settle down to quiescence upon receiving no input. On the other hand, recurrent networks have been known (Banerjee, 2006) to have complex dynamics, in general, even on receiving no input for unbounded periods of time. Also, we do not treat synaptic plasticity effects here.

In the past, a large body of work has been done assuming neurons and their networks are entities for which only the instantaneous rate of incoming spikes and outgoing spikes is relevant. The theory of perceptrons (Rosenblatt, 1988; Minsky & Papert, 1969) and multilayer perceptrons (Rumelhart et al., 1986), for instance, makes such assumptions. However, there is widespread evidence (Strehler & Lestienne, 1986; Rieke et al., 1997) that precise spike times play a role in information processing in the brain. Therefore rate-based models may not adequately explain all aspects of brain function.

Several researchers have studied related questions. In (Poirazi et al., 2003), Poirazi et al., model a compartmental model of a pyramidal neuron using a two layer neural network, assuming rate codes. Bohte et al., (Bohte et al., 2002) derive a supervised learning rule for a network of spiking neurons, where the output is restricted to a single spike in a given period of observation. Gutig and Sompolinsky (Gutig & Sompolinsky, 2006) describe a model to learn spike time decisions. They also have a task with two outcomes, which are mapped to notions of a presence or absence of spikes. Maass (Maass, 1996) investigates the computational power of networks of neurons relating them to well-known models of computation such as Turing Machines and Random Access Machines. Finally Bartlett and

¹ while the term “feedforward” network is widely used in the literature to refer to this type of network, we prefer to call these acyclic networks to emphasize that these networks are not feedforward in the system-theoretic sense.

² owing to the membrane potential also depending on past output spikes to account for effects during the absolute and relative refractory periods

Maass ([Maass, 1995](#)) analyze the discriminative capacity of a pulse-coded neuron from a statistical learning theory perspective.

In contrast to all the above approaches, we seek to investigate the relative complexity of the physical spike train to spike train transformations that are instantiated by systems of spiking neurons, without making overarching assumptions about the underlying computational dynamics of the system. Our results are therefore at a level more fundamental than the computational framework most other work assumes and are therefore more widely applicable.

The remainder of the dissertation is organized as follows. In Chapter 2, we review neurobiology relevant to understanding and evaluating the assumptions of our model, which is elaborated in Chapter 3. Armed with this model, in Chapter 4, we seek to study acyclic networks as spike-train to spike-train transformations. To our surprise, we find that the notion is not always well-defined. We provide counterexamples that illustrate such cases. This pessimistic view is then surmounted by the recognition that under spiking regimes usually seen in the brain, one can consistently treat acyclic networks as the said transformations. The rest of the chapter is devoted to rigorously defining sufficient conditions where this is the case and proving that notions of transformations do hold under these conditions. We then proceed, in Chapter 5, to define notions of complexity to address questions of what transformations specific architectures cannot accomplish that networks of other architectures can. We also prove an important technical lemma that significantly eases the task of proving complexity results. The rest of the chapter then establishes some results of this form for some specific classes of architectures. While attempting to ask how the set of transformations effected is constrained when the depth of the network is increased, we were surprised to find that the abstract model formulated in Chapter 3 does not so constrain it for networks with depth equal to two or more. Thus, Chapter 6 is devoted to proving that with the current abstract model, every acyclic network has an equivalent acyclic network of depth two, that effects exactly the same

transformation. In Chapter 7, we develop some additional theoretical tools to answer questions of this type. Specifically, we study the space of all spike-train to spike-train transformations, define notions of a hierarchy of sets in that space and try to relate sets in such hierarchies to the set of transformations spanned by specific network architectures. We conclude in Chapter 8 with a discussion and describe some directions for future work.

CHAPTER 2 BACKGROUND: NEURONAL BIOPHYSICS

2.1 Introduction

The Human Brain is one of the most remarkable entities in the Universe. This approximately 3 pound mass of soft tissue has not only ensured its survival, but enabled it to rule the earth, explore the universe and indeed attempt to understand itself. Understanding the way the brain works is one of the last great frontiers in science.

The two predominant type of cells in the brain are neurons and glia. Glia are believed primarily to be support cells that provide structural support and nourishment to neurons (Bear et al., 2007).¹ Neurons are believed to be the cells responsible for information processing and cognition. The human brain is estimated to have a 100 billion neurons and a 100 trillion connections between them.

The purpose of this chapter is to provide the reader unfamiliar with Neurobiology, the necessary background to understand the context of this work and more specifically to evaluate the abstract model of the neuron proposed in Chapter 3. We start with a brief description of neuron function before delving into more detailed biophysics. The exposition here is based on material from (Dayan & Abbott, 2005) and (Banerjee, 2001b).

2.2 Neuron: Basic Structure and Function

In this section, we briefly describe how a neuron works. The goal is to give the reader a sense of the big picture, to facilitate the reading of subsequent sections where the biophysics is described in much greater detail.

A neuron is a cell which receives input from certain neurons and communicates its output to certain other neurons. Figure 2-1 illustrates morphological regions in various types of neurons. These diagrams, circa 1911, were hand-drawn by the pioneering Spanish histologist Santiago Ramon y Cajal. The three primary morphological regions in a neuron

¹ Interestingly, glia outnumber neurons by a factor of 10.

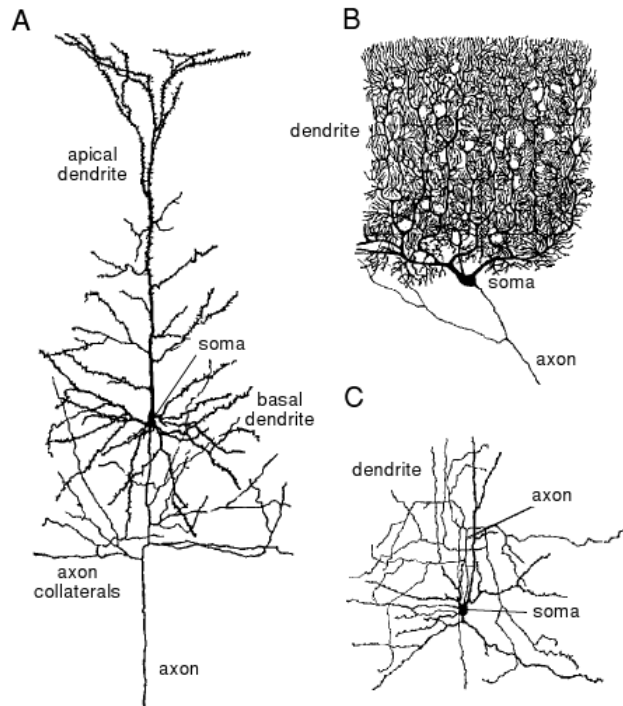


Figure 2-1. Diagrams showing morphology of three neurons. A) A cortical pyramidal neuron. B) A Purkinje cell of the cerebellum. C) A stellate cell of the cerebral cortex. [Reprinted with permission from (Dayan & Abbott, 2005) (Page 5, Figure 1.1).]

are dendrites, axon and the soma. The dendrites receive input from other neurons, usually through elaborate dendritic trees. The neuron communicates its output to other neurons via the axon and its collaterals. Axons often tend to cover large distances including entire lengths of bodies of the organism.

Neurons communicate with other neurons by firing stereotypical pulses called “action potentials” or “spikes”. Spikes have a stereotypical shape; it is only their timing that is important. Neurons, therefore represent and communicate information via temporal patterns of spikes. The generation of spikes is mediated by the potential difference between the inside of the neuron and outside, called the “membrane potential”. The membrane potential is an internal variable, in the sense that it is not explicitly communicated to other neurons. Under resting conditions (i.e. when the neuron does not receive input for sufficiently long), the membrane potential is maintained at around

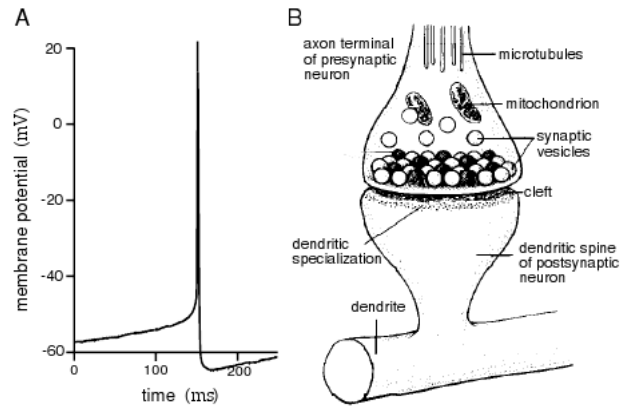


Figure 2-2. Action potentials and the synapse.

A) An action potential recorded intracellularly from a cultured rat neocortical pyramidal cell. (Recorded by L. Rutherford in the lab of G. Turrigiano).

B) Diagram of a synapse. [Reprinted with permission from (Dayan & Abbott, 2005) (Page 6, Figure 1.2).]

-70 mV. This is called the “resting membrane potential” or simply the resting potential and the neuron is said to be “polarized” at this stage.

Input spikes received from other neurons modulate this potential. They may either elevate the potential in which case they are called “excitatory” or depress the potential, in which case they are called “inhibitory”. Elevation of the membrane potential above the resting potential is called “depolarization” and depressing it below the resting potential is called “hyperpolarization”. If a neuron is sufficiently depolarized, so that its membrane potential approaches a certain “threshold” (from below), the neuron produces a spike which travels down the axon and is communicated to neurons that the present neuron is connected to. The stereotypical change in membrane potential that an action potential causes is shown in Figure 2-2(A). Once the neuron spikes, due to physiological constraints, it cannot spike for a certain small period of time called the “absolute refractory period” which is typically of the order of a few milliseconds. Beyond this time, for a certain interval of time, the threshold to fire is elevated, i.e. more than usual depolarization is required to elicit a spike. This period is called the “relative refractory period” and is usually the order of tens of milliseconds. Action potentials are the only type of membrane

fluctuation that are communicated without attenuation by a positive feedback process which is a consequence of the the unique electrophysiological properties of the cell membrane on the axon. This is expounded in greater detail later when we discuss the biophysics of the active membrane.

Axons contain “synapses” at which dendrites of other neurons make contact with the present neuron. Those neurons are said to be “postsynaptic” to the present neuron. A schematic diagram of a synapse is provided in Figure 2-2(B). The generation of a spike on the current neuron is communicated to postsynaptic neurons via the synapse. As illustrated, the synapse consists of two boutons in close proximity to each other with a small gap in between called the “synaptic cleft”. The boutons on the presynaptic side have neurotransmitters stored in small sacs called “vesicles”. On generation of a spike, neurotransmitters are released into the synaptic cleft which bind to receptors on the postsynaptic side and cause modulation of the membrane potential of the postsynaptic neuron. This process is called “synaptic transmission”.

Armed with this broad outline of how neurons work, we treat its biophysics in more detail in the rest of this chapter. More detailed expositions of these topics are available in (Tuckwell, 1988; Koch, 2004; Squire et al., 2003). For a first look at Neurobiology, the reader is referred to (Bear et al., 2007).

2.3 The Cell Membrane and the Membrane Potential

The cell membrane of a neuron is a lipid bilayer with thickness approximately 3nm. Figure 2-3 provides an illustration. The cell membrane is impermeable to most ions. Embedded in the membrane are various water-filled protein molecules that act as “ion channels”, which allow specific ions to pass through them. These ion channels are usually very selective in that they only allow certain types of ions to pass under certain conditions. These conditions could include specific ranges of membrane potential when the channels are activated, or presence of intracellular messengers such as Ca^{2+} or extracellular entities such as neurotransmitters. A typical neuron may have tens of different types of ion

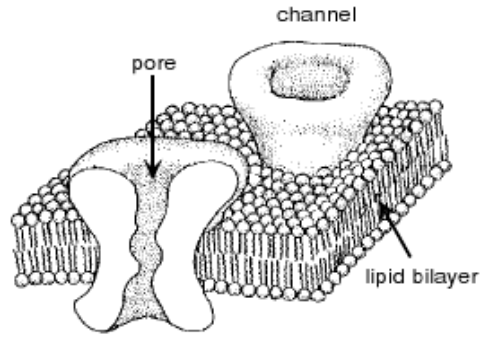


Figure 2-3. A schematic diagram of a section of the lipid bilayer that forms the cell membrane of the neuron with two ion channels embedded in it. [Reprinted with permission from (Dayan & Abbott, 2005)(Page 154, Figure 5.1).]

channels and of the order of millions of channels in it. The membrane also contains “ion pumps” that constantly expend energy to exchange ions in order to maintain a certain difference in concentration between the inside and the outside of the neuron.

As mentioned earlier, the resting potential is about -70 mV, although this number varies depending on the organism and the type of neuron under consideration. We now explore the underlying mechanisms that are responsible for maintaining this potential difference. The primary cause is the different concentration of Na^+ and K^+ ions across the membrane. The outside of the neuron has a higher concentration of Na^+ than the inside. Likewise, the inside of the neuron has a higher concentration of K^+ than the outside. This concentration gradient is maintained by $\text{Na}^+ - \text{K}^+$ ion pumps that send out 3 Na^+ ions for every 2 K^+ taken in. Furthermore, the membrane is semi-permeable to these ions via ion channels. This allows for diffusion of ions across the membrane, the rate of diffusion being a function of the concentration gradient, the permeability of the ion in question and the electric field across the membrane. Therefore, K^+ ions tend to diffuse out of the cell and Na^+ ions tend to diffuse into the cell. However, at rest, the permeability of Na^+ is much lower than that of K^+ and therefore more K^+ ions are expelled than Na^+ ions admitted, which leads to an equilibrium membrane potential in which the inside of the neuron is negatively charged with respect to the outside. Also, Cl^- ions are present in

higher concentration outside. In these circumstances, the equilibrium membrane potential can be computed using the Goldman-Hodgkin-Katz Equation ([Goldman, 1943](#)):

$$V_m = \frac{RT}{F} \ln \frac{P_K[K^+]_o + P_{Na}[Na^+]_o + P_{Cl}[Cl^-]_i}{P_K[K^+]_i + P_{Na}[Na^+]_i + P_{Cl}[Cl^-]_o} \quad (2-1)$$

where R is the gas constant, T the absolute temperature, F Faraday's constant, $[\cdot]_i$ s and $[\cdot]_o$ s the concentration of the corresponding ions on the inside and outside respectively and P s the permeability of the membrane to the ion in question.

2.4 Passive Conductance in Dendrites - Cable Theory

Change in potentials in a postsynaptic neuron due to an incoming spike are conducted passively across the dendrite en route to the soma. This process is modeled using Cable Theory ([Rall, 1960](#)).

The dendrite is modeled as a cylinder of uniform radius with membrane properties such as capacitance and resistivity being uniform. Let V_m be the membrane potential along the axis of the cable, $r = r_i + r_o$ be the resistance per unit length of cable and i the current flowing across the axis. Let ΔV_m be the potential dropped along Δx of the cable. Then, by Ohm's Law, we have,

$$\Delta V_m = -ri\Delta x \quad (2-2)$$

Taking $\Delta x \rightarrow 0$, we have,

$$\frac{\partial V_m}{\partial x} = -ri$$

and therefore,

$$\frac{\partial^2 V_m}{\partial x^2} = -r \frac{\partial i}{\partial x} \quad (2-3)$$

Now, let i_m be the current flowing perpendicular to the axis of the cable, per unit length of cable. Let Δi be the change in current for a distance Δx along the axis of the cable. By Kirchoff's First Law, we have

$$\Delta i = -i_m \Delta x \quad (2-4)$$

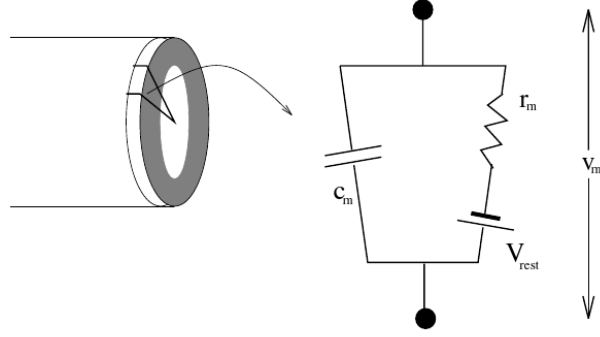


Figure 2-4. Schematic diagram of the equivalent circuit for a passive membrane.
 [Reprinted with permission from (Banerjee, 2001b)(Page 28, Figure 3.2).]

Again, taking $\Delta x \rightarrow 0$, we have,

$$i_m = -\frac{\partial i}{\partial x} \quad (2-5)$$

Substituting Equation 2-5 in Equation 2.4, we have

$$\frac{\partial^2 V_m}{\partial x^2} = r i_m \quad (2-6)$$

As shown in Figure 2-4, a unit length of the membrane is modeled as a circuit which has a cell (modeling the resting potential) and a resistance connected in series, which are in turn connected in parallel to a capacitor. Let c_m be the membrane capacitance per unit length and r_m is the membrane resistance per unit length. We therefore have the Equation,

$$i_m = c_m \frac{\partial v_m}{\partial t} + \frac{v_m}{r_m} \quad (2-7)$$

Let $g_m = \frac{1}{r_m}$ represent the membrane conductance per unit length. The previous Equation thus becomes,

$$i_m = c_m \frac{\partial v_m}{\partial t} + g_m v_m \quad (2-8)$$

The Partial Differential Equations 2-6 and 2-8 are the starting point of Cable Theory which is used to describe passive conductance of postsynaptic potential across dendrites.

Solutions of these equations have been obtained for various boundary conditions, which we will not describe here. The reader is referred to (Dayan & Abbott, 2005) for more.

2.5 Active Conductance in the Axon - Hodgkin Huxley Equations

The axon has ion channels that allow for passive conductance in the same way as the dendrites. However, in addition, they also allow for active conductance, which is crucial in producing and propagating action potentials. Recall that when the membrane potential is elevated above a certain threshold, the neuron produces an action potential which has a stereotypical shape as shown in Figure 2-2(A). The membrane potential at which a spike is initiated varies between -55 mV and -45 mV, depending on the type of neuron in question. The action potential causes the membrane potential to depolarize to about +55 mV momentarily after which it quickly repolarizes to a value slightly below the resting potential; the latter phenomenon is called “afterhyperpolarization (AHP)”. After this, the membrane slowly gets back towards the resting potential. As alluded to earlier, unlike passive conduction, the action potential is transmitted down the axon without any loss in amplitude.

The mechanism of action potential generation involves a sudden change in the permeability of the membrane. Due to the concentration gradient, this causes a rapid change in the membrane potential. In what follows, we discuss the exact mechanism. The axon hillock and the axon contain voltage-gated ion channels. Recall that ion channels were simply proteins embedded in the cell membrane. Voltage-gated ion channels are activated only at certain voltage ranges. That is, at those voltage ranges, the proteins constituting the channel undergo a conformational change that makes them permeable to the corresponding ion. When the membrane potential approaches the threshold, voltage-gated Na^+ channels start getting activated. This causes an influx of Na^+ ions from outside the cell, which further activates more of the voltage-gated Na^+ channels. This process, called Na^+ current “activation”, momentarily raises the membrane potential to about +55 mV. At this point, the voltage-gated Na^+ channels undergo another

conformational change that makes them impermeable again. This is called Na⁺ current “inactivation”. At the same time, there is a further increase in the K⁺ permeability via voltage-gated K⁺ channels, which is called “delayed rectification”. Afterhyperpolarization is the consequence of Na⁺ current inactivation and delayed rectification. This is followed by “deactivation” of the voltage-gated K⁺ channels enabling the membrane potential to return to the resting potential. The voltage-gated Na⁺ channels are also “deinactivated” at this point, which is to say that they undergo another conformational change to bring them back to the conformation they were in before the current spike was initiated. This makes them ready to initiate the next spike. Recall that the absolute refractory period was a period of time (of the order of a few milliseconds) immediately after the initiation of a spike during which a subsequent spike cannot be initiated. The reason for the absolute refractory period is the deactivation of the Na⁺ channels. Also recall the relative refractory period, which was a time period of the order of tens of milliseconds after a spike during which more than usual depolarization was necessary to initiate a spike. This is because of the outward K⁺ current which is a consequence of delayed rectification.

In 1952, Alan Hodgkin and Andrew Huxley, in Nobel prize winning work, created a mathematical model for the above process that fit data from the squid giant axon. They started with a more general version of Equation 2-8,

$$i_m = c_m \frac{\partial v_m}{\partial t} + g_{Na}(v_m - V_{Na}) + g_K(v_m - V_K) + g_L(v_m - V_L) \quad (2-9)$$

where g .'s are the membrane conductances for the corresponding ions and V .'s are their respective equilibrium potentials.² They then determined the relationship between the various conductances and the potential using experimental data, leading to the following

² L stands for leak and represents effects of Cl⁻ and other ions put together.

set of equations (Hodgkin & Huxley, 1952b,c,d,a; Hodgkin & Katz, 1952).

$$i_m = c_m \frac{\partial v_m}{\partial t} + \bar{g}_{Na} m^3 h (v_m - V_{Na}) + \bar{g}_K n^4 (v_m - V_K) + \bar{g}_L (v_m - V_L), \quad (2-10)$$

$$\frac{dm}{dt} = \alpha_m (1 - m) - \beta_m m, \quad (2-11)$$

$$\frac{dh}{dt} = \alpha_h (1 - h) - \beta_h h, \quad (2-12)$$

$$\frac{dn}{dt} = \alpha_n (1 - n) - \beta_n n, \quad (2-13)$$

These are called the Hodgkin-Huxley Equations. $\alpha_m, \beta_m, \alpha_h, \beta_h, \alpha_n$ and β_n are functions of v_m as below:

$$\alpha_m = \frac{0.1(v_m + 25)}{e^{\frac{v_m + 25}{10}} - 1}, \quad \beta_m = 4e^{\frac{v_m}{18}}, \quad (2-14)$$

$$\alpha_h = 0.07e^{\frac{v_m}{20}}, \quad \beta_h = \frac{1}{e^{\frac{v_m + 30}{10}} + 1}, \quad (2-15)$$

$$\alpha_n = \frac{0.01(v_m + 10)}{e^{\frac{v_m + 10}{10}} - 1}, \quad \beta_n = 0.125e^{\frac{v_m}{80}}. \quad (2-16)$$

2.6 Other Varieties of Neuronal Behavior

The biophysics described thus far does not account for several classes of neuronal behavior that have been seen. Primarily this is because such behavior is mediated by a variety of ion channels that were not treated previously. In this section, we briefly describe some of these modes and their electrophysiological underpinnings.

“Spike frequency adaptation” is a slight modification of the basic behavior wherein on continuous depolarization over a period of time, the neuron’s output spike rate goes down or “adapts” to the sustained input. Cortical and hippocampal pyramidal neurons exhibit this behavior (Shepherd, 2004). Activation of a slow K^+ hyperpolarization current that adds up over time has been demonstrated to cause spike frequency adaptation.

Some neurons are inherently “bursty”. That is they tend to emit bursts of spikes rather than single spikes. Examples include Thalamic relay neurons, some pyramidal neurons, stellate cells and neurons in the inferior olive (Shepherd, 2004). These are caused

by specialized Ca^{2+} currents with slow kinetics, which keep the membrane depolarized for long enough, so that there can be a sequence of Na^+ - K^+ spikes.

Inhibitory interneurons in the cortex, thalamus and hippocampus fire short-duration spikes of less than a millisecond duration at frequencies greater than 300 Hz.

Then, there are “pacemaker neurons” that spontaneously generate low frequency spike trains (1-10Hz). These are found in the locus coeruleus, dorsal raphe, medial habenula nuclei and cerebellar Purkinje cells ([Shepherd, 2004](#)). These are caused by persistent Na^+ channels, where steady influx of Na^+ ions depolarizes the cell to above threshold, consequently triggering baseline activity.

CHAPTER 3 MODEL

In this chapter, we describe the abstract model used in the rest of the work. First we define some notation that will be used frequently. Next, we informally discuss the assumptions underlying our model. We then proceed to formally define the neuron as an abstract mathematical object and define other structures that we will find ourselves using.

3.1 Notation and Preliminaries

An “action potential” or “spike” is a stereotypical event characterized by the time instant at which it is initiated¹, which is referred to as its “spike time”. Spike times are represented relative to the present by real numbers, with positive values denoting past spike times and negative values denoting future spike times². A “spike train” $\vec{x} = \langle x^1, x^2, \dots, x^k, \dots \rangle$ is an increasing sequence of spike times, with every pair of spike times being more than α apart, where $\alpha > 0$ is (called) the absolute refractory period³ and x^i is the spike time of spike i . An “empty spike train”, denoted by $\vec{\phi}$, is one which has no spikes. Let \mathcal{S} denote the set of all spike trains. A “time-bounded spike train” (with “bound” $[a, b]$) is one where all spike times lie in the bounded interval $[a, b]$, for $a, b \in \mathbb{R}$. Note that, owing to the absolute refractory period, a time-bounded spike train is also a finite-length sequence. Let $\bar{\mathcal{S}}_{[a,b]}$ denote the set of all time-bounded spike trains with bound $[a, b]$. A spike train is said to have a “gap” in the interval $[c, d]$, if it has no spikes in that time interval. Furthermore, this gap is said to be of “length” $d - c$. Two spike

¹ The exact location of spike initiation is not important as long as the spike time is measured as the time at which the spike passes through a certain fixed point in the axon, upon initiation.

² Note that this convention implies that larger a spike’s spike time, the more ancient the spike is.

³ We assume a single fixed absolute refractory period for all neurons, for convenience, although our results would be no different if different neurons had different absolute refractory periods.

trains are said to be “identical” in the interval $[a, b]$, if they have exactly the same spike times in that time interval.

A “spike configuration” $\chi = \langle \vec{x}_1, \dots, \vec{x}_m \rangle$ is a tuple of spike trains. The “order” of a spike configuration is the number of spike trains in it. For example, χ above is a spike configuration of order m . A “time-bounded spike configuration” (with “bound” $[a, b]$) is one which has each of its spike trains to be time-bounded (with “bound” $[a, b]$). Two spike configurations are said to be “identical” in the interval $[a, b]$, if they have the same order and their corresponding spike trains are identical in the same time interval. A spike configuration χ is said have a “gap” in the interval $[c, d]$, if each of its spike trains has a gap in the interval $[c, d]$.

Next, we define some operators to time-shift spike trains/configurations, create bounded-time spike trains/configurations from spike trains/configurations and to assemble/disassemble spike trains into/from spike configurations. Let $\vec{x} = \langle x^1, x^2, \dots, x^k, \dots \rangle$ be a spike train and $\chi = \langle \vec{x}_1, \dots, \vec{x}_m \rangle$ be a spike configuration. We define some operations on these objects. The “shift operator for spike trains” is defined as $\sigma_t(\vec{x}) = \langle x^1 - t, x^2 - t, \dots, x^k - t, \dots \rangle$. The “shift operator for spike configurations” is defined as $\sigma_t(\chi) = \langle \sigma_t(\vec{x}_1), \dots, \sigma_t(\vec{x}_m) \rangle$. The “truncation operator for spike trains” is defined as follows: $\Xi_{[a,b]}(\vec{x})$ is the time-bounded spike train with bound $[a, b]$ that is identical to \vec{x} in the interval $[a, b]$. $\Xi_{(a,b]}(\vec{x})$ and $\Xi_{[a,b)}(\vec{x})$ are defined likewise. $\Xi_{[a,\infty)}(\vec{x})$ is the spike train that is identical to \vec{x} in the interval $[a, \infty)$ and has no spikes in the interval $(-\infty, a)$. Similarly, $\Xi_{(-\infty,b]}(\vec{x})$ is the spike train that is identical to \vec{x} in the interval $(-\infty, b]$ and has no spikes in the interval (b, ∞) . The “truncation operator for spike configurations” is defined as $\Xi_{[a,b]}(\chi) = \langle \Xi_{[a,b]}(\vec{x}_1), \dots, \Xi_{[a,b]}(\vec{x}_m) \rangle$; $\Xi_{(a,b]}(\chi)$ and $\Xi_{[a,b)}(\chi)$ are defined likewise. Similarly, $\Xi_{[a,\infty)}(\chi) = \langle \Xi_{[a,\infty)}(\vec{x}_1), \dots, \Xi_{[a,\infty)}(\vec{x}_m) \rangle$ and $\Xi_{(-\infty,b]}(\chi) = \langle \Xi_{(-\infty,b]}(\vec{x}_1), \dots, \Xi_{(-\infty,b]}(\vec{x}_m) \rangle$. Furthermore, $\Xi_t(\cdot)$ is shorthand for $\Xi_{[t,t]}(\cdot)$. The “projection operator for spike configurations” is defined as $\Pi_i(\chi) = \vec{x}_i$, for

$1 \leq i \leq m$. Let $\vec{y}_1, \vec{y}_2, \dots, \vec{y}_n$ be spike trains. The “join operator for spike trains” is defined as $\vec{y}_1 \sqcup \vec{y}_2 \sqcup \dots \sqcup \vec{y}_n = \langle \vec{y}_1, \vec{y}_2, \dots, \vec{y}_n \rangle$. For brevity, $\vec{y}_1 \sqcup \vec{y}_2 \sqcup \dots \sqcup \vec{y}_n$ is also written as $\bigsqcup_{i=1}^n \vec{y}_i$.

3.2 Assumptions Underlying the Model

Our neurons are abstract mathematical objects that are constrained to satisfy a small set of axioms. First, we informally describe the assumptions that underlie the axioms. Notable cases where the assumptions do not hold are also pointed out. This is followed by formal definitions. The approach taken here closely follows the one in (Banerjee, 2001a).

The following are our assumptions:

1. We assume that the neuron is a device that receives input from other neurons exclusively by spikes which are received via chemical synapses.⁴
2. The neuron is a finite-precision device with fading memory. Hence, the underlying potential function can be determined from a bounded past. That is, we assume that, for each neuron, there exist real numbers Υ and ρ , so that the current membrane potential of the neuron can be determined as a function of the input spikes received in the past Υ seconds and the spikes produced by the neuron in the past ρ seconds.⁵
3. We assume that the membrane potential of the neuron can be written down as a real-valued, everywhere bounded function of the form $P(\chi; \vec{x}_0)$, where \vec{x}_0 is a time-bounded spike train, with bound $[0, \rho]$ and $\chi = \langle \vec{x}_1, \dots, \vec{x}_m \rangle$ is a time-bounded spike configuration with bound $[0, \Upsilon]$. Informally, \vec{x}_i , $1 \leq i \leq m$, is the sequence of spikes afferent in synapse i in the past Υ seconds and \vec{x}_0 is the sequence of spikes efferent from the current neuron in the past ρ seconds. The function $P(\cdot)$ characterizes the entire spatiotemporal response of the neuron to spikes including

⁴ In this work, we do not treat electrical synapses or ephaptic interactions (Shepherd, 2004).

⁵ ρ corresponds to the notion of “relative refractory period.”

synaptic strengths, their location on dendrites, and their modulation of each other's effects at the soma, spike-propagation delays, and the postspike hyperpolarization.

4. Without loss of generality⁶, we assume the resting membrane potential⁷ to be 0. Let $\tau > 0$ be the threshold membrane potential⁸.
5. The neuron outputs a spike whenever $P(\cdot) = \tau$. Additionally, when a new output spike is produced, we assume that the first derivative is sufficiently high, so as to keep the membrane potential from going above the threshold.⁹ Therefore the membrane potential function in the abstract model takes values that are at most that of the threshold.
6. Owing to the absolute refractory period $\alpha > 0$, no two input or output spikes can occur closer than α . That is, suppose $\vec{x}_0 = \langle x_0^1, x_0^2, \dots, x_0^k \rangle$, where $x_0^1 < \alpha$. Then $P(\chi; \vec{x}_0) < \tau$, for all “legal” χ .
7. Past output spikes have an inhibitory effect, in the following sense¹⁰:

$$P(\chi; \vec{x}_0) \leq P(\chi; \vec{\phi}),$$
for all “legal” χ and \vec{x}_0 .
8. Finally, on receiving no input spikes in the past Υ seconds and no output spikes in the past ρ seconds, the neuron settles to its resting potential. That is,

$$P(\langle \vec{\phi}, \vec{\phi}, \dots, \vec{\phi} \rangle; \vec{\phi}) = 0.$$

⁶ There is no loss of generality since the membrane potential can be rescaled to fit this constraint, without affecting the spike times.

⁷ say measured at the soma.

⁸ Observe that the model allows for variable thresholds, as long as the threshold itself is a function of spikes afferent in the past Υ seconds and spikes efferent from the present neuron in the past ρ seconds.

⁹ Note that this assumption is made without loss of generality and helps ease the formulation of the abstract model.

¹⁰ This is notably violated in case of neurons that have a post-inhibitory rebound.

An “acyclic network of neurons”, informally, is a Directed Acyclic Graph where each vertex corresponds to an instantiation of the neuron model, with some vertices designated input vertices (which are placeholders for input spike trains), and one neuron designated the output neuron. The “depth” of an acyclic network is the length of the longest path from an input vertex to the output vertex.

3.3 Formal Definition of the Model

Next, we present a formal definition of a neuron as an abstract mathematical object. Some more definitions follow.

Definition 1 (Neuron). A “neuron” \mathbf{N} is a 7-tuple $\langle \alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau] \rangle$, where $\alpha, \Upsilon, \rho, \tau \in \mathbb{R}^+$ with $\rho \geq \alpha$, $\lambda \in \mathbb{R}^-$ and $m \in \mathbb{Z}^+$. Furthermore,

1. If $\vec{x}_0 = \langle x_0^1, x_0^2, \dots, x_0^k \rangle$ with $x_0^1 < \alpha$, then $P(\chi; \vec{x}_0) < \tau$, for all $\chi \in \bar{\mathcal{S}}_{[0, \Upsilon]}^m$ and for all $\vec{x}_0 \in \bar{\mathcal{S}}_{[0, \rho]}$.
2. $P(\chi; \vec{x}_0) \leq P(\chi; \vec{\phi})$, for all $\chi \in \bar{\mathcal{S}}_{[0, \Upsilon]}^m$ and for all $\vec{x}_0 \in \bar{\mathcal{S}}_{[0, \rho]}$.
3. $P(\langle \vec{\phi}, \vec{\phi}, \dots, \vec{\phi} \rangle; \vec{\phi}) = 0$.

A neuron is said to “generate a spike” whenever $P(\cdot) = \tau$.

Definition 2 (Network Architecture). A Network Architecture \mathcal{A} is a 5-tuple $\langle G(V, E), I, o, \# : \{1, 2, \dots, |I|\} \rightarrow I, \{\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V \mid v \in (V - I)\} \rangle$ where $G(V, E)$ is a directed graph¹¹ with input vertices $I \subset V$, an¹² output vertex $o \in V$, a bijection $\# : \{1, 2, \dots, |I|\} \rightarrow I$ that numbers the input vertices and a bijection $\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V$, for each vertex $v \in (V - I)$, that numbers the vertices incident on the in-edges of v . Furthermore, for each $v \in I$, we have $\text{indegree}(v) = 0$ and $\text{outdegree}(o) = 0$.

¹¹ A directed edge represents the fact that there is “atleast” one synapse from the neuron represented by the outgoing vertex to the neuron represented by the incoming vertex. The $P(\cdot)$ function of the efferent neuron gets only one spike train from the afferent neuron and encodes the response of all the synapses from the afferent neuron.

¹² In this work, we only study networks with a single output neuron.

Definition 3 (Acyclic Network Architecture). A Network Architecture $\mathcal{A}\langle G\langle V, E \rangle, I, o, \# : \{1, 2, \dots, |I|\} \rightarrow I, \{\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V \mid v \in (V - I)\} \rangle$ is called an “Acyclic Network Architecture” if $G\langle V, E \rangle$ is an acyclic graph.

Definition 4 (Network). A “Network” \mathcal{N} is a 3-tuple $\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$ where $\mathcal{A}\langle G\langle V, E \rangle, I, o, \# : \{1, 2, \dots, |I|\} \rightarrow I, \{\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V \mid v \in (V - I)\} \rangle$ is a network architecture, N is a set of neurons and $L : (V - I) \rightarrow N$ is a bijection that maps non-input vertices of G to neurons in N such that the following is true: For every $v \in (V - I)$, that has $L(v) = \mathbf{N}\langle \alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau] \rangle$, we have $\text{indegree}(v) = m$. Further, $L(o)$ is called the “output neuron”.

Definition 5 (Acyclic Network). An “Acyclic Network” $\mathcal{N}\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$ is one whose network architecture \mathcal{A} is acyclic.

Definition 6. In an acyclic network $\mathcal{N}\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$, the “depth” of a neuron $\mathbf{N} \in N$ is the length of the longest path from an input vertex to $L^{-1}(\mathbf{N})$.

Definition 7. The “depth” of an acyclic network $\mathcal{N}\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$ is the depth of its output neuron.

CHAPTER 4 ACYCLIC NETWORKS AS SPIKE-TRAIN TO SPIKE-TRAIN TRANSFORMATIONS

As stated in the introduction, we wish to look at acyclic networks of neurons as transformations that map input spike trains to output spike trains. Therefore, we first need to define in what sense, if at all, these networks constitute the said transformations.

We begin by demonstrating that notions of input/output transformations are not always well-defined, even for a single neuron. We then show, notwithstanding this dismaying prognosis, that such transformations are indeed well-defined under spiking regimes normally encountered in the brain. In particular, we derive precise sufficient conditions under which these transformations are well-defined which correspond to mapping finite-length input spike trains to finite-length output spike trains. We then observe that while this sufficient condition is biologically well-founded, it is mathematically unwieldy. We therefore proceed to establish more mathematically tractable sufficient conditions, albeit seemingly at the cost of some biological realism. However, in the next chapter, we demonstrate that in fact no biological well-foundedness is lost. We show this by establishing an equivalence theorem between these two conditions, insofar as proving results involving complexity¹ is concerned.

4.1 On the Consistency of Defining Input/Output Transformations on Neurons

Let us first consider the simplest acyclic network, namely the single neuron. Given that our abstract neuron model does not explicitly prescribe an output spike train for a given input spike train, we need to ask what it means for a neuron to “produce” an output spike train, when supplied with an input spike train. Recall, from the previous section, that the membrane potential of the neuron depends not only on the input spikes received in the past Υ seconds, it also depends on the output spikes produced by it in

¹ whose definition we also introduce in the next chapter.

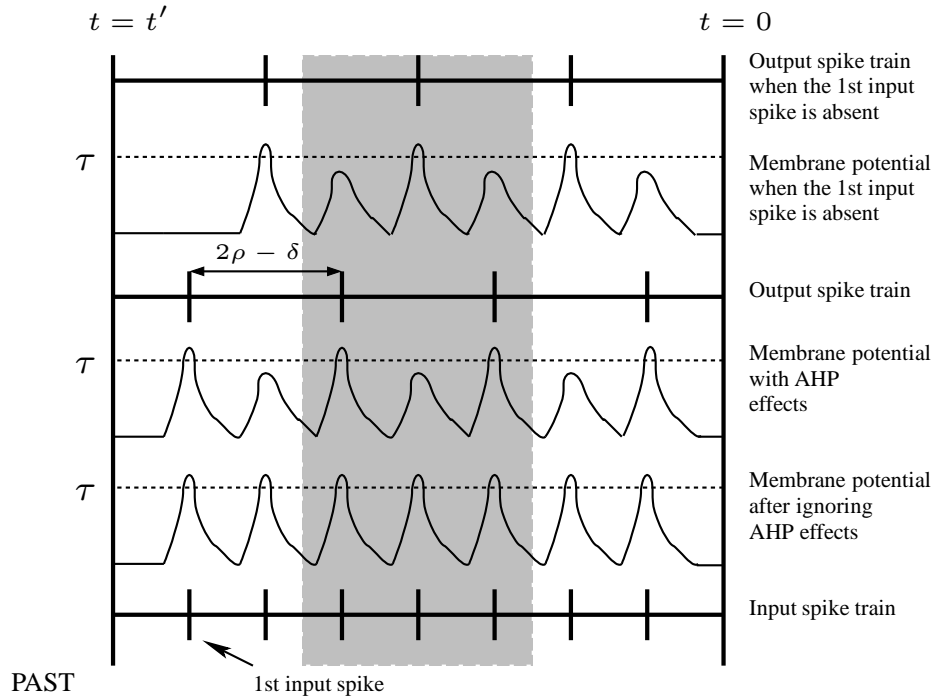


Figure 4-1. This example describes a single neuron which has just one afferent synapse. Until time t' in the past, it received no input. After this time, its input was spikes that arrived every $\rho - \delta/2$ seconds, where $\delta > 0$. An input spike alone (if there were no output spikes in the past ρ seconds) can cause this neuron to produce an output spike. However, if there were an output spike within the past ρ seconds, the AHP due to that spike is sufficient to bring the potential below threshold, so that the neuron does not spike currently. We therefore observe that if the first spike is absent, then the output spike train drastically changes. Note that this change occurs no matter how often the shaded segment in the middle is repeated, i.e. it does not depend on how long ago the first spike occurred.

the past ρ seconds. Therefore, knowledge of just input spikes in the past Υ seconds does not uniquely determine the current membrane potential (and therefore the output spike train produced from it). It might be tempting to then somehow use the fact that the past output spikes are themselves a function of input and output received in the more distant past, and attempt to make the membrane potential a function of a bounded albeit larger “window” of input spikes alone. The example in Figure 4-1 shows that this does not work. In particular, the current membrane potential of the neuron may depend on the position of the input spike that has occurred arbitrarily long time ago in the past (if we

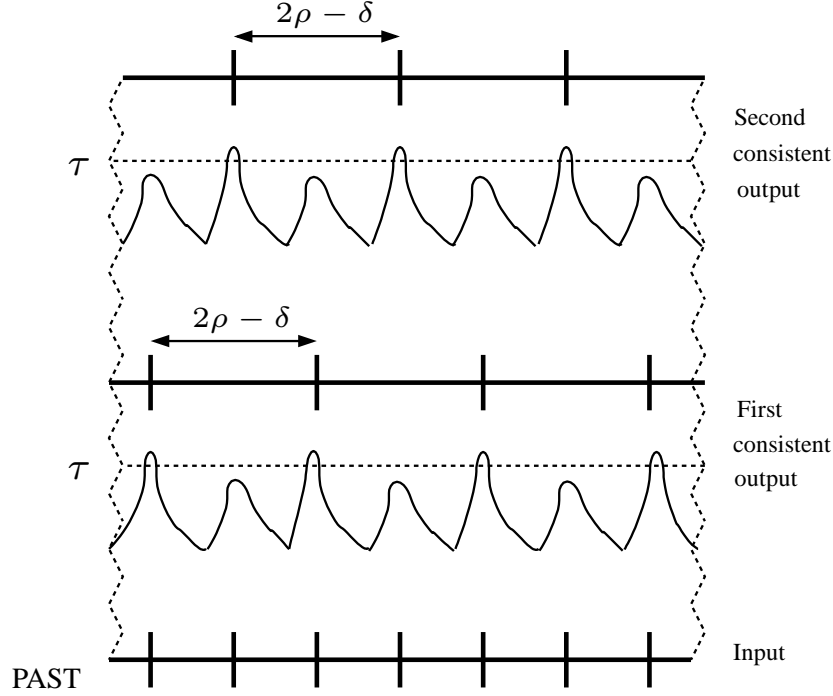


Figure 4-2. The example here is very similar to the one in Figure 4-1, except that, instead of there being no input spikes before t' , we have an unbounded input spike configuration, with the same periodic input spikes occurring since the infinite past. Observe that both output spike trains are consistent with this input, for each $t \in \mathbb{R}$.

attempt to characterize the membrane potential as a function of input spikes alone). One is then forced to ask if given the infinite history of input spikes received by the neuron, the membrane potential is then uniquely determined. Before we can answer this question, we need to rigorously define when we can consistently associate an (unbounded) output spike train with an (unbounded) input spike configuration, for a single neuron.

Definition 8. An output spike train \vec{x}_o is said to be “consistent” with an input spike configuration χ , with respect to a neuron $\mathbf{N}\langle\alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau]\rangle$, if the following holds. For every $t \in \mathbb{R}$, $t \in \vec{x}_o$ if and only if $P(\bar{\Xi}_{[0, \Upsilon]}(\sigma_t(\chi)), \bar{\Xi}_{[0, \rho]}(\sigma_t(\vec{x}_o))) = \tau$.

The question we posed now becomes equivalent to the following : For every (unbounded) input spike configuration χ , does there exist exactly one (unbounded)

output spike train \vec{x}_o , so that \vec{x}_o is consistent with χ for a given neuron N ? Interestingly, we find that the answer is still in the negative. The example in Figure 4-2 describes a neuron and an infinitely long input spike train, which has two consistent output spike trains.

It could be argued that the input spike train cannot possibly be infinitely long, since every neuron begins existence at a certain point in time. However, this begs the question of if the neuron was at the resting potential when the first input spikes arrived². An assumption to this effect would be significant, particularly if the current membrane potential depended on it. It is easy to construct an example along the lines of the example described in Figure 4-1, where the current membrane potential is different depending on whether this assumption is made or not. Assuming infinitely long input spike configurations, on the other hand, obviates the need to make any such assumption. We retain this viewpoint for the rest of the dissertation with the understanding that the alternative viewpoint discussed at the beginning of this paragraph can also be expounded along similar lines.

Nevertheless, the underlying difficulty in defining even single neurons as spike train to spike train transformations, with both viewpoints discussed above is dependence, in general, of current membrane potential to “initial state”. However, this still leaves open the possibility of considering just a subset of input/output spike trains, which have the property of the current membrane potential being independent of the input spike train beyond a certain time in the past. Such a subset would exclude the examples discussed in this section. The caveat, of course, is that even if such a subset exists, unless it is also biologically well-motivated, claims of subsequent results having biological relevance are questionable.

² Note that our axiomatic definition of a neuron does not address this question.

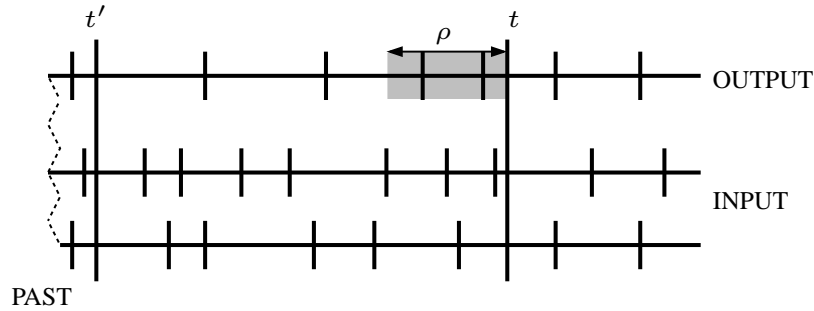


Figure 4-3. The idea behind the Gap Lemma.

In what follows, we come up with a biologically-realistic condition that implies independence as alluded to above; roughly speaking, the condition is that if a neuron has had a recent gap in its output spike train equal to at least twice its relative refractory period, then its current membrane potential is independent of the input beyond the relatively recent past. We show that this leads to the notion of spike-train to spike-train transformations to be well defined for acyclic networks.

4.2 Gap Criteria

In this section, we come up with a biologically well-motivated condition that guarantees independence of current membrane potential from input spikes beyond the recent past. This condition is used in constructing a criterion for single neurons which when satisfied, guarantees a unique consistent (unbounded) output spike train. Next, similar criteria are defined for acyclic networks.

For a neuron, the way input spikes that happened sufficiently earlier affect current membrane potential is via a causal sequence of output spikes, causal in the sense that each of them had an effect on the membrane potential while the subsequent one in the sequence was being produced. The condition in the Gap Lemma basically seeks to break this causal chain.

4.2.1 Gap Lemma

To see the main idea that leads to the condition, see Figure 4-3. Suppose, the spikes in the shaded region (which is an interval of length ρ) occurred at the exact same position

$P(\Xi_{[0,\Upsilon]}(\sigma_{t'}(\chi)), \Xi_{[0,\rho]}(\sigma_{t'}(\vec{x}_0)))$ is at most the value of $P(\Xi_{[0,\Upsilon]}(\sigma_{t'}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t'}(\vec{x}_0^*)))$, because $\Xi_{[0,\rho]}(\sigma_{t'}(\vec{x}_0^*))$ is $\vec{\phi}$. Since $P(\Xi_{[0,\Upsilon]}(\sigma_{t'}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t'}(\vec{x}_0^*)))$ is less than τ for every $t' \in [t, t + \rho]$, $P(\Xi_{[0,\Upsilon]}(\sigma_{t'}(\chi)), \Xi_{[0,\rho]}(\sigma_{t'}(\vec{x}_0)))$ is less than τ in the same interval, as well. Therefore, \vec{x}_0 has no output spikes in $[t, t + \rho]$.

That 2ρ is the smallest possible gap length in \vec{x}_0^* for this to hold, follows from the example in Figure 4-1, where the conclusion did not hold, when \vec{x}_0^* had gaps of length $2\rho - \delta$, for arbitrary $\delta > 0$. □

Corollary 1. *Consider a neuron $\mathbf{N}(\alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0,\Upsilon]}^m \times \bar{\mathcal{S}}_{[0,\rho]} \rightarrow [\lambda, \tau])$, a spike configuration χ^* of order m and a spike train \vec{x}_0^* which has a gap in the interval $[t, t + 2\rho]$ so that \vec{x}_0^* is consistent with χ^* , with respect to \mathbf{N} . Then*

1. *Every \vec{x}_0 consistent with χ^* , with respect to \mathbf{N} , has a gap in the interval $[t, t + \rho]$.*
2. *Every \vec{x}_0 consistent with χ^* , with respect to \mathbf{N} , is identical to \vec{x}_0^* in the interval $(-\infty, t + \rho]$.*
3. *For every t' more recent than $(t + \rho)$, the membrane potential at t' , is precisely a function of spikes in $\Xi_{[t', t + \Upsilon + \rho]}(\chi^*)$.*

Proof. (1) is immediate from the lemma, when we set $\chi = \chi^*$.

For (2), the proof is by strong induction on the number of spikes since t . Let \vec{x}_0 be an arbitrary spike train that is consistent with χ^* , with respect to \mathbf{N} . Notice that from (1) we have that every \vec{x}_0 is identical to \vec{x}_0^* in $[t, t + \rho]$. The base case is to show that both \vec{x}_0^* and \vec{x}_0 have their first spike after t at the same time. Assume, without loss of generality, that the first spike of \vec{x}_0 at $t_1 < t$, is no later than the first spike of \vec{x}_0^* . We have $P(\Xi_{[0,\Upsilon]}(\sigma_{t_1}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t_1}(\vec{x}_0^*))) = P(\Xi_{[0,\Upsilon]}(\sigma_{t_1}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t_1}(\vec{x}_0)))$ since $\Xi_{[0,\rho]}(\sigma_{t_1}(\vec{x}_0^*)) = \Xi_{[0,\rho]}(\sigma_{t_1}(\vec{x}_0)) = \vec{\phi}$. Therefore \vec{x}_0^* also has its first spike at t_1 . Let the induction hypothesis be that both \vec{x}_0^* and \vec{x}_0 have their first k spikes since t at the same times. We show that this implies that the $(k + 1)^{th}$ spike in each spike train is also at the same time instant. Assume, without loss of generality, that the $(k + 1)^{th}$ spike since t of \vec{x}_0 at t_{k+1} , is no later than the $(k + 1)^{th}$ spike since t of \vec{x}_0^* . Now, $\Xi_{[0,\rho]}(\sigma_{t_{k+1}}(\vec{x}_0^*))$ is

identical to $\Xi_{[0,\rho]}(\sigma_{t_{k+1}}(\vec{x}_0))$ from the induction hypothesis since $(t + \rho) - t_{k+1} \geq \rho$. Thus, $P(\Xi_{[0,\Upsilon]}(\sigma_{t_{k+1}}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t_{k+1}}(\vec{x}_0^*))) = P(\Xi_{[0,\Upsilon]}(\sigma_{t_{k+1}}(\chi^*)), \Xi_{[0,\rho]}(\sigma_{t_{k+1}}(\vec{x}_0)))$ and therefore \vec{x}_0^* also has its $(k + 1)^{th}$ spike at t_{k+1} . This completes the proof of (2). (3) follows from the Gap Lemma and (2). \square

The upshot of the Gap Lemma and its corollary is that whenever a neuron goes through a period of time equal to twice its relative refractory period where it has produced no output spikes, its membrane potential from then on becomes independent of input spikes that are older than $\Upsilon + \rho$ seconds before the end of the gap.

Large gaps in the output spike trains of neurons seem to be extensively prevalent in the human brain. In parts of the brain where the neurons spike persistently, such as in the frontal cortex, the spike rate is very low (0.1Hz-10Hz) (Shepherd, 2004). In contrast, the typical spike rate of retinal ganglion cells can be very high but the activity is generally interspersed with large gaps during which no spikes are emitted (Nirenberg et al., 2001).

4.2.2 Gap Criterion for a Neuron

These observations motivate our definition of a criterion for input spike configurations efferent on single neurons. The criterion dictates that there be intermittent gaps of length atleast twice the relative refractory period in an output spike train consistent with the spike configuration.

Definition 9 (Gap Criterion for a single neuron). For $T \in \mathbb{R}^+$, a spike configuration χ is said to satisfy a “ T -Gap Criterion” for a neuron $\mathbf{N}\langle\alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0,\Upsilon]}^m \times \bar{\mathcal{S}}_{[0,\rho]} \rightarrow [\lambda, \tau]\rangle$ if the following is true: χ is of order m and there exists a spike train \vec{x}_0 with at least one gap of length 2ρ in every interval of time of length $T - \Upsilon + 2\rho$, so that \vec{x}_0 is consistent with χ with respect to \mathbf{N} .

Such input spike configurations also have exactly one consistent output spike train. We prove this in the next proposition.

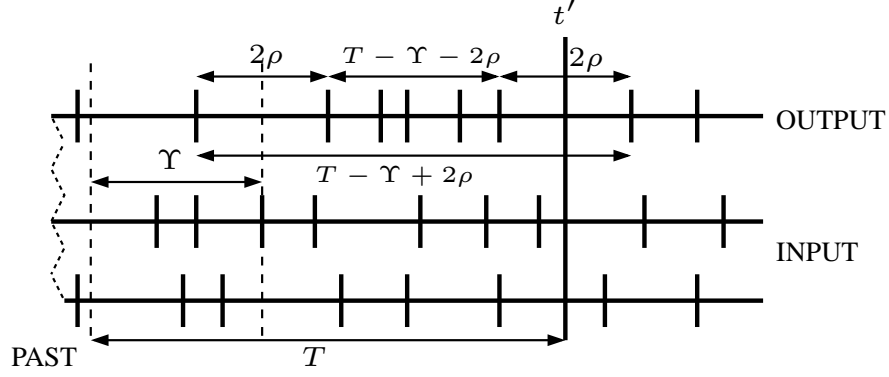


Figure 4-5. Illustration demonstrating that for an input spike configuration χ that satisfies a T -Gap criterion, the membrane potential at any point in time is dependent on at most T seconds of input spikes in χ before it.

Proposition 1. Let χ be a spike configuration that satisfies a T -Gap criterion for a neuron $\mathbf{N}\langle\alpha, \Upsilon, \rho, \tau, \lambda, m, P : \bar{\mathcal{S}}_{[0, \Upsilon]}^m \times \bar{\mathcal{S}}_{[0, \rho]} \rightarrow [\lambda, \tau]\rangle$. Then, there is exactly one spike train \vec{x}_0 , such that \vec{x}_0 is consistent with χ , with respect to \mathbf{N} .

Proof. Since χ satisfies a T -Gap criterion, there exists a spike train \vec{x}_0 with at least one gap of length 2ρ in every interval of time of length $T - \Upsilon + 2\rho$, so that \vec{x}_0 is consistent with χ with respect to \mathbf{N} . For the sake of contradiction, assume that there exists another spike train \vec{x}'_0 , not identical to \vec{x}_0 , which is consistent with χ , with respect to \mathbf{N} . Let t' be the time at which one spike train has a spike but another doesn't. Let $t > t'$ be such that \vec{x}_0 has a gap in the interval $[t, t + \rho]$. By Corollary 1 to the Gap Lemma, it follows that \vec{x}'_0 is identical to \vec{x}_0 starting from time instant $t + \rho$. This contradicts the hypothesis that \vec{x}'_0 is different from \vec{x}_0 at t' . \square

For an input spike configuration χ that satisfies a T -Gap criterion, the membrane potential at any point in time is dependent on at most T seconds of input spikes in χ before it. This can be seen from Figure 4-5, which illustrates a section of the input spike configuration and the output spike train. Because of the T -Gap criterion the distance between any two gaps of length 2ρ on the output spike train is at most $T - \Upsilon - 2\rho$. Upto the earlier half of a 2ρ gap (whose latest point is denoted by t') is dependent on input

corresponding to the previous 2ρ gap. The membrane potential at t' depends on input spikes in the interval of length T , as depicted.

With inputs that satisfy the T -Gap Criterion, here is what we need to do to physically determine the current membrane potential, even if the neuron has been receiving input since the infinite past: Start off the neuron from an arbitrary state, and drive it with input that the neuron received in the past T seconds. The Gap Lemma guarantees that the membrane potential we see now will be identical to the actual membrane potential.

The Gap Criterion we have defined for single neurons can be naturally extended to acyclic networks. The criterion is simply that the input spike configuration to the network is such that every neuron's input obeys a Gap criterion for single neurons.

Definition 10 (Gap Criterion for an acyclic network). Consider an acyclic network $\mathcal{N}\langle\mathcal{A}, N, L : (V - I) \rightarrow N\rangle$. Let d be the depth of \mathcal{N} . Let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i . For $T \in \mathbb{R}^+$, a spike configuration χ is said to satisfy a “ T -Gap Criterion” for the acyclic network \mathcal{N} if χ is of order $|I|$ and the following are true:

1. For each $\mathbf{N} \in N_1$, $\bigsqcup_{i=1}^{\text{indegree}(v)} \Pi_{\#^{-1}(\#_v(i))}(\chi)$ satisfies a $(\frac{T}{d})$ -Gap Criterion for \mathbf{N} , where $v = L^{-1}(\mathbf{N})$.
2. For $i = 2, \dots, d$ respectively, for each $\mathbf{N} \in N_i$, the input to \mathbf{N} satisfies a $(\frac{T}{d})$ -Gap Criterion for \mathbf{N} .

As with the criterion for the single neuron, the membrane potential of the output neuron at any point is dependent on atmost T seconds of past input, if the input spike configuration to the acyclic network satisfies a T -Gap criterion. The situation is illustrated in Figure 4-6. Additionally, the output spike train is unique.

Lemma 2. *Consider an acyclic network $\mathcal{N}\langle\mathcal{A}, N, L : (V - I) \rightarrow N\rangle$. Let χ satisfy a T -Gap criterion for \mathcal{N} . Then \mathcal{N} produces a unique output spike train when it receives χ as input. Furthermore, the membrane potential of the output neuron at any time instant depends on atmost the past T seconds of input in χ .*

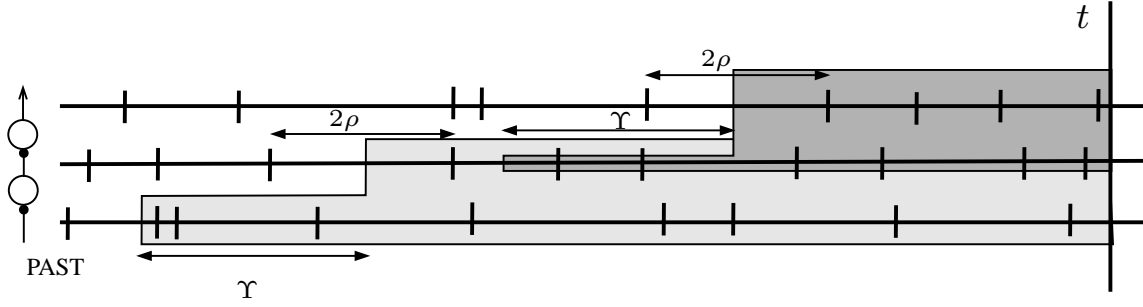


Figure 4-6. Schematic diagram illustrating how the Gap criterion works for the simple two-neuron network on the left. The membrane potential of the output neuron at t depends on input received from the “intermediate” neuron, as depicted in the darkly-shaded region, owing to the Gap Lemma. The output of the intermediate neuron in the darkly-shaded region, in turn, depends on input it received in the lightly-shaded region. Thus, transitively, membrane potential of the output neuron at t is dependent at most on input received by the network in the lightly-shaded region.

Proof. We prove that the output of the network is unique by strong induction on depth. As before, let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i . Each neuron $N \in N_1$ receives all inputs from spike trains in χ . Since, N satisfies a Gap criterion with those input spike trains, its output is unique. The induction hypothesis then is that for all $i \leq k < d$, each neuron $N \in N_i$ produces a unique output when \mathcal{N} is driven by χ . Consider arbitrary $N' \in N_{k+1}$. It is clear that all inputs to N' are from spike trains from χ or neurons in $\bigcup_{i=1}^k N_i$, for otherwise the depth of N' would be greater than $k + 1$. Since, all its inputs are unique by the induction hypothesis and they satisfy a Gap criterion for N' , its output is also unique.

Next, we show that the membrane potential of the output neuron at any time instant depends on at most the past T seconds of input in χ . Since the output neuron satisfies a $(\frac{T}{d})$ -Gap Criterion, its membrane potential at any point depends on at most the past $(\frac{T}{d})$ seconds of the inputs it receives (some of which are output spike trains of other neurons). Consider one such “penultimate layer” neuron. Again, its output membrane potential at any time instant, likewise, depends on its inputs in the past $(\frac{T}{d})$ seconds. Therefore, the current potential of the output neuron is dependent on the input received by the

penultimate layer neuron in the past ($\frac{2T}{d}$) seconds. Similar arguments can be put forth until, for each path, one reaches a neuron, all of whose inputs do not come from other neurons. Since the longest such path is of length d , it is straightforward to verify that the membrane potential of the output neuron depends on at most T seconds of past input. \square

We are thus at a juncture where questions we initially asked can even be posed in a coherent manner that is also biologically well-motivated. Before we proceed, we introduce some more notation.

Given an acyclic network \mathcal{N} , let $\mathcal{G}_{\mathcal{N}}^T$ be the set of all input spike configurations that satisfy a T -Gap Criterion for \mathcal{N} . Let $\mathcal{G}_{\mathcal{N}} = \bigcup_{T \in \mathbb{R}^+} \mathcal{G}_{\mathcal{N}}^T$. Therefore, every acyclic network \mathcal{N} induces a transformation $\mathcal{T}_{\mathcal{N}} : \mathcal{G}_{\mathcal{N}} \rightarrow \mathcal{S}$ that maps each element of $\mathcal{G}_{\mathcal{N}}$ to a unique output spike train in \mathcal{S} . Suppose $\mathcal{G}' \subseteq \mathcal{G}_{\mathcal{N}}$. Then, let $\mathcal{T}_{\mathcal{N}}|_{\mathcal{G}'} : \mathcal{G}' \rightarrow \mathcal{S}$ be the map defined as $\mathcal{T}_{\mathcal{N}}|_{\mathcal{G}'}(\chi) = \mathcal{T}_{\mathcal{N}}(\chi)$, for all $\chi \in \mathcal{G}'$.

4.2.3 Practical Issues with the Gap Criteria

The Gap Criteria are very general and biologically realistic. However, given a neuron or an acyclic network, there does not appear to be an easy way to characterize all the input spike configurations that satisfy a certain Gap Criterion for it. Whether a spike configuration satisfies a certain gap criterion seems to be intimately dependent on the particular form of the potential function of each neuron. For an acyclic network, this is even more complex, since intermediate neurons must satisfy Gap Criteria, with the inputs they get being outputs of other neurons. Each neuron or acyclic network could potentially induce a different set of spike configurations that satisfy a T-Gap Criterion for it. This appears to make the problem of comparing the transformations performed by two different neurons/acyclic networks difficult, because of the difficulty in finding spike configurations that satisfy Gap Criteria for both of them.

This brings up the question of the existence of another criterion according to which the set of spike configurations is easier to characterize and is common across different networks. Next, we propose one such criterion and we show that it induces spike

configurations which are a subset of those induced by the Gap criteria for all acyclic networks. Briefly, these are input spike configurations which, before a certain time instant in the past, have had no spikes. We emphasize that this is a purely theoretical construct made for mathematical expedience; that is, no biological relevance is claimed. The spike configurations satisfying the said criterion, which we call the Flush criterion, allow us to sidestep the difficult issues mentioned in the previous paragraph. Significantly, in a subsequent section, after having defined notions of complexity, we show that there is no loss by restricting ourselves to the Flush criterion. That is, not only is a complexity result proved using the Flush criterion applicable with the Gap criterion, every complexity result true for Gap criterion can be proved by using the Flush criterion exclusively.

4.3 Flush Criterion

The idea of the Flush Criterion is to force the neuron to produce no output spikes for sufficiently long so as to guarantee that a Gap criterion is being satisfied. This is done by having a sufficiently long interval with no input spikes. This “flushes” the neuron by bringing it to the resting potential and keeps it there for an appropriately long time. In an acyclic network, the flush is propagated so that all neurons have had a sufficiently long gap in their output spike trains. Note that the Flush Criterion is not defined with reference to any acyclic network. We formalize this notion below.

Definition 11 (Flush Criterion). A spike configuration χ is said to satisfy a T -Flush Criterion, if all its spikes lie in the interval $[0, T]$, i.e. it has no spikes before time instant T and after time instant 0.

First, we show that an input spike configuration to a neuron that satisfies a Flush criterion also satisfies a Gap criterion. Figure 4-7 accompanies the following lemma.

Lemma 3. *An input spike configuration χ for a neuron that satisfies a T -Flush Criterion also satisfies a $(T + 2\Upsilon + 2\rho)$ -Gap Criterion for that neuron.*

Proof. The neuron on being driven by χ cannot have output spikes outside the interval $[-\Upsilon, T]$. This easily follows from Axiom 2 and 3 of the neuron because the neuron does

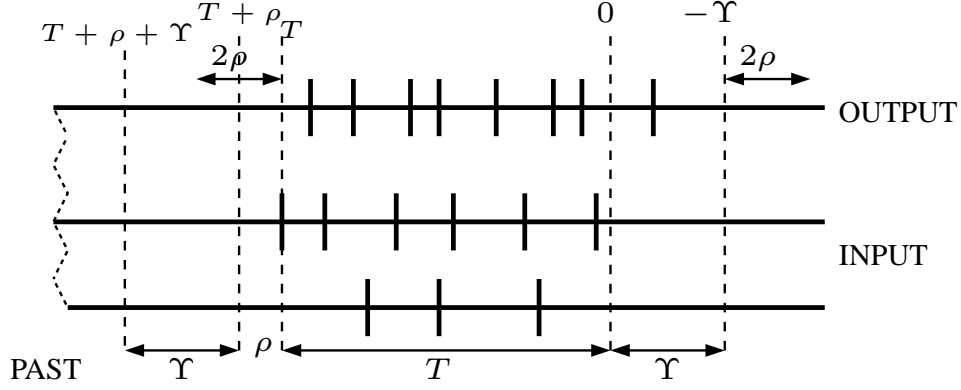


Figure 4-7. Illustration showing that an input spike configuration satisfying a Flush Criterion also satisfies a Gap Criterion.

not have input spikes before time instant T and in the interval $[-\Upsilon, 0]$ and onwards. Now to see that χ satisfies a $(T + 2\Upsilon + 2\rho)$ -Gap Criterion, recall that with a T' -Gap Criterion, distance between any two gaps of length 2ρ on the output spike train is at most $T' - \Upsilon - 2\rho$. With χ , we observe that the distance between any two 2ρ gaps on the output spike train is at most $T + \Upsilon$. Thus, $T' - \Upsilon - 2\rho = T + \Upsilon$, which gives us $T' = T + 2\Upsilon + 2\rho$. The result follows. \square

Next, it is shown that an input spike configuration to an acyclic network satisfying a Flush criterion also satisfies a Gap criterion for that network.

Lemma 4. *An input spike configuration χ for an acyclic network that satisfies a T -Flush Criterion also satisfies a $(dT + d(d + 1)\Upsilon + 2d\rho)$ -Gap Criterion for that network, where Υ , ρ are upper bounds on the same parameters taken over all the neurons in the network and d is the depth of the network.*

Proof. Following the proof of the previous lemma, we know that neurons that receive all their inputs from χ have no output spikes outside the interval $[-\Upsilon, T]$. Similarly, neurons that have depth 2 with respect to the input vertices of the network have no output spikes outside $[-2\Upsilon, T]$. Likewise, the output neuron, which has depth d , has no output spikes outside $[-d\Upsilon, T]$. It follows that the output neuron obeys a $(T + (d + 1)\Upsilon + 2\rho)$ -Gap Criterion. Also, every other neuron follows this criterion because the distance between the

2ρ output gaps for every neuron is at most that of the output neuron, since their depth is bounded from above by the depth of the output neuron. Thus, from the definition of the Gap criterion for acyclic networks, we have that χ satisfies a $(dT + d(d + 1)\Upsilon + 2d\rho)$ -Gap Criterion for the current network. \square

We introduce some more notation. Let the set of spike configurations containing exactly m spike trains that satisfy the T-Flush criterion be \mathcal{F}_m^T . Let $\mathcal{F}_m = \bigcup_{T \in \mathbb{R}^+} \mathcal{F}_m^T$. What we have established in this section is that $\mathcal{F}_m \subseteq \mathcal{G}_{\mathcal{N}}$, where \mathcal{N} has exactly m input vertices.

CHAPTER 5
TRANSFORMATIONAL COMPLEXITY: DEFINITIONS AND RESULTS

5.1 Motivation and Definition

In this section, we define notions of relative complexity of sets of acyclic networks of neurons, with respect to transformations effected by them. For brevity, we refer to these notions as “Transformational Complexity”. What we would like to capture with the definition is the following: Given two classes of networks with the second class subsuming the first, we wish to ask if there are transformations in the second class that cannot be performed by networks in the first class. That is, do the extra networks in the second class make it richer in terms of transformational power? The classes could correspond to network architectures, although for the purpose of the definition, there is no reason to require this to be the case. While comparing a set of networks, we restrict ourselves to inputs for which all the networks satisfy a certain Gap Criterion (though, not necessarily for the same T), so that the notion of a transformation is well-defined on the input set, for all networks under consideration.

Definition 12. Let Σ_1 and Σ_2 be two sets of acyclic networks, each network being of order m , with $\Sigma_1 \subseteq \Sigma_2$. Define $\mathcal{G}_{12} = \bigcap_{\mathcal{N} \in \Sigma_1 \cup \Sigma_2} \mathcal{G}_{\mathcal{N}}$. Σ_2 is said to be “more complex than” Σ_1 , if $\exists \mathcal{N}' \in \Sigma_2$ such that $\forall \mathcal{N} \in \Sigma_1, \mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}$.

Note that \mathcal{G}_{12} is always nonempty because $\mathcal{F}_m \subseteq \mathcal{G}_{12}$. Henceforth, any result that establishes a relationship of the form defined above is called a “complexity result”.

5.2 Gap-Flush Equivalence Lemma

Next, is the main lemma of this section. We show that if one class of networks is more complex than another, then inputs that satisfy the Flush Criterion are sufficient to prove this. That is, to prove this type of complexity result, one can work exclusively with Flush inputs without losing any generality. This is not obvious because Flush inputs form a subset of the more biologically-realistic Gap inputs.

Lemma 5 (Equivalence of Flush and Gap Criteria with respect to Complexity). *Let Σ_1 and Σ_2 be two sets of acyclic networks, each network being of order m , with $\Sigma_1 \subseteq \Sigma_2$. Then, Σ_2 is more complex than Σ_1 if and only if $\exists \mathcal{N}' \in \Sigma_2$ such that $\forall \mathcal{N} \in \Sigma_1, \mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m}$.*

Proof. We prove the easy direction first. If $\exists \mathcal{N}' \in \Sigma_2$ such that $\forall \mathcal{N} \in \Sigma_1, \mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m}$, then it follows that $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}$ because $\mathcal{F}_m \subseteq \mathcal{G}_{\mathcal{N}}$.

For the other direction, let $\exists \mathcal{N}' \in \Sigma_2$ such that $\forall \mathcal{N} \in \Sigma_1, \mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}$. We construct $\mathcal{F}' \subseteq \mathcal{F}_m$, so that $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}'} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}'}$. This immediately implies $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m}$. Consider arbitrary $\mathcal{N} \in \Sigma_1$. From the hypothesis we have, $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}$. Therefore $\exists \chi \in \mathcal{G}_{12}$ such that $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}}(\chi) \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}(\chi)$. Additionally, there exist $T_1, T_2 \in \mathbb{R}^+$, so that χ satisfies a T_1 -Gap Criterion for \mathcal{N} and a T_2 -Gap Criterion for \mathcal{N}' . Let $T = \max(T_1, T_2)$. Let $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{G}_{12}}(\chi) = \vec{x}'_0$ and $\mathcal{T}_{\mathcal{N}}|_{\mathcal{G}_{12}}(\chi) = \vec{x}_0$. Let $\tilde{\mathcal{F}} = \bigcup_{t \in \mathbb{R}} \Xi_{[0, 2T]}(\sigma_t(\chi))$. Note that each element of $\tilde{\mathcal{F}}$ satisfies a $2T$ -Flush Criterion. The claim, then, is that $\mathcal{T}_{\mathcal{N}'}|_{\tilde{\mathcal{F}}} \neq \mathcal{T}_{\mathcal{N}}|_{\tilde{\mathcal{F}}}$. We have $\Xi_{[0, T]}(\mathcal{T}_{\mathcal{N}'}(\Xi_{[0, 2T]}(\sigma_t(\chi)))) = \Xi_{[0, T]}(\sigma_t(\vec{x}'_0))$ and $\Xi_{[0, T]}(\mathcal{T}_{\mathcal{N}}(\Xi_{[0, 2T]}(\sigma_t(\chi)))) = \Xi_{[0, T]}(\sigma_t(\vec{x}_0))$. This follows from the fact that χ satisfies the T -Gap Criterion with both \mathcal{N} and \mathcal{N}' and therefore when \mathcal{N} and \mathcal{N}' are driven by any segment of χ of length $2T$, the output produced in the last T seconds of that interval agrees with \vec{x}_0 and \vec{x}'_0 respectively. Therefore, if $\vec{x}_0 \neq \vec{x}'_0$, it is clear that there exists a t , so that $\mathcal{T}_{\mathcal{N}'}(\Xi_{[0, 2T]}(\sigma_t(\chi))) \neq \mathcal{T}_{\mathcal{N}}(\Xi_{[0, 2T]}(\sigma_t(\chi)))$. \mathcal{F}' is obtained by taking the union of such $\tilde{\mathcal{F}}$ for every $\mathcal{N} \in \Sigma_1$. The result follows. \square

Assured by this theoretical guarantee that there is no loss of generality by doing so, we will henceforth only work with inputs satisfying the Flush Criterion while faced with the task of proving complexity results. This buys us a great deal of mathematical expedience at no cost. From now on, unless otherwise mentioned, when we say a “transformation of order m ”, we mean a map $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$.

5.3 Complexity Results for the Abstract Model

In this section, we prove some complexity results for networks whose neurons obey the abstract model of the neuron described in Chapter 3. Before we proceed, a couple of remarks are in order.

First, the perceptive reader might wonder if allowing inputs satisfying the T -Gap/Flush Criterion for arbitrarily large T is contrived, particularly since physiologically it is hard for a neuron to sustain high firing rates for arbitrarily large periods of time. Moreover, is this unrealistic assumption, being used to push through results, thereby making them only of theoretical interest? Suppose, instead, we put a global bound on T , then any transformation can be effected even by a single neuron, by making its Υ suitably (and unrealistically) high. Then one might propose putting a global bound on Υ , which brings us to the difficult question of having to quantify the relationship between the two bounds, with complexity results depending on the answer to this question. We sidestep these issues with our current formulation, but whenever we present a transformation that cannot be effected by a network architecture we do the following. Given an upper bound on the values of Υ and ρ of the neurons, we write down T as a function of these bounds, so that for all $T' \geq T$, inputs in the prescribed transformation satisfying a T' -Flush criterion cannot (provably) be mapped to the prescribed output by any network with the same architecture which also obeys the bounds on Υ and ρ .

The second remark concerns our abstract model of the neuron. The model admits a wide variety of membrane potential functions, which are meant to subsume the class of biologically-realistic potential functions. Therefore, when one proves a negative result¹ about the abstract model, the negative result applies to all of the biologically-realistic

¹ By a negative result, we mean a result that shows that a particular transformation cannot be accomplished by a certain set of networks. Likewise, by a positive result we mean a demonstration that a member of a certain set of networks can effect the transformation in question.

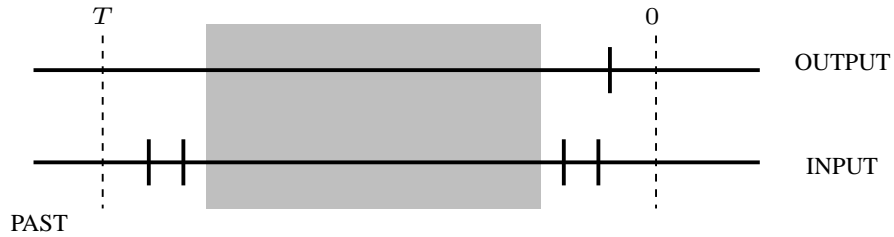


Figure 5-1. Example of a transformation that no acyclic network can effect. The shaded region is replicated over, to obtain mappings for larger and larger values of T

potential functions as well. However, when we wish to prove a positive result, some caution is required. So, for example, suppose we are trying to come up with a network that can effect a transformation, coming up with a network whose neurons merely satisfy the abstract model, leaves open the possibility that the potential functions of the neurons of the network are biologically unrealistic. Therefore, when we are proving such positive results we must restrict ourselves to as simple a model as possible. What we do is describe the neuron in the construction so that it can certainly be effected using Gerstner’s Spike Response Model SRM_0 [Gerstner & Kistler \(2002\)](#).

A final remark about causality. Whenever one prescribes a transformation, we require it to be causal. That is, if two input spike configurations are identical upto some time t , then each of their prescribed outputs must also be identical upto t . This is to forbid non-causal transformations, which are obviously unrealistic.

We now mention some complexity results. First, we point out that it is straightforward to construct a transformation that cannot be effected by any acyclic network. One of its input spike configurations with the prescribed output is shown in Figure 5-1. For larger T , the shaded region is simply replicated over and over again. Informally, the reason this transformation cannot be effected by any network is that, for any network, beyond a certain value of T , the shaded region tends to act as a flush erasing memory of the first spike. When the network receives another input spike, it is in the exact same “state” it was when it received the first spike, and therefore produces no output spike. Whether

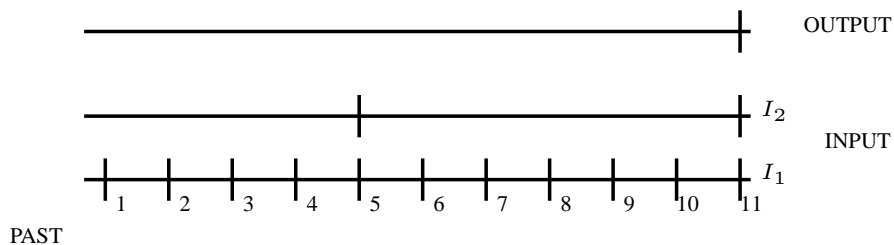


Figure 5-2. A transformation that no single neuron can effect, that a network with two neurons can.

there are other types of transformation that cannot be done by any acyclic network remains to be investigated.

Next, we prove that the set of networks with atmost two neurons is more complex than the set of single neurons. The proof is by prescribing a transformation which cannot be done by any single neuron. We then construct a network with two neurons that can effect this transformation.

Theorem 1. *For $m \geq 2$, the set of acyclic networks with atmost two neurons which have m input vertices is more complex than the set of single neurons with m input vertices.*

Proof. We first prove the result for $m = 2$ and indicate how it can be trivially extended for larger values of m . The following transformation is prescribed for $m = 2$. Let the two input spike trains in each input spike configuration, which satisfies a Flush Criterion be I_1 and I_2 . I_1 has evenly-spaced spikes starting at time instant T until 0. For the sake of exposition, we call the distance between consecutive spikes, one time unit and we number the spikes of I_1 with the first spike being the oldest one. The i th input spike configuration in the prescribed transformation satisfies a T -Flush criterion, where $T = 4i + 3$ time units. In the i th configuration, I_2 has spikes at time instants at which spike numbers $2i + 1$ and $4i + 3$ occur in I_1 . Finally, the output spike train corresponding to the the i th input spike configuration has exactly one spike at the time instant at which I_1 has spike number $4i + 3$. Figure 5-2 illustrates the transformation for $i = 2$.

Next, we prove that the transformation prescribed above cannot be effected by any single neuron. For the sake of contradiction, assume it can, by a neuron with associated Υ and ρ . Let $\max(\Upsilon, \rho)$ be bounded from above by k time units. We show that the $\frac{k}{2}$ th input spike configuration and above cannot be mapped by this neuron to the prescribed output spike train. Consider the output of the neuron at the time instants corresponding to the $(k + 1)$ th spike number and $(2k + 3)$ rd spike number of I_1 . At each of these time instants, the input received in the past k time units and the output produced by the neuron in the past k time units are the same. Therefore, the neuron's membrane potential must be identical. However, the transformation prescribes no spike in the first time instant and a spike in the second, which is a contradiction. It follows that no single neuron can effect the prescribed transformation.

We now construct a two neuron network which can carry out the transformation prescribed in Figure 5-2. The idea is to use the “infinite-time memory” mechanism of Figure 4-1 here to “remember” if the first spike on I_2 has occurred or not, and to switch the output spike pattern when it does. The network is shown in Figure 5-3(A). I_1 and I_2 arrive instantaneously at N_2 . I_1 arrives instantaneously at N_1 but I_2 arrives at N_1 after a delay of 1 time unit. Spikes output by N_1 take one time unit to arrive at N_2 , which is the output neuron of the network. The functioning of this network for $i = 2$ is described in Figure 5-3(B). The generalization for larger i is straightforward. All inputs are excitatory. N_1 is akin to the neuron described in example of Figure 4-1, in that it while the depolarization due to a spike in I_1 causes potential to cross threshold, if, additionally, the previous output spike happened one time unit ago, the associated hyperpolarization is sufficient to keep the membrane potential below threshold now. However, if there is a spike from I_2 also at the same time, the depolarization is sufficient to cause an output spike, irrespective of if there was an output spike one time unit ago. The Υ corresponding to N_2 is shorter than 1 time unit. Further, N_2 produces a spike if and only if all three of its afferent synapses receive spikes at the same time. In the figure,

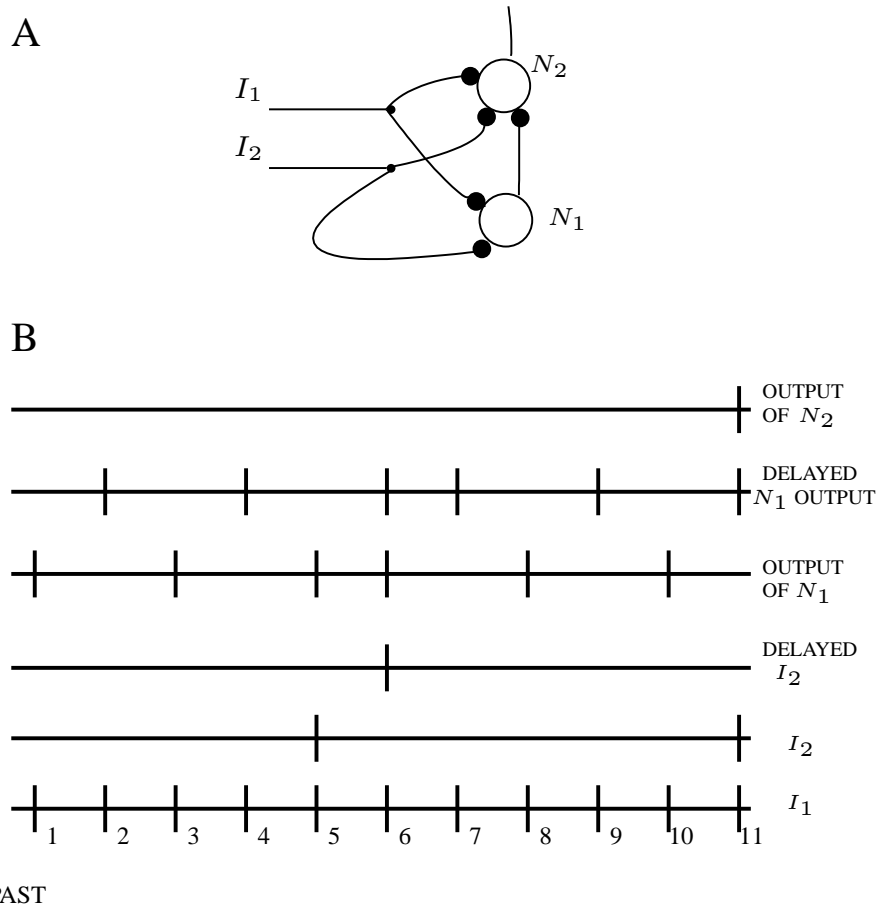


Figure 5-3. A two-neuron network effecting the transformation prescribed in Figure 5-2. A) The network that can effect the transformation described in Figure 5-2. B) Figure describing the operation of this network.

I_1 spikes at times 1, 3, 5. It spikes at 6 because it received spikes both from I_1 and I_2 at that time instant. Subsequently, it spikes at 8 and 10. The only time wherein N_2 received spikes at all three synapses at the same time is at 11, which is the prescribed time. The generalization for larger i is straightforward.

For larger m one can just have no input on the extra input spike trains and the same proof generalizes trivially. This completes the proof. \square

The above proof also suggests a large class of transformations that cannot be done by a single neuron. Informally, these are transformations for which there is no fixed bound, so

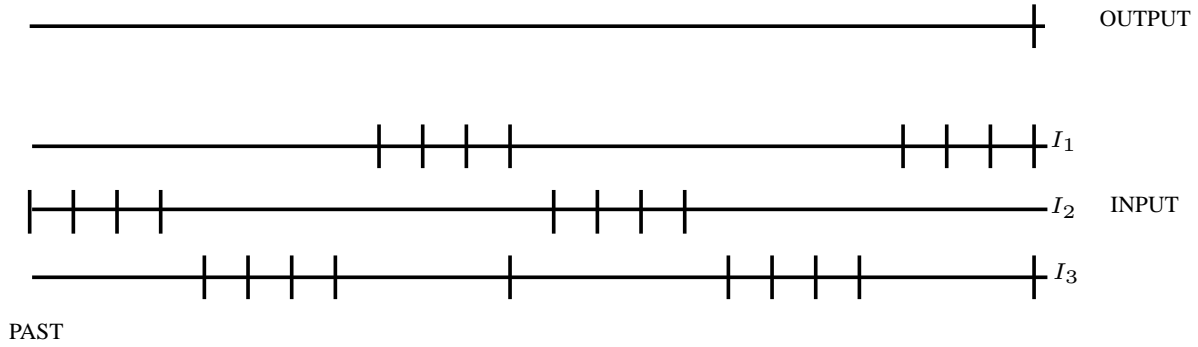


Figure 5-4. A transformation that no network with a path-disjoint architecture can effect.

that one can always determine whether there is an output spike or not, just by looking a window of past input and past output, so that the window has length atmost this bound.

The previous result might suggest that the more the number of neurons the larger the variety of transformations possible. The next complexity result demonstrates that the structure of the network architecture is crucial. That is, we can construct network architectures with arbitrarily large number of neurons which cannot perform transformations that a simple 2-neuron network can. This also shows a class of architectures that share a certain abstract graph-theoretic property also share in their inability in effecting a particular class of transformations. First, we describe a definition that makes precise the abstract graph-theoretic property that characterizes this class of architectures.

Definition 13 (Path-disjoint Architecture). An acyclic network architecture $\mathcal{A}\langle G\langle V, E \rangle, I, o, \# : \{1, 2, \dots, |I|\} \rightarrow I, \{\#_v : \{1, \dots, \text{indegree}(v)\} \rightarrow V \mid v \in (V - I)\} \rangle$, where $|I| = m$ is called “path-disjoint” if for every set of m paths, where the i th path starts at input vertex i and ends at the output vertex, the intersection of the m paths is exactly the output vertex.

Theorem 2. For $m \geq 3$, let Σ_1 be the set of all acyclic networks whose architecture is path-disjoint. Let Σ_2 be the union of Σ_1 with the set of all networks with atmost 2 neurons which have m input vertices. Then Σ_2 is more complex than Σ_1 .

Proof. We prove the theorem for $m = 3$; the generalization for larger m is straightforward.

The following transformation is prescribed for $m = 3$. Let the three input spike trains

in each input spike configuration, which satisfies a Flush Criterion by I_1 , I_2 and I_3 . As before, we will use regularly spaced spikes; we call the distance between two such consecutive spikes one time unit and number these spike time instants with the oldest being numbered 1; we call this numbering the spike index. The transformation is prescribed for inputs of various lengths with the i th input spike configuration in the prescribed transformation satisfying a T -Flush Criterion for $T = 4im$ time units. The first $2i$ time units have spikes on I_2 spaced one time unit apart, the next $2i$ on I_3 and so forth. In addition, at spike index $2im$, I_m has a single spike. The input spike pattern from the beginning is repeated once again for the latter $2im$ time units. The output spike train has exactly one spike at spike index $4im$. Figure 5-4 illustrates the transformation for $i = 2$.

Next we prove that the transformation prescribed above cannot be effected by any network in Σ_1 . For the sake of contradiction, assume that $\mathcal{N} \in \Sigma_1$ can effect the transformation. Let Υ and ρ be upper bounds on the same parameters over all of the neurons in \mathcal{N} and let d be the depth of \mathcal{N} . By construction of Σ_1 , every neuron in \mathcal{N} that is efferent on the output neuron receives input from at most $m - 1$ of the input spike trains; for, otherwise there would exist a set of m paths, one from each input vertex to the output neuron, whose intersection would contain the neuron in question. The claim, now, is that for $i > \frac{\Upsilon d}{2} + \rho$, the output neuron of \mathcal{N} has the same membrane potential at spike index $2im$ and $4im$, and therefore either has to spike at both those instants or not. Intuitively, this is so because each neuron efferent on the output neuron receives a “flush” at some point after $2im$, so that the output produced by it Υ seconds before time index $2im$ and Υ seconds before time index $4im$ are the same. We will leave the task of verifying this as an exercise to the reader.

We now construct a two-neuron network that can effect this transformation. The construction is similar to the one used in Theorem 1. For $m = 3$, the network is shown in Figure 5-5(A). I_1 , I_2 and I_3 arrive instantaneously at N_1 and N_2 . Spikes output by N_1 take two time units to arrive at N_2 , which is the output neuron of the network. The

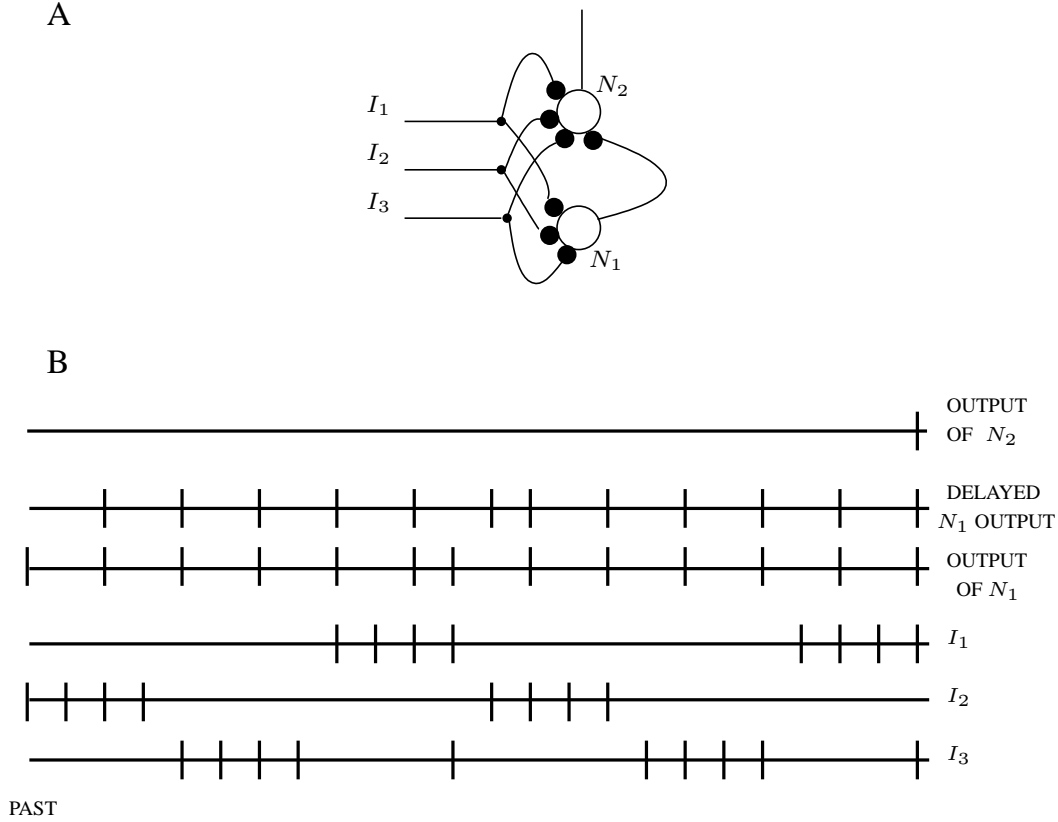


Figure 5-5. A two-neuron network effecting the transformation prescribed in Figure 5-4.

A) Network that can effect the transformation described in Figure 5-4.

B) Figure describing the operation of this network.

functioning of this network for $i = 2$ is described in Figure 5-5(B). The generalization for larger i is straightforward. All inputs are excitatory. N_1 is akin to the the N_1 used in the network in Theorem 1 except that that periodic input may arrive from any one of I_1 , I_2 or I_3 . As before, if two input spikes arrive at the same time, as in spike index $2im$, the depolarization is sufficient to cause an output spike, irrespective of if there was an output spike one time unit ago. Again, the Υ corresponding to N_2 is shorter than 1 time unit and N_2 produces a spike if and only if three of its afferent synapses receive spikes at the same time. As before the idea is that at time $2im$, N_2 , receives two spikes, but not a spike from N_1 , since it is “out of sync”. However, at time $4im$, additionally, there is a spike from N_1 arriving at N_2 , which causes N_2 to spike. □

The theorem therefore demonstrates that it is feasible to attempt to link transformational properties of networks to abstract graph-theoretic properties of their architecture. It also establishes that network structure is crucial in determining transformational ability.

CHAPTER 6 COMPLEXITY RESULTS AND NETWORK DEPTH: A BARRIER AT DEPTH TWO

While, in the previous chapter, we have proved a number of complexity results, the perceptive reader will observe that we did not explicitly address the question of complexity results pertaining to depth, in general. In other words, the question is, does increase in depth of the network, in general, buy us a larger variety of transformations effectable by it? More formally, is the set of “all” acyclic networks more complex than the set of acyclic networks of depth k , for $k \geq 2$? Much to our surprise¹, it turns out that the answer is No, even for $k = 2$, and we can prove it. That is, given an arbitrary acyclic network (consisting of neurons obeying the abstract model described in Chapter 3, there exists a network of depth two (equipped with neurons obeying the same abstract model), so that the latter network induces “exactly” the same transformation as the former. The implication of this result is that one needs to add more axioms to the abstract model of Chapter 3, in order to break this barrier to the manifestation of complexity results.²

The difficulty in proving that every acyclic network, having arbitrary depth, has an equivalent network of depth two, appears to be in devising a way of “collapsing” the depth of the former network, while keeping the effected transformation the same. Our proof actually does not demonstrate this head-on, but instead proves it to be the case indirectly. The broad attack consists of starting off with a certain subset of the set of all possible transformations and showing that every transformation that lies outside this subset cannot be effected by “any” acyclic network. Thereafter, we prove that every transformation

¹ Indeed, the author spent many a moon attempting to prove the aforementioned type of complexity results. In hindsight, of course, something like this should have been expected at some point, given that the abstract model assumed so little about single neurons.

² Note that the complexity results proved thus far still hold, if the “new” abstract model also has all the axioms of the one described in Chapter 3.

in this subset can in fact be effected by an acyclic network of depth two, by providing a construction.

6.1 Technical Structure of the Proof

The main theorem of this chapter is the following.

Theorem. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by an acyclic network, then it can be effected by an acyclic network of depth two.*

The above theorem follows from the following two lemmas which are proved in the two subsections that follow:

Lemma. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by an acyclic network, then it is causal, time-invariant and resettable.*

Lemma. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, time-invariant and resettable, then it can be effected by an acyclic network of depth two.*

6.2 Causal, Time-Invariant and Resettable Transformations

In this section, we first define notions of causal, time-invariant and resettable transformations³. Transformations that are causal, time-invariant and resettable form a strict subset of the set of all transformations. Note that these notions exist independent of the existence of neurons or their networks. We then show that transformations lying outside this subset cannot be effected by any acyclic network. This is the relatively easy part of the proof. The next section proves the harder part, namely that every transformation in this subset can indeed be effected by an acyclic network of depth equal to two.

As in systems theory, informally, a “causal transformation” is one whose current output depends only on its current input and past input (and not future input).

Abstractly, it is convenient to define a causal transformation as one that given two

³ Recall that when we say transformation without further qualification, we mean one, of the form $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$.

different inputs that are identical up to a certain point in time, also have their outputs, according to the transformation, be identical up to (atleast) the same point.

Definition 14 (Causal Transformation). A transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be “causal” if, for every $\chi_1, \chi_2 \in \mathcal{F}_m$, with $\Xi_{[t, \infty)}\chi_1 = \Xi_{[t, \infty)}\chi_2$, for some $t \in \mathbb{R}$, we have $\Xi_{[t, \infty)}\mathcal{T}(\chi_1) = \Xi_{[t, \infty)}\mathcal{T}(\chi_2)$.

Again, as in systems theory, a “time-invariant transformation” is one for which an input which is a time-shifted version of another input has as output the time-shifted version of the output corresponding to the latter input. To keep the definition kosher, we also need to ensure that the time-shifted input in fact also satisfies the flush criterion.

Definition 15 (Time-Invariant Transformation). A transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be “time-invariant” if, for every $\chi \in \mathcal{F}_m$ and every $t \in \mathbb{R}$ with $\sigma_t(\chi) \in \mathcal{F}_m$, we have $\mathcal{T}(\sigma_t(\chi)) = \sigma_t(\mathcal{T}(\chi))$.

A “resettable transformation” is one for which there exists a positive real number W , so that an input gap of the form $(t, t + W]$ “resets” it, i.e. output beyond t is independent of input received before it. Again, abstractly, it becomes convenient to say that the output in this case is identical to that produced by an input which has no spikes before t , but is identical to the present input thereafter.

Definition 16 (W -Resettable Transformation). For $W \in \mathbb{R}^+$, a transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be “ W -resettable” if, for every $\chi \in \mathcal{F}_m$ which has a gap in the interval $(t, t + W]$, for some $t \in \mathbb{R}$, we have $\Xi_{(-\infty, t]}\mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t]}\chi)$.

Definition 17 (Resettable Transformation). A transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be “resettable” if there exists a $W \in \mathbb{R}^+$, so that it is W -resettable.

Next, we prove that every transformation that can be effected by an acyclic network is causal, time-invariant and resettable. This implies that every transformation that is not causal, time-invariant and resettable cannot be effected by any acyclic network.

Lemma 6. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by an acyclic network then it is causal, time-invariant and resettable.*

Proof. Let $\mathcal{N}\langle \mathcal{A}, N, L : (V - I) \rightarrow N \rangle$ be a network that effects $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$. We first show that it is causal. Consider arbitrary $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_{[t, \infty)}\chi_1 = \Xi_{[t, \infty)}\chi_2$, for some $t \in \mathbb{R}$. We wish to show that $\Xi_{[t, \infty)}\mathcal{T}(\chi_1) = \Xi_{[t, \infty)}\mathcal{T}(\chi_2)$. Let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i , where d is the depth of \mathcal{N} . Each neuron $\mathbf{N} \in N_1$ receives all its inputs from spike trains in χ . Let $\chi'_1 = \bigsqcup_{i=1}^{\text{indegree}(v)} \Pi_{\#^{-1}(\#_v(i))}(\chi_1)$ and $\chi'_2 = \bigsqcup_{i=1}^{\text{indegree}(v)} \Pi_{\#^{-1}(\#_v(i))}(\chi_2)$, where $v = L^{-1}(\mathbf{N})$. When the network receives χ_1 and χ_2 as input, \mathbf{N} receives χ'_1 and χ'_2 respectively as input. Also, clearly $\Xi_{[t, \infty)}\chi'_1 = \Xi_{[t, \infty)}\chi'_2$. Let \vec{x}'_1 and \vec{x}'_2 be the output produced by \mathbf{N} on receiving χ'_1 and χ'_2 respectively. Since $\chi'_1, \chi'_2 \in \mathcal{F}_m$, there exists a $T \in \mathbb{R}^+$, so that $\Xi_{[T, \infty)}\chi'_1 = \Xi_{[T, \infty)}\chi'_2 = \vec{\phi}^{m'}$, where m' is the number of inputs to \mathbf{N} . Therefore, by Axiom (3) of the neuron, we have $\Xi_{[T, \infty)}\vec{x}'_1 = \Xi_{[T, \infty)}\vec{x}'_2 = \vec{\phi}$. Now, for all $t' \in \mathbb{R}$, $t' \in \vec{x}'_j$ if and only if $P_{\mathbf{N}}(\Xi_{[0, \Upsilon_{\mathbf{N}}]}(\sigma_{t'}(\chi'_j)), \Xi_{[0, \rho_{\mathbf{N}}]}(\sigma_{t'}(\vec{x}'_j))) = \tau_{\mathbf{N}}$, for $j = 1, 2$. It is immediate that for $t' \leq t$, we have $\Xi_{[0, \Upsilon_{\mathbf{N}}]}(\sigma_{t'}(\chi'_1)) = \Xi_{[0, \Upsilon_{\mathbf{N}}]}(\sigma_{t'}(\chi'_2))$. Now, by an induction argument on the spike number since T , it is straightforward to show that for all $t' \leq t$, $\Xi_{[0, \rho_{\mathbf{N}}]}(\sigma_{t'}(\vec{x}'_1)) = \Xi_{[0, \rho_{\mathbf{N}}]}(\sigma_{t'}(\vec{x}'_2))$. Thus, we have $\Xi_{[t, \infty)}\vec{x}'_1 = \Xi_{[t, \infty)}\vec{x}'_2$. Similarly, using a straightforward induction argument on depth, one can show that for every neuron in the network, its output until time instant t is identical in either case. We therefore have $\Xi_{[t, \infty)}\mathcal{T}(\chi_1) = \Xi_{[t, \infty)}\mathcal{T}(\chi_2)$.

Next, we show that it is time-invariant. Consider arbitrary $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$ with $\sigma_t(\chi) \in \mathcal{F}_m$. We wish to show that $\mathcal{T}(\sigma_t(\chi)) = \sigma_t(\mathcal{T}(\chi))$. As before, let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i , where d is the depth of \mathcal{N} . Each neuron $\mathbf{N} \in N_1$ receives all its inputs from spike trains in χ . Let $\chi' = \bigsqcup_{i=1}^{\text{indegree}(v)} \Pi_{\#^{-1}(\#_v(i))}(\chi)$. When the network receives χ and $\sigma_t(\chi)$ as input, \mathbf{N} receives χ' and $\sigma_t(\chi')$ respectively as input. Let \vec{x}'_1 and \vec{x}'_2 be the output produced by \mathbf{N} on receiving χ' and $\sigma_t(\chi')$ as input respectively. We wish to show that $\vec{x}'_2 = \sigma_t(\vec{x}'_1)$. Since $\chi' \in \mathcal{F}_m$, there exists a $T \in \mathbb{R}^+$, so that $\Xi_{[T, \infty)}\chi' = \Xi_{[T-t, \infty)}\sigma_t(\chi') = \vec{\phi}^{m'}$, where m' is the number of inputs to \mathbf{N} . Therefore, by Axiom (3) of the neuron, we have $\Xi_{[T, \infty)}\vec{x}'_1 = \Xi_{[T-t, \infty)}\vec{x}'_2 = \vec{\phi}$. Now, for all $t' \in \mathbb{R}$, $t' \in \vec{x}'_1$ if and only if $P_{\mathbf{N}}(\Xi_{[0, \Upsilon_{\mathbf{N}}]}(\sigma_{t'}(\chi')), \Xi_{[0, \rho_{\mathbf{N}}]}(\sigma_{t'}(\vec{x}'_1))) = \tau_{\mathbf{N}}$. It is therefore straightforward to make

an induction argument on the spike number, starting from the oldest spike in \vec{x}'_1 to show that \vec{x}'_1 has a spike at some t' iff \vec{x}'_2 has a spike at $t' - t$ and therefore we have $\vec{x}'_2 = \sigma_t(\vec{x}'_1)$. Similarly, using a straightforward induction argument on depth, one can show that for every neuron in the network, its output in the second case is a time-shifted version of the one in the first case. We therefore have $\mathcal{T}(\sigma_t(\chi)) = \sigma_t(\mathcal{T}(\chi))$.

Finally, we show that \mathcal{T} is resettable. Let Υ and ρ be upper bounds on those parameters over all the neurons in \mathcal{N} . If $\Upsilon < \rho$, then set $\Upsilon = \rho$. The claim is that for $W = d(\Upsilon + \rho) + \rho + \delta$, for arbitrary $\delta > 0$, \mathcal{T} is W -resettable. Consider arbitrary $\chi \in \mathcal{F}_m$ so that χ has a gap in the interval $(t, t + \delta + d(\Upsilon + \rho) + \rho]$, for some $t \in \mathbb{R}^+$. Thus χ has a gap in the interval $[t + \delta, t + \delta + d(\Upsilon + \rho) + \rho]$. We will show that $\Xi_{(-\infty, t + \delta]} \mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t + \delta]} \chi)$ which implies the required result $\Xi_{(-\infty, t]} \mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t]} \chi)$. As before, let $N_i \subseteq N$, for $1 \leq i \leq d$, be the set of neurons in N of depth i , where d is the depth of \mathcal{N} . Each neuron $\mathbf{N} \in N_1$ receives all its inputs from spike trains in χ . Therefore by Axiom (3) of the neuron, it is straightforward to see that the output of \mathbf{N} has a gap in the interval $(t, t + \delta + (d - 1)(\Upsilon + \rho) + 2\rho]$. By similar arguments, we have that output of each neuron $\mathbf{N} \in N_i$, for $1 \leq i \leq d$ has a gap in the interval $(t, t + \delta + (d - i)(\Upsilon + \rho) + (i + 1)\rho]$. Thus in particular the output neuron has a gap in the interval $[t + \delta, t + \delta + (d + 1)\rho]$. Since $d \geq 1$, the Gap Lemma applies, and at time instant $t + \delta$ the output of the output neuron depends on spikes in the interval $[t + \delta, t + \delta + (\Upsilon + \rho)]$ of its inputs. All inputs to the output neuron have a gap in the interval $[t + \delta, t + \delta + (\Upsilon + \rho) + d\rho]$, since they have depth at most $(d - 1)$. Since those inputs have a gap in the interval $[t + \delta + (\Upsilon + \rho), t + \delta + (\Upsilon + \rho) + d\rho]$, for $d \geq 2$, the Gap Lemma applies and the output neuron's output at time instant $t + \delta$ depends on outputs of the "penultimate" layer in the interval $[t + \delta, t + \delta + 2(\Upsilon + \rho)]$. Therefore by similar arguments, the output of the output neuron at time instant $t + \delta$ at most depends on inputs from χ in the interval $[t + \delta, t + \delta + d(\Upsilon + \rho)]$. That is to say that $\mathcal{T}(\chi')$, for every χ' identical to χ in the interval $(-\infty, t + \delta]$, has the same output as $\mathcal{T}(\chi)$ in that same interval, following the corollary to the Gap Lemma. In particular,

$\Xi_{(-\infty, t+\delta]}\chi$ is one such χ' . We therefore have $\Xi_{(-\infty, t+\delta]}\mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t+\delta]}\chi)$ which implies $\Xi_{(-\infty, t]}\mathcal{T}(\chi) = \mathcal{T}(\Xi_{(-\infty, t]}\chi)$. Thus, \mathcal{T} is resettable. \square

6.3 Construction of a Depth Two Acyclic Network for every Causal, Time-Invariant and Resettable Transformation

In this section, we describe a construction of a depth two acyclic network for every causal, time-invariant and resettable transformation. More formally, we prove the following lemma.

Lemma. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, time-invariant and resettable, then it can be effected by an acyclic network of depth two.*

Before we dive into the proofs, we offer some intuition.

Suppose we had a transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ which is causal, time-invariant and resettable. For the moment, pretend it satisfies the following property: For every input spike configuration satisfying a flush criterion, there exist constant input and output “windows”, so that at any point in time, just given knowledge of spikes in those windows of past input and output, one can unambiguously determine if the transformation prescribes an output spike or not. Intuitively, it seems reasonable that such a transformation can in fact be effected by a single neuron⁴ by setting the Υ and ρ of the neuron to the sizes of the input and output window mentioned above.

Of course, one easily sees that not every transformation that is causal, time-invariant and resettable satisfies the aforementioned property. That is, there could exist two different input instances, which have past inputs and outputs be identical in the aforementioned windows at some points in time; yet in one instance, the transformation prescribes an output spike, whereas it prescribes none in the other. Indeed, the two input instances do differ at some point in the past, for otherwise the transformation would not be causal. Therefore, in such a situation, it is natural to ask if a single “intermediate”

⁴ Presuming that no axioms of the neuron are violated.

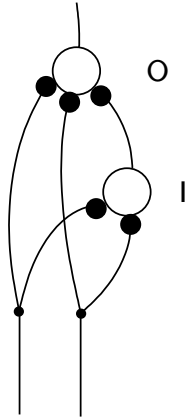


Figure 6-1. Network architecture for $m = 2$.

neuron can “break the tie”. That is, if two input instances differ at some point in the past, the output of the intermediate neuron since then, in any interval of time of length U , must be different. This is so that it can disambiguate the two, were an output spike demanded for one instance but not the other. Unfortunately, this exact property cannot be achieved by any single neuron because the transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is resettable, and so is every transformation induced by the intermediate neuron. In other words, the problem is that, suppose two input instances differ at a certain point in time; however, since then, both have had an arbitrarily large input gap. The input gap serves to “erase memory” in any network that received it and therefore it cannot disambiguate the two inputs beyond this gap. Fortunately, it does not have to, since this gap also causes a “reset” in the transformation (which is resettable). Thus, we can make do with a slightly weaker condition; that the intermediate neuron is only guaranteed to disambiguate, when it is required to do so. That is, suppose there are two input instances, whose outputs according to $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ are different at certain points in time. Then, the corresponding inputs must be different too at some point in the past with no resets in the intervening time and therefore the intermediate neuron should break the tie. Additionally, we say that the outputs of the intermediate neuron in the preceding U seconds are guaranteed to be different, only if the inputs themselves in the past U seconds are not different.

The network we have in mind is illustrated in Figure 6-1, for $m = 2$. In the following proposition, we prove that if the intermediate neuron satisfies the weaker condition alluded to above, then there exists an output neuron, so that the network effects the transformation in question.

Proposition 2. Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be causal, time-invariant and resettable. Let I be a neuron with $\mathcal{T}_{\mathsf{I}} : \mathcal{F}_m \rightarrow \mathcal{S}$, so that for each $\chi \in \mathcal{F}_m$, $\mathcal{T}_{\mathsf{I}}(\chi)$ is consistent with χ with respect to I . Further, suppose there exists a $U \in \mathbb{R}^+$ so that for all $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, we have $\Xi_{[0,U]}(\sigma_{t_1}(\mathcal{T}_{\mathsf{I}}(\chi_1) \sqcup \chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2}(\mathcal{T}_{\mathsf{I}}(\chi_2) \sqcup \chi_2))$.

Then there exists a neuron O , so that for every $\chi \in \mathcal{F}_m$, $\mathcal{T}(\chi)$ is consistent with $\mathcal{T}_{\mathsf{I}}(\chi) \sqcup \chi$ with respect to O .

Proof. Assume that the hypothesis in the proposition above is true. Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be W -Resettable for some $W \in \mathbb{R}^+$.

We first show a construction for the neuron O , prove that it obeys all the axioms and then show that it has the property that for every $\chi \in \mathcal{F}_m$, $\mathcal{T}(\chi)$ is consistent with $\mathcal{T}_{\mathsf{I}}(\chi) \sqcup \chi$ with respect to O .

We first construct the neuron $\mathsf{O} \langle \alpha_{\mathsf{O}}, \Upsilon_{\mathsf{O}}, \rho_{\mathsf{O}}, \tau_{\mathsf{O}}, \lambda_{\mathsf{O}}, m_{\mathsf{O}}, P_{\mathsf{O}} : \bar{\mathcal{S}}_{[0, \Upsilon_{\mathsf{O}}]}^{m_{\mathsf{O}}} \times \bar{\mathcal{S}}_{[0, \rho_{\mathsf{O}}]} \rightarrow [\lambda_{\mathsf{O}}, \tau_{\mathsf{O}}] \rangle$. Set $\alpha_{\mathsf{O}} = \alpha$ and $\rho_{\mathsf{O}}, \tau_{\mathsf{O}} \in \mathbb{R}^+$, $\lambda_{\mathsf{O}} \in \mathbb{R}^-$ arbitrarily with $\rho_{\mathsf{O}} \geq \alpha_{\mathsf{O}}$. Set $\Upsilon_{\mathsf{O}} = \max\{U, W\}$ and $m_{\mathsf{O}} = m + 1$. The function $P_{\mathsf{O}} : \bar{\mathcal{S}}_{[0, \Upsilon_{\mathsf{O}}]}^{m_{\mathsf{O}}} \times \bar{\mathcal{S}}_{[0, \rho_{\mathsf{O}}]} \rightarrow [\lambda_{\mathsf{O}}, \tau_{\mathsf{O}}]$ is constructed as follows.

For $\chi' \in \bar{\mathcal{S}}_{[0, \Upsilon_{\mathsf{O}}]}^{m_{\mathsf{O}}}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_{\mathsf{O}}]}$, set $P_{\mathsf{O}}(\chi', \vec{x}'_0) = \tau_{\mathsf{O}}$ and $P_{\mathsf{O}}(\chi', \vec{\phi}) = \tau_{\mathsf{O}}$ if and only if there exists $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$ so that $\Xi_t \mathcal{T}(\chi) = \langle t \rangle$ and $\chi' = \Xi_{[0, \Upsilon_{\mathsf{O}}]}(\sigma_t(\mathcal{T}_{\mathsf{I}}(\chi) \sqcup \chi))$ and $\vec{x}'_0 = \Xi_{[0, \rho_{\mathsf{O}}]}(\sigma_t(\mathcal{T}(\chi)))$. Everywhere else, the value of this function is set to zero. Next, we show it obeys all of the axioms of the single neuron.

We prove that O satisfies Axiom (1) by showing that its contrapositive is true. Let $\chi' \in \bar{\mathcal{S}}_{[0, \Upsilon_{\mathsf{O}}]}^{m_{\mathsf{O}}}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_{\mathsf{O}}]}$ be arbitrary so that $P_{\mathsf{O}}(\chi', \vec{x}'_0) = \tau_{\mathsf{O}}$. If $\vec{x}'_0 = \vec{\phi}$, Axiom (1) is immediately satisfied. Thus consider the case when $\vec{x}'_0 = \langle x_0^1, x_0^2, \dots, x_0^{k'} \rangle$. Then $x_0^1 \geq \alpha$, otherwise, from the construction of $P_{\mathsf{O}}(\cdot)$, it is straightforward to show that there exists a $\chi \in \mathcal{F}_m$ with $\mathcal{T}(\chi) \notin \mathcal{S}$.

Next, we prove that \mathbf{O} satisfies Axiom (2). Let $\chi' \in \bar{\mathcal{S}}_{[0, \Upsilon_{\mathbf{O}}]}^{m_{\mathbf{O}}}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_{\mathbf{O}}]}$ be arbitrary. If $P_{\mathbf{O}}(\chi', \vec{x}'_0) = \tau_{\mathbf{O}}$, then it is immediate from the construction that $P_{\mathbf{O}}(\chi', \vec{\phi}) = \tau_{\mathbf{O}}$. On the contrary, if $P_{\mathbf{O}}(\chi', \vec{x}'_0) \neq \tau_{\mathbf{O}}$, from the construction of \mathbf{O} , we have $P_{\mathbf{O}}(\chi', \vec{x}'_0) = 0$. Then the condition in the hypothesis implies that $P_{\mathbf{O}}(\chi', \vec{\phi}) \neq \tau_{\mathbf{O}}$. Therefore, $P_{\mathbf{O}}(\chi', \vec{\phi}) = 0$. Thus, Axiom (2) is satisfied either way.

With Axiom (3), we wish to show $P_{\mathbf{O}}(\vec{\phi}^{m+1}, \vec{\phi}) = 0$. Here, we will show that $P_{\mathbf{O}}(\vec{x}_1 \sqcup \vec{\phi}^m, \vec{x}'_0) = 0$, for all $\vec{x}_1 \in \bar{\mathcal{S}}_{[0, \Upsilon_{\mathbf{O}}]}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_{\mathbf{O}}]}$ which implies the required result. Assume, for the sake of contradiction, that there exists a $\vec{x}_1 \in \bar{\mathcal{S}}_{[0, \Upsilon_{\mathbf{O}}]}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_{\mathbf{O}}]}$, so that $P_{\mathbf{O}}(\vec{x}_1 \sqcup \vec{\phi}^m, \vec{x}'_0) = \tau_{\mathbf{O}}$. From the construction of \mathbf{O} , this implies that there exists $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$ so that $\Xi_t \mathcal{T}(\chi) = \langle t \rangle$ and $\Xi_{[0, \Upsilon_{\mathbf{O}}]}(\sigma_t(\chi)) = \vec{\phi}^m$. That is, χ has a gap in the interval $[t, t + W]$, since $\Upsilon_{\mathbf{O}} \geq W$. Since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, time-invariant and W -resettable, by Corollary 3, we have $\Xi_t \mathcal{T}(\chi) = \vec{\phi}$, which is a contradiction. Therefore, we have $P_{\mathbf{O}}(\vec{x}_1 \sqcup \vec{\phi}^m, \vec{x}'_0) \neq \tau_{\mathbf{O}}$ and by construction of \mathbf{O} , $P_{\mathbf{O}}(\vec{x}_1 \sqcup \vec{\phi}^m, \vec{x}'_0) = 0$, for all $\vec{x}_1 \in \bar{\mathcal{S}}_{[0, \Upsilon_{\mathbf{O}}]}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_{\mathbf{O}}]}$. This implies $P_{\mathbf{O}}(\vec{\phi}^{m+1}, \vec{\phi}) = 0$.

Finally, we wish to show that for every $\chi \in \mathcal{F}_m$, $\mathcal{T}(\chi)$ is consistent with $\mathcal{T}_1(\chi) \sqcup \chi$ with respect to \mathbf{O} . That is, we wish to show that for every $\chi \in \mathcal{F}_m$ and for every $t \in \mathbb{R}$, $\Xi_0 \sigma_t(\mathcal{T}(\chi)) = \langle 0 \rangle$ if and only if $P_{\mathbf{O}}(\Xi_{[0, \Upsilon_{\mathbf{O}}]}(\sigma_t(\mathcal{T}_1(\chi) \sqcup \chi)), \Xi_{[0, \rho_{\mathbf{O}}]}(\sigma_t(\mathcal{T}(\chi)))) = \tau_{\mathbf{O}}$. Consider arbitrary $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$. If $\Xi_0 \sigma_t(\mathcal{T}(\chi)) = \langle 0 \rangle$, then it is immediate from the construction of \mathbf{O} that $P_{\mathbf{O}}(\Xi_{[0, \Upsilon_{\mathbf{O}}]}(\sigma_t(\mathcal{T}_1(\chi) \sqcup \chi)), \Xi_{[0, \rho_{\mathbf{O}}]}(\sigma_t(\mathcal{T}(\chi)))) = \tau_{\mathbf{O}}$. To prove the converse, suppose $\Xi_0 \sigma_t(\mathcal{T}(\chi)) \neq \langle 0 \rangle$. Then, from the condition in the hypothesis, it follows that for all $\tilde{\chi} \in \mathcal{F}_m$ and for all $\tilde{t} \in \mathbb{R}$ with $\Xi_{[0, \Upsilon_{\mathbf{O}}]}(\sigma_{\tilde{t}}(\mathcal{T}_1(\tilde{\chi}) \sqcup \tilde{\chi})) = \Xi_{[0, \Upsilon_{\mathbf{O}}]}(\sigma_{\tilde{t}}(\mathcal{T}_1(\chi) \sqcup \chi))$, we have $\Xi_0 \sigma_{\tilde{t}}(\mathcal{T}(\tilde{\chi})) = \Xi_0 \sigma_{\tilde{t}}(\mathcal{T}(\chi)) \neq \langle 0 \rangle$. Therefore, from the construction, we have $P_{\mathbf{O}}(\Xi_{[0, \Upsilon]}(\sigma_t(\mathcal{T}_1(\chi) \sqcup \chi)), \Xi_{[0, \rho]}(\sigma_t(\mathcal{T}(\chi)))) \neq \tau_{\mathbf{O}}$. \square

We now give an intuitive description of how the intermediate neuron is constructed, so as to satisfy the said condition. The basic idea is to “encode”, in the time difference of two successive output spikes, the positions of all the input spikes that occurred since the last output gap of the form $(t, t + W]$, which we call a “reset gap” from now on, for the

sake of exposition. Such pairs of output spikes occur once every p seconds, with the time difference within each pair being a function of the time difference within the previous pair and the input spikes encountered since. Intuitively, it is convenient to think of this encoding as one from which we can “reconstruct” the entire past input spike train since the last reset. We first describe the encoding function for the case of a single input spike train after which we remark on how it can be generalized.

So, suppose the time difference of the successive spikes lies in the interval $[0, 1)$. Define the encoding function as $\varepsilon_0 : [0, 1) \times \bar{\mathcal{S}}_{(0,p]} \rightarrow [0, 1)$. p is chosen to be such that there are at most 8 spikes in any interval of the form $(t, t + p]$. We now describe how $\varepsilon_0(e, \vec{x})$ is computed, given $e \in [0, 1)$ and $\vec{x} = \langle x^1, x^2, \dots, x^k \rangle$, such that each spike time lies in the interval $(0, p]$. Let e have a decimal expansion, so that $e = 0.c_1s_1c_2s_2c_3s_3\dots$. Accordingly, let $c = 0.c_1c_2c_3\dots$ and $s = 0.s_1s_2s_3\dots$. c is a real number that encodes the number of spikes in each interval of length p encountered, since the last reset. Since each interval of length p has between 0 and 8 spikes, the digit 9 is used as a “termination symbol”. So, for example, suppose there have been 4 intervals of length p , since the last reset with 5, 0, 8 and 2 spikes apiece respectively, then $c = 0.8059$ and $c' = 0.28059$, where c' is the “updated” value of c . Likewise, s is a real number that stores the positions of all input spikes encountered since the last reset. Let each spike time be of the form $x^i = 0.x_1^i x_2^i x_3^i \dots \times 10^q$, for appropriate q . Then the updated value of s is $s' = 0.x_1^1 x_1^2 \dots x_1^k s_1 x_2^1 x_2^2 \dots x_2^k s_2 \dots$. Suppose the c' and s' obtained above were of the form $c' = 0.c'_1 c'_2 c'_3 \dots$ and $s' = 0.s'_1 s'_2 s'_3 \dots$, then $\varepsilon_0(e, \vec{x}) = 0.c'_1 s'_1 c'_2 s'_2 \dots$. Suppose the input were a spike configuration of order m , then for each spike train an encoding would be computed as above and in the final step, the m real numbers obtained would be interleaved together, so as to produce the encoding.

Given knowledge of the encoding function, we now describe how \mathcal{I} exactly works. Figure 6-2 provides an illustration. Suppose $\chi \in \mathcal{F}_m$ is an input spike train. Let its oldest spike be T seconds ago. Then \mathcal{I} produces a spike at time $T - p$ and at every $T - kp$, for

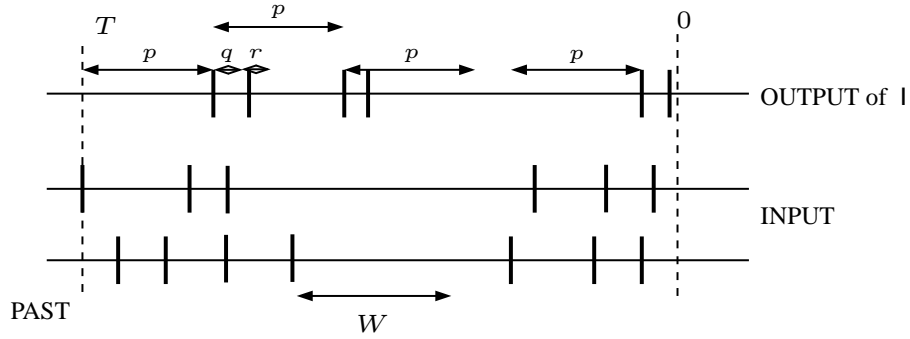


Figure 6-2. Example illustrating the operation of I .

$k \in \mathbb{Z}^+$, unless in the previous p seconds to when it is to spike there is a gap⁵ of the form $(t, t + W]$. For the sake of exposition, let's call these the “clock” spikes. Now, suppose there is a gap of the form $(t, t + W]$ in the input and there is an input spike at time t , then the neuron spikes at time $t - p$ and every p seconds thereafter subject to the same “rules” as above. These clock spikes are followed by “encoding” spikes, which occur at least q seconds after the clock spike, but at most $q + r$ seconds after, where q is greater than the absolute refractory period α . As expected, the position of the current encoding spike is a function of the time difference between the previous encoding and clock spike⁶ and the positions of the input spikes in the past p seconds. The output of the encoding function is appropriately scaled to “fit” in this interval of length r ; the details are available in the proof.

The claim then is that if two input spike trains are different at some point with no intervening “reset” gaps, then the output of I in the past U seconds, where $U = p + q + r$ will be different. Intuitively, this is because the difference between the latest encoding and clock spike in each case would be different, as they encode different “histories” of input spikes.

⁵ We set $W > p$ to force a spike at $T - p$.

⁶ unless the present clock spike is the first after a reset gap in the input.

Finally, we remark that the above is just an informal description that glosses over several technical details that the reader is invited to enjoy in the proof.

Proposition 3. Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be causal, time-invariant and resettable. Then there exists a neuron l and $U \in \mathbb{R}^+$ so that for all $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, we have $\Xi_{[0,U]}(\sigma_{t_1}(\mathcal{T}_1(\chi_1) \sqcup \chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2}(\mathcal{T}_1(\chi_2) \sqcup \chi_2))$, where $\mathcal{T}_1 : \mathcal{F}_m \rightarrow \mathcal{S}$ such that for each $\chi \in \mathcal{F}_m$, $\mathcal{T}_1(\chi)$ is consistent with χ with respect to l .

Proof. Assume that the hypothesis in the proposition above is true. Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be W' -Resettable for some $W' \in \mathbb{R}^+$. Set $W = \max\{W', 12\alpha\}$. One readily verifies that $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is also W -resettable.

We first show a construction for the neuron l , prove that it obeys all the axioms and then show that it has the property that there exists a $U \in \mathbb{R}^+$ so that for all $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, we have $\Xi_{[0,U]}(\sigma_{t_1}(\mathcal{T}_1(\chi_1) \sqcup \chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2}(\mathcal{T}_1(\chi_2) \sqcup \chi_2))$, where $\mathcal{T}_1 : \mathcal{F}_m \rightarrow \mathcal{S}$ such that for each $\chi \in \mathcal{F}_m$, $\mathcal{T}_1(\chi)$ is consistent with χ with respect to l .

We first construct the neuron $\mathsf{l}(\alpha_1, \Upsilon_1, \rho_1, \tau_1, \lambda_1, m_1, P_1 : \bar{\mathcal{S}}_{[0,\Upsilon_1]}^{m_1} \times \bar{\mathcal{S}}_{[0,\rho_1]} \rightarrow [\lambda_1, \tau_1])$. Set $\alpha_1 = \alpha$. Let $p, q, r \in \mathbb{R}^+$, with⁷ $p = 8\alpha, q = 2\alpha$ and $r = \alpha$. Set $\Upsilon_1 = p + q + r + W$, $\rho_1 = p + q + r$ and $m_1 = m$. Let $\tau_1 \in \mathbb{R}^+, \lambda_1 \in \mathbb{R}^-$ be chosen arbitrarily. The function $P_1 : \bar{\mathcal{S}}_{[0,\Upsilon_1]}^{m_1} \times \bar{\mathcal{S}}_{[0,\rho_1]} \rightarrow [\lambda_1, \tau_1]$ is constructed as follows.

For $\chi \in \bar{\mathcal{S}}_{[0,\Upsilon_1]}^{m_1}$ and $\vec{x}_0 \in \bar{\mathcal{S}}_{[0,\rho_1]}$, set $P_1(\chi, \vec{x}_0) = \tau_1$ and $P_1(\chi, \vec{\phi}) = \tau_1$ if and only if one of the following is true; everywhere else, the function is set to zero.

1. $\Xi_{(p,p+W]}\chi = \vec{\phi}^{m_1}, \Xi_p\chi \neq \vec{\phi}^{m_1}$ and $\Xi_{[0,p]}\vec{x}_0 = \vec{\phi}$.
2. $\Xi_{[0,p+q]}\vec{x}_0 = \langle t \rangle$, where $q \leq t < (q+r)$ and $(t-q) = \varepsilon(0, \Xi_{(0,p]}\sigma_t(\chi))$. Moreover, $\Xi_{(t+p,t+p+W]}\chi = \vec{\phi}^{m_1}$ and $\Xi_{(p+t)}\chi \neq \vec{\phi}^{m_1}$.

⁷ The choice of values for p, q, r and W was made so as to satisfy the following inequalities, which we will need in the proof: $p < W, p > 2(q+r)$ and $q > \alpha$.

3. $\Xi_{[0,2p-(q+r)]}\vec{x}_0 = \langle t_x, t_y \rangle$ with $(p - (q + r)) \leq t_x \leq (p - q) \leq t_y = p$. Also, for all $t' \in [0, p]$, $\Xi_{(t', t'+W)}\chi \neq \vec{\phi}^{m_1}$.
4. $\Xi_{[0,2p-r]}\vec{x}_0 = \langle t, t_x, t_y \rangle$ with $q \leq t \leq (q + r) < (p - r) \leq t_x \leq p \leq t_y = p + t$ and $(t - q) = \varepsilon((t_y - t_x - q), \Xi_{(0,p)}\sigma_t(\chi))$. Furthermore, for all $t' \in [0, p+t]$, $\Xi_{(t', t'+W)}\chi \neq \vec{\phi}^{m_1}$.

where $\varepsilon : [0, r] \times \bar{\mathcal{S}}_{(0,p)}^{m_1} \rightarrow [0, r]$ is as defined below.

For convenience, we define an operator $\iota_j^k : [0, 1) \rightarrow [0, 1)$, for $j, k \in \mathbb{Z}^+$, that constructs a new number obtained by concatenating every i^{th} digit of a given number, where $i \equiv j \pmod k$. More formally, for $x \in [0, 1)$, $\iota_j^k(x) = \sum_{i=1}^{\infty} ((\lfloor x \times 10^{j+(i-1)k} \rfloor - 10 \lfloor x \times 10^{j+(i-1)k-1} \rfloor) \times 10^{-i})$.

Also, we define another operator $\zeta_k : [0, 1)^k \rightarrow [0, 1)$, for $k \in \mathbb{Z}^+$ which ‘‘interleaves’’ the digits of k given numbers in order to produce a new number. More formally, for $x_0, x_1, \dots, x_{k-1} \in [0, 1)$, $\zeta_k(x_0, x_1, \dots, x_{k-1}) = \sum_{i=0}^{\infty} ((\lfloor x_{k(\frac{i}{k} - \lfloor \frac{i}{k} \rfloor)} \times 10^{1+\lfloor \frac{i}{k} \rfloor} \rfloor - 10 \lfloor x_{k(\frac{i}{k} - \lfloor \frac{i}{k} \rfloor)} \times 10^{\lfloor \frac{i}{k} \rfloor} \rfloor) \times 10^{-(i+1)})$.

Let d be the largest integer so that, for all $x' \in [0, r]$, we have $x' \times 10^d < 1$. For $x' \in [0, r]$, let $x = x' \times 10^d$. For $\chi \in \bar{\mathcal{S}}_{(0,p)}^{m_1}$, define⁸ $\varepsilon(x', \chi) = 10^{-d} \times \zeta_{m_1}(\varepsilon_0(\iota_1^{m_1}(x), \Pi_1(\chi)), \varepsilon_0(\iota_2^{m_1}(x), \Pi_2(\chi)), \dots, \varepsilon_0(\iota_{m_1}^{m_1}(x), \Pi_{m_1}(\chi)))$, where $\varepsilon_0 : [0, 1) \times \bar{\mathcal{S}}_{(0,p)} \rightarrow [0, 1)$ is as defined below.

Let $n \in [0, 1)$ and $\vec{x} \in \bar{\mathcal{S}}_{(0,p)}$. Furthermore, let $c = \iota_1^2(n)$ and $s = \iota_2^2(n)$. Let $\vec{x} = \langle x^1, x^2, \dots, x^k \rangle$. We have $0 \leq k \leq 8$, because $p = 8\alpha$. Also, since $p = 8r$, we have $x^i \times 10^{d-1} < 1$, for $1 \leq i \leq k$. Let $s' = \zeta_{k+1}(x^1 \times 10^{d-1}, x^2 \times 10^{d-1}, \dots, x^k \times 10^{d-1}, s)$. If $c = 0$, then let $c' = \frac{k}{10} + 0.09$ else let $c' = \frac{k}{10} + \frac{c}{10}$. Finally, define $\varepsilon_0(n, \vec{x}) = \zeta_2(c', s')$.

Next, we show that \mathbb{I} satisfies all the axioms of the neuron.

It is immediate that \mathbb{I} satisfies Axiom (1), since all output spikes in the above construction are atleast q seconds apart, and $q = 2\alpha$.

⁸ Recall that the ‘‘projection operator for spike configurations’’ is defined as $\Pi_i(\chi) = \vec{x}_i$, for $1 \leq i \leq m$, where $\chi = \langle \vec{x}_1, \vec{x}_2, \dots, \vec{x}_m \rangle$.

We now prove that \mathbb{I} satisfies Axiom (2). Let $\chi' \in \bar{\mathcal{S}}_{[0, \gamma_1]}^{m_1}$ and $\vec{x}'_0 \in \bar{\mathcal{S}}_{[0, \rho_1]}$ be arbitrary. If $P_1(\chi', \vec{x}'_0) = \tau_1$, then it is immediate from the construction that $P_1(\chi', \vec{\phi}) = \tau_1$ which satisfies Axiom (2). On the contrary, if $P_1(\chi', \vec{x}'_0) \neq \tau_1$, from the construction of \mathbb{I} , we have $P_1(\chi', \vec{x}'_0) = 0$. Also, from the construction we have either $P_1(\chi', \vec{\phi}) = 0$ or $P_1(\chi', \vec{\phi}) = \tau_1$. Axiom (2) is satisfied in either case.

Also, \mathbb{I} satisfies Axiom (3), since it is clear that $\chi = \vec{\phi}^{m_1}$ does not satisfy any of the conditions enumerated above. We therefore have $P_1(\vec{\phi}^{m_1}, \vec{\phi}) = 0$.

Finally, we show that there exists a $U \in \mathbb{R}^+$ so that for all $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, we have $\Xi_{[0, U]}(\sigma_{t_1}(\mathcal{T}_I(\chi_1) \sqcup \chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2}(\mathcal{T}_I(\chi_2) \sqcup \chi_2))$, where $\mathcal{T}_I : \mathcal{F}_m \rightarrow \mathcal{S}$ such that for each $\chi \in \mathcal{F}_m$, $\mathcal{T}_I(\chi)$ is consistent with χ with respect to \mathbb{I} . From Proposition 5, it follows that there exist $V_1, V_2 \in \mathbb{R}^+$ so that $\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[0, V_2]}(\sigma_{t_2}(\chi_2))$. Let $U = p + q + r$. If $\Xi_{[0, U]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2}(\chi_2))$, it is immediate that $\Xi_{[0, U]}(\sigma_{t_1}(\mathcal{T}_I(\chi_1) \sqcup \chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2}(\mathcal{T}_I(\chi_2) \sqcup \chi_2))$. It therefore suffices to prove that if $\Xi_{[U, V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[U, V_2]}(\sigma_{t_2}(\chi_2))$ then $\Xi_{[0, U]}(\sigma_{t_1} \mathcal{T}_I(\chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2} \mathcal{T}_I(\chi_2))$. Proposition 5 implies that $\Xi_{(V_1, V_1 + w]}(\sigma_{t_1}(\chi_1)) = \vec{\phi}^m$ and $\Xi_{V_1}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$. Therefore, by Case (1) of the construction, $\Xi_{(V_1 - p)} \sigma_{t_1} \mathcal{T}_I(\chi_1) = \langle V_1 - p \rangle$. Moreover, since Proposition 5 implies that for all $t'_1 \in [0, V_1)$, $\Xi_{(t'_1, t'_1 + w]}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$, from Case (3) of the construction, we have that for every $k \in \mathbb{Z}^+$ with $V_1 - kp > 0$, $\Xi_{(V_1 - kp)} \sigma_{t_1} \mathcal{T}_I(\chi_1) = \langle V_1 - kp \rangle$. Let k_1 be⁹ the smallest positive integer, so that $V_1 - k_1 p \leq U$. From the previous arguments, we have $\Xi_{(V_1 - k_1 p)} \sigma_{t_1} \mathcal{T}_I(\chi_1) = \langle V_1 - k_1 p \rangle$. Also, it is easy to see that $V_1 - k_1 p \geq (q + r)$. Let k_2 be similarly defined with respect to χ_2 so that $\Xi_{(V_2 - k_2 p)} \sigma_{t_2} \mathcal{T}_I(\chi_2) = \langle V_2 - k_2 p \rangle$ and $V_2 - k_2 p \leq U$. Now, there are two cases:

1. If $V_1 - k_1 p \neq V_2 - k_2 p$, we now show that $\Xi_{[0, U]}(\sigma_{t_1} \mathcal{T}_I(\chi_1)) \neq \Xi_{[0, U]}(\sigma_{t_2} \mathcal{T}_I(\chi_2))$, which is the required result. Assume, without loss of generality, that $V_1 - k_1 p < V_2 - k_2 p$. If these two quantities are less than $p - r$ apart, we have $\Xi_{[0, U]}(\sigma_{t_1} \mathcal{T}_I(\chi_1)) \neq$

⁹ k_1 exists because $U > p$.

$\Xi_{[0,U]}(\sigma_{t_2} \mathcal{T}_1(\chi_2))$, because by Case (4) of the construction $\mathcal{T}_1(\chi_1)$ has a spike in the interval $[V_1 - k_1p - (q + r), V_1 - k_1p - q]$ and by Case (3) of the construction, $\mathcal{T}_1(\chi_2)$ has no spike in the interval $(V_2 - k_2p, V_2 - k_2p + p - (q + r)]$. In other words, the spike following the one at $V_1 - k_1p$ in $\mathcal{T}_1(\chi_1)$ has no counterpart in $\mathcal{T}_1(\chi_2)$. On the other hand, if they are less than p apart but at most $p - r$ apart, by similar arguments, it is easy to show that the spike at $V_2 - k_2p$ in $\mathcal{T}_1(\chi_2)$ has no counterpart in $\mathcal{T}_1(\chi_1)$. Finally, if they are atleast p apart, then k_2 does not satisfy the property that it is the smallest positive integer, so that $V_2 - k_2p \leq U$, which is a contradiction.

2. On the contrary, consider the case when $V_1 - k_1p = V_2 - k_2p$. We have two cases:

(a) Suppose $k_1 \neq k_2$. Let t'_1 be the largest positive integer so that $\Xi_{t'_1} \sigma_{t'_1} \mathcal{T}_1(\chi_1) = \langle t'_1 \rangle$ and $t'_1 < V_1 - k_1p$. From Case (4) of the construction, we have that $q \leq (V_1 - k_1p) - t'_1 \leq q + r$. Let t'_2 be defined likewise, with respect to χ_2 . Further, let $n'_1 = (V_1 - k_1p) - t'_1 - q$ and $n'_2 = (V_2 - k_2p) - t'_2 - q$ and $n_1 = n'_1 \times 10^d$ and $n_2 = n'_2 \times 10^d$. Since $k_1 \neq k_2$, it is straightforward to verify that for all j with $1 \leq j \leq m_1$, $\iota_1^2(\iota_j^{m_1}(n_1)) \neq \iota_1^2(\iota_j^{m_1}(n_2))$, for the former number has 9 in the $(k_1 + 1)^{th}$ decimal place, while the latter number does in the $(k_2 + 1)^{th}$ decimal place and not in the $(k_1 + 1)^{th}$ decimal place since $k_1 \neq k_2$. Therefore, $n_1 \neq n_2$ and consequently $t'_1 \neq t'_2$ which gives us $\Xi_{[0,U]}(\sigma_{t'_1} \mathcal{T}_1(\chi_1)) \neq \Xi_{[0,U]}(\sigma_{t'_2} \mathcal{T}_1(\chi_2))$, which is the required result.

(b) On the other hand, suppose $k_1 = k_2$. Again, we have two cases:

i. Suppose, there exists a j with $1 \leq j \leq m_1$ and a $k' \leq k_1$, so that

$\Xi_{(V_1 - k'p, V_1 - (k' - 1)p]} \Pi_j(\sigma_{t_1}(\chi_1))$ has a different number of spikes when compared to $\Xi_{(V_2 - k'p, V_2 - (k' - 1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$. Let n_1, n_2 be defined as before. It is straightforward to verify that $\iota_1^2(\iota_j^{m_1}(n_1)) \neq \iota_1^2(\iota_j^{m_1}(n_2))$,

because they differ in the $(k_1 - k' + 1)^{th}$ decimal place¹⁰. Therefore,

$$\Xi_{[0,U]}(\sigma_{t_1} \mathcal{T}_1(\chi_1)) \neq \Xi_{[0,U]}(\sigma_{t_2} \mathcal{T}_1(\chi_2)).$$

ii. Now consider the case where for all j with $1 \leq j \leq m_1$ and $k' \leq k_1$, we

have $\Xi_{(V_1-k'p, V_1-(k'-1)p]} \Pi_j(\sigma_{t_1}(\chi_1))$ have the same number of spikes when compared to $\Xi_{(V_2-k'p, V_2-(k'-1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$. Now, by hypothesis, we have

$$\Xi_{[U, V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[U, V_2]}(\sigma_{t_2}(\chi_2)).$$

Therefore there must exist a $1 \leq j \leq m_1$ and $k' \leq k_1$, so that there is a point in time where one of the spike trains

$$\Xi_{(V_1-k'p, V_1-(k'-1)p]} \Pi_j(\sigma_{t_1}(\chi_1)) \text{ and } \Xi_{(V_2-k'p, V_2-(k'-1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$$

has a spike, while the other does not. Let t' be the latest time instant at which this is

$$\text{so. Also, assume without loss of generality that } \Xi_{(V_1-k'p, V_1-(k'-1)p]} \Pi_j(\sigma_{t_1}(\chi_1)) =$$

$$\langle x^1, \dots, x^q \rangle \text{ has a spike at time instant } t' \text{ while } \Xi_{(V_2-k'p, V_2-(k'-1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$$

does not. Let p be the number so that $t' = x^p$. Let n_1, n_2 be defined

as before. Also, for each h with $1 \leq h \leq k_1$, let r_h be the number of

spikes in $\Xi_{(V_1-hp, V_1-(h-1)p]} \Pi_j(\sigma_{t_1}(\chi_1))$. Each r_h can be determined from

$$n_1. \text{ Then, it is straightforward to verify}^{11} \text{ that } l_p^{r_{k'}} l_{r_{k'-1}}^{r_{k'-1}} \dots l_{r_1}^{r_1} l_2^2 l_j^{m_1} n_1 \neq$$

$$l_p^{r_{k'}} l_{r_{k'-1}}^{r_{k'-1}} \dots l_{r_1}^{r_1} l_2^2 l_j^{m_1} n_2. \text{ Therefore, } n_1 \neq n_2 \text{ and it follows that } \Xi_{[0,U]}(\sigma_{t_1} \mathcal{T}_1(\chi_1)) \neq$$

$$\Xi_{[0,U]}(\sigma_{t_2} \mathcal{T}_1(\chi_2)).$$

□

The preceding two propositions thus imply the following lemma.

Lemma 7. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, time-invariant and resettable, then it can be effected by an acyclic network of depth two.*

¹⁰ Which in n_1 and n_2 encodes the number of spikes in the interval $(V_2 - k'p, V_2 - (k' - 1)p]$ on the j^{th} spike train of χ_1 and χ_2 respectively.

¹¹ The expression on either side of the inequality is a real number that encodes for the p^{th} spike time in the spike trains $\Xi_{(V_1-k'p, V_1-(k'-1)p]} \Pi_j(\sigma_{t_1}(\chi_1))$ and $\Xi_{(V_2-k'p, V_2-(k'-1)p]} \Pi_j(\sigma_{t_2}(\chi_2))$ respectively.

Lemma 6 and 7 imply the following theorem. From this it follows that, with the abstract model currently under consideration, the set of all acyclic networks is not more complex than the set of acyclic networks of depth two.

Theorem 3. *If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ can be effected by an acyclic network, then it can be effected by an acyclic network of depth two.*

Corollary 2. *The set of all acyclic networks is not more complex than the set of acyclic networks of depth equal to two.*

6.4 Directions for Further Constraining the Abstract Model

The meticulous reader would have observed that the proof above exploited a number of phenomena that the abstract model allowed for, but are not particularly biologically well-founded. These are good candidates for new axioms that we hope will overcome the “depth barrier” that this section describes. One is that spike-times in the abstract model are real numbers, i.e. numbers with infinite precision. When noise is taken into account, this assumption is no longer true. The other easy observation is that we did not assume much about the membrane potential function here, which in nature, changes smoothly with time. And, finally, an assumption consistent with Dale’s principle, that each neuron has either an excitatory effect at all its postsynaptic neurons or an inhibitory effect could also help in this direction.

6.5 Some Auxiliary Propositions

Proposition 4. If $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is time-invariant, then $\mathcal{T}(\vec{\phi}^m) = \vec{\phi}$.

Proof. For the sake of contradiction, suppose $\mathcal{T}(\vec{\phi}^m) = \vec{x}_0$, where $\vec{x}_0 \neq \vec{\phi}$. That is, there exists a $t \in \mathbb{R}$ with $\Xi_t \vec{x}_0 = \langle t \rangle$. Let $\delta < \alpha$. Clearly, $\sigma_\delta(\vec{\phi}^m) = \vec{\phi}^m \in \mathcal{F}_m$. Since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is time-invariant, $\mathcal{T}(\sigma_\delta(\vec{\phi}^m)) = \sigma_\delta(\mathcal{T}(\vec{\phi}^m)) = \sigma_\delta(\vec{x}_0)$. Now, $\sigma_\delta(\vec{x}_0) \neq \vec{x}_0$ since $\Xi_{(t-\delta)} \sigma_\delta(\vec{x}_0) = \langle t - \delta \rangle$ whereas $\Xi_{(t-\delta)} \vec{x}_0 = \vec{\phi}$, for otherwise $\vec{x}_0 \notin \mathcal{S}$. This is a contradiction. Therefore, $\mathcal{T}(\vec{\phi}^m) = \vec{\phi}$. □

Corollary 3. *Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be causal, time-invariant and W -resettable, for some $W \in \mathbb{R}^+$. If $\chi \in \mathcal{F}_m$ has a gap in the interval $[t, t + W]$, then $\Xi_t \mathcal{T}(\chi) = \vec{\phi}$.*

Proof. Assume the hypothesis of the above statement. One readily sees that $\Xi_t \mathcal{T}(\chi) = \Xi_{[t,\infty)} \Xi_{(-\infty,t]} \mathcal{T}(\chi)$. Now, since χ has a gap in the interval $(t, t + W]$ and $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is W -resettable, we have $\Xi_{[t,\infty)} \Xi_{(-\infty,t]} \mathcal{T}(\chi) = \Xi_{[t,\infty)} \mathcal{T}(\Xi_{(-\infty,t]} \chi)$. Further, since $\Xi_t \chi = \vec{\phi}^m$, we have $\Xi_{[t,\infty)} \Xi_{(-\infty,t]} \chi = \Xi_{[t,\infty)} \vec{\phi}^m$. Therefore, since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, it follows that $\Xi_{[t,\infty)} \mathcal{T}(\Xi_{(-\infty,t]} \chi) = \Xi_{[t,\infty)} \mathcal{T}(\vec{\phi}^m) = \vec{\phi}$, with the last equality following from the previous proposition. \square

Proposition 5. Let $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ be causal, time-invariant and W' -resettable, for some $W' \in \mathbb{R}^+$. Then for all $W \in \mathbb{R}^+$ with $W \geq W'$, $t_1, t_2 \in \mathbb{R}$ and $\chi_1, \chi_2 \in \mathcal{F}_m$ with $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) \neq \Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2))$, there exist $V_1, V_2 \in \mathbb{R}^+$ so that the following are true.

1. $\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[0, V_2]}(\sigma_{t_2}(\chi_2))$
2. $\Xi_{(V_1, V_1 + W]}(\sigma_{t_1}(\chi_1)) = \vec{\phi}^m$, $\Xi_{V_1}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$ and $\Xi_{(V_2, V_2 + W]}(\sigma_{t_2}(\chi_2)) = \vec{\phi}^m$, $\Xi_{V_2}(\sigma_{t_2}(\chi_2)) \neq \vec{\phi}^m$
3. For all $t'_1 \in [0, V_1)$, $\Xi_{(t'_1, t'_1 + W]}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$ and for all $t'_2 \in [0, V_2)$, $\Xi_{(t'_2, t'_2 + W]}(\sigma_{t_2}(\chi_2)) \neq \vec{\phi}^m$.

Proof. Since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is causal, we have $\Xi_{[t_1, \infty)} \mathcal{T}(\chi_1) = \Xi_{[t_1, \infty)} \mathcal{T}(\Xi_{[t_1, \infty)} \chi_1)$. This implies $\sigma_{t_1}(\Xi_{[t_1, \infty)} \mathcal{T}(\chi_1)) = \sigma_{t_1}(\Xi_{[t_1, \infty)} \mathcal{T}(\Xi_{[t_1, \infty)} \chi_1))$ which gives us $\Xi_{[0, \infty)} \sigma_{t_1}(\mathcal{T}(\chi_1)) = \Xi_{[0, \infty)} \sigma_{t_1}(\mathcal{T}(\Xi_{[t_1, \infty)} \chi_1))$. Since $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is time-invariant and $\sigma_{t_1}(\Xi_{[t_1, \infty)} \chi_1) = \Xi_{[0, \infty)} \sigma_{t_1}(\chi_1) \in \mathcal{F}_m$, we have $\Xi_{[0, \infty)} \sigma_{t_1}(\mathcal{T}(\Xi_{[t_1, \infty)} \chi_1)) = \Xi_{[0, \infty)} \mathcal{T}(\Xi_{[0, \infty)} \sigma_{t_1}(\chi_1))$. In short, $\Xi_{[0, \infty)} \sigma_{t_1}(\mathcal{T}(\chi_1)) = \Xi_{[0, \infty)} \mathcal{T}(\Xi_{[0, \infty)} \sigma_{t_1}(\chi_1))$ which implies $\Xi_0 \sigma_{t_1}(\mathcal{T}(\chi_1)) = \Xi_0 \mathcal{T}(\Xi_{[0, \infty)} \sigma_{t_1}(\chi_1))$. Similarly, $\Xi_0 \sigma_{t_2}(\mathcal{T}(\chi_2)) = \Xi_0 \mathcal{T}(\Xi_{[0, \infty)} \sigma_{t_2}(\chi_2))$. Therefore, it follows from the hypothesis that $\Xi_0 \mathcal{T}(\Xi_{[0, \infty)}(\sigma_{t_1}(\chi_1))) \neq \Xi_0 \mathcal{T}(\Xi_{[0, \infty)}(\sigma_{t_2}(\chi_2)))$.

Let $V_1, V_2 \in \mathbb{R}^+$ be the smallest positive real numbers so that $\Xi_{[0, \infty)}(\sigma_{t_1}(\chi_1))$ and $\Xi_{[0, \infty)}(\sigma_{t_2}(\chi_2))$ have gaps in the intervals $(V_1, V_1 + W]$ and $(V_2, V_2 + W]$ respectively. That such V_1, V_2 exist follows from the fact that $\chi_1, \chi_2 \in \mathcal{F}_m$. Since, $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is W' -resettable, it is also W -resettable for $W \geq W'$. It therefore follows that $\Xi_{(-\infty, V_1]} \mathcal{T}(\Xi_{[0, \infty)}(\sigma_{t_1}(\chi_1))) = \mathcal{T}(\Xi_{(-\infty, V_1]} \Xi_{[0, \infty)}(\sigma_{t_1}(\chi_1)))$ which equals $\mathcal{T}(\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1)))$. This implies that $\Xi_0 \Xi_{(-\infty, V_1]} \mathcal{T}(\Xi_{[0, \infty)}(\sigma_{t_1}(\chi_1))) = \Xi_0 \mathcal{T}(\Xi_{[0, V_1]}(\sigma_{t_1}(\chi_1)))$ due to which we

have $\Xi_0 \mathcal{T}(\Xi_{[0,\infty)}(\sigma_{t_1}(\chi_1))) = \Xi_0 \mathcal{T}(\Xi_{[0,V_1]}(\sigma_{t_1}(\chi_1)))$. Likewise, $\Xi_0 \mathcal{T}(\Xi_{[0,\infty)}(\sigma_{t_2}(\chi_2))) = \Xi_0 \mathcal{T}(\Xi_{[0,V_2]}(\sigma_{t_2}(\chi_2)))$. We therefore have $\Xi_0 \mathcal{T}(\Xi_{[0,V_1]}(\sigma_{t_1}(\chi_1))) \neq \Xi_0 \mathcal{T}(\Xi_{[0,V_2]}(\sigma_{t_2}(\chi_2)))$. This readily implies $\Xi_{[0,V_1]}(\sigma_{t_1}(\chi_1)) \neq \Xi_{[0,V_2]}(\sigma_{t_2}(\chi_2))$ and, from the construction, it follows that $\Xi_{(V_1,V_1+W]}(\sigma_{t_1}(\chi_1)) = \vec{\phi}^m$, $\Xi_{V_1}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$ and $\Xi_{(V_2,V_2+W]}(\sigma_{t_2}(\chi_2)) = \vec{\phi}^m$, $\Xi_{V_2}(\sigma_{t_2}(\chi_2)) \neq \vec{\phi}^m$, for otherwise V_1 or V_2 would not be the smallest choice of numbers with the said property. Furthermore, for the same reasons, for all $t'_1 \in [0, V_1)$, $\Xi_{(t'_1,t'_1+W]}(\sigma_{t_1}(\chi_1)) \neq \vec{\phi}^m$ and for all $t'_2 \in [0, V_2)$, $\Xi_{(t'_2,t'_2+W]}(\sigma_{t_2}(\chi_2)) \neq \vec{\phi}^m$. □

CHAPTER 7 COMPLEXITY CLASSES, TRANSFORMATION HIERARCHIES AND HIERARCHY CLASSES

In this section, we develop some additional theoretical tools useful in proving complexity results. The idea is to study the space of all possible transformations of the form $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ that map spike configurations of order m which satisfy the Flush criterion to output spike trains. Specifically, we are interested in asking how the subset of transformations spanned by certain sets of networks are related to subsets of the space that can be characterized by specific mathematical properties of transformations. While the former types of subsets are related to networks of neurons, the latter type are related to transformations alone and mathematical properties thereof.

More concretely, using mathematical properties of the transformations (i.e. without any reference to networks), we identify a sequence of subsets of this space, with each subset contained in the subsequent one in the sequence. We then attempt to relate sets in this “hierarchy” of subsets to sets of transformations spanned by specific network architectures. We call this sequence of subsets a “Transformation Hierarchy” and subsets of transformations spanned by specific network architectures to be the “Complexity Classes” of those architectures.

There are atleast two reasons for trying to relate complexity classes to transformation hierarchies. First, complexity classes themselves seem to be hard to characterize succinctly in terms of properties of transformations they contain. Instead we try to understand complexity classes of network architectures relative to these sets in the hierarchy which are easier to characterize using mathematical properties of transformations. We do this by finding the “smallest” set in the hierarchy that contains the complexity class in question as a subset. This set is called the “Hierarchy Class” of the architecture with respect to the hierarchy in question. Figure 7-1 provides an illustration. The second reason for this approach is that it provides us another, and possibly more wholesale, way to prove that one set of networks is more complex than another.

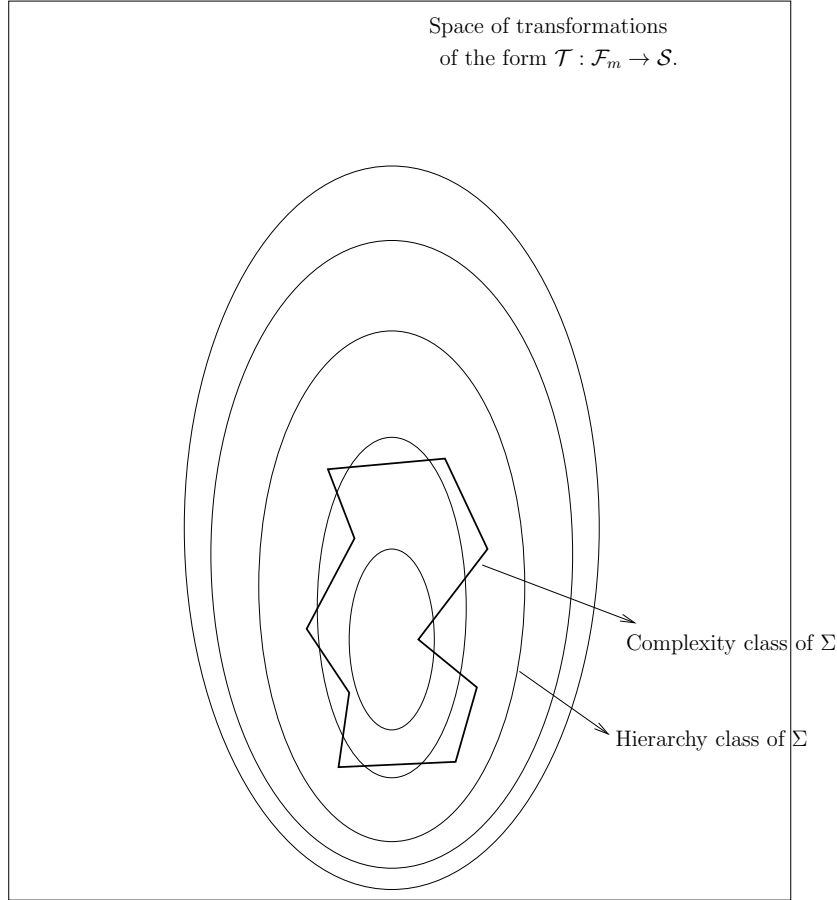


Figure 7-1. Venn Diagram illustrating the basic notions of a transformation hierarchy, complexity class and hierarchy class.

These notions are made precise next. As the reader familiar with Theoretical Computer Science will undoubtedly observe, the approach here is reminiscent of the one in Computational Complexity Theory.

7.1 Definitions and Preliminaries

Let \mathfrak{F}_m be the space of all possible transformations of the form $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ that map spike configurations of order m which satisfy the Flush criterion to output spike trains. Each acyclic network of order m induces one such transformation. A set of networks of order m therefore induces a class of such transformations, which we call the “complexity class” of that set.

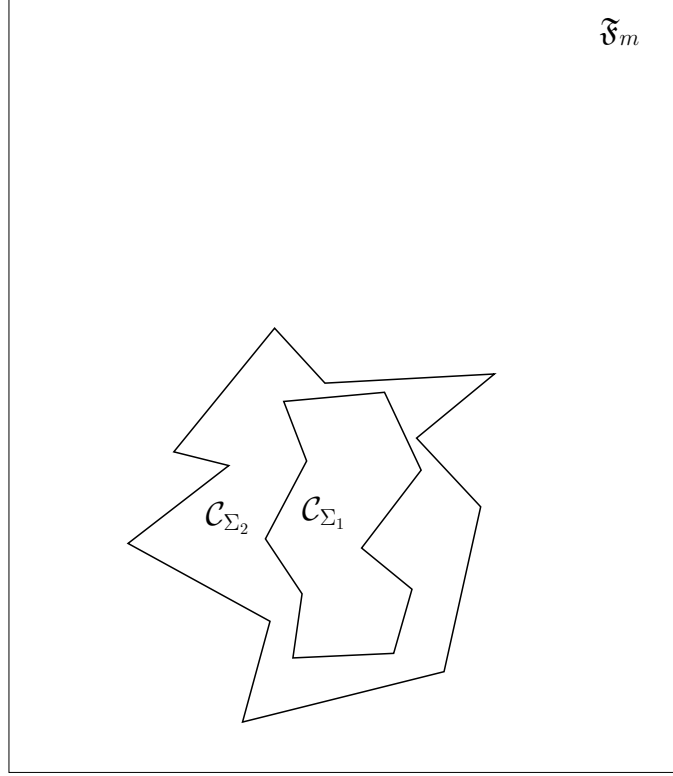


Figure 7-2. Venn Diagram showing complexity classes when one set of networks is more complex than another.

Definition 18 (Complexity Class). Let Σ be a set of acyclic networks of order m . The “complexity class” of Σ , \mathcal{C}_{Σ} , is defined to be the set $\bigcup_{\mathcal{N} \in \Sigma} \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m}$.

As is clear from Lemma 6, no complexity class spans the entire space \mathfrak{F}_m . Lemma 5 implies that questions of relative complexity of sets of networks can be posed in terms of questions about containment of their complexity classes. Figure 7-2 illustrates the situation and the next proposition formalizes it.

Lemma 8. *Let Σ_1 and Σ_2 be two sets of acyclic networks, each network being of order m , with $\Sigma_1 \subseteq \Sigma_2$. Further, let \mathcal{C}_{Σ_1} and \mathcal{C}_{Σ_2} be the corresponding complexity classes. Then, Σ_2 is more complex than Σ_1 if and only if $\mathcal{C}_{\Sigma_1} \subset \mathcal{C}_{\Sigma_2}$.*

Proof. First, suppose Σ_2 is more complex than Σ_1 . That $\mathcal{C}_{\Sigma_1} \subseteq \mathcal{C}_{\Sigma_2}$ follows immediately from the fact that $\Sigma_1 \subseteq \Sigma_2$. Lemma 5 implies that $\exists \mathcal{N}' \in \Sigma_2$ such that $\forall \mathcal{N} \in \Sigma_1, \mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \neq \mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m}$. That is, $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \in \mathcal{C}_{\Sigma_2}$ and $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \notin \mathcal{C}_{\Sigma_1}$. Since $\mathcal{C}_{\Sigma_1} \subseteq \mathcal{C}_{\Sigma_2}$, it follows that $\mathcal{C}_{\Sigma_1} \subset \mathcal{C}_{\Sigma_2}$.

To prove the other direction, assume $\mathcal{C}_{\Sigma_1} \subset \mathcal{C}_{\Sigma_2}$. Therefore, $\exists \mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ such that $\mathcal{T} \in \mathcal{C}_{\Sigma_2}$ and $\mathcal{T} \notin \mathcal{C}_{\Sigma_1}$. By definition of \mathcal{C}_{Σ_2} , $\exists \mathcal{N} \in \Sigma_2$, so that $\mathcal{T}_{\mathcal{N}}|_{\mathcal{F}_m} = \mathcal{T}$. Since $\mathcal{T} \notin \mathcal{C}_{\Sigma_1}$, the definition of \mathcal{C}_{Σ_1} implies that $\forall \mathcal{N}' \in \Sigma_1$, $\mathcal{T}_{\mathcal{N}'}|_{\mathcal{F}_m} \neq \mathcal{T}$. Therefore, Σ_2 is more complex than Σ_1 . \square

Next, we make precise the notion of a Transformation Hierarchy. Informally, a Transformation Hierarchy is a sequence of subsets of this space, with each subset contained in the subsequent one in the sequence.

Definition 19 (Transformation Hierarchy). A “Transformation Hierarchy” \mathcal{H} in \mathfrak{F}_m is a sequence of subsets $\langle H_1, H_2, \dots, H_i, \dots, \mathfrak{F}_m \rangle$ of \mathfrak{F}_m with $H_i \subset H_{i+1}, \forall i = 1, 2, \dots$

Note that the above definition exists independent of the existence of networks. That is, a hierarchy is defined only in terms of properties of transformations in its constituent sets. Figure 7-3 provides an illustration. The next definition provides a connection between sets of acyclic networks and hierarchies. Each set of acyclic networks Σ is associated with a specific set in the hierarchy called its hierarchy class, so that the hierarchy class is the smallest set in the hierarchy that contains the complexity class of Σ .

Definition 20 (Hierarchy Class). Let Σ be a set of acyclic networks, each of order m , and let $\mathcal{H} = \langle H_1, H_2, \dots, H_i, \dots, \mathfrak{F}_m \rangle$ be a transformation hierarchy in \mathfrak{F}_m . The “Hierarchy Class” H_Σ of Σ in \mathcal{H} is the set H_i with $\mathcal{C}_\Sigma \subseteq H_i$ and $\mathcal{C}_\Sigma \not\subseteq H_{i-1}$, if such an H_i exists and \mathfrak{F}_m otherwise.

Note that the Hierarchy class of a set of networks is well-defined. This is because $\mathcal{C}_\Sigma \subset \mathfrak{F}_m$, so there is atleast one set in the hierarchy that contains every complexity class. Also, the sets in the hierarchy are well-ordered¹ due to which every collection of sets from the hierarchy (which contain \mathcal{C}_Σ) has a smallest set.

¹ The ordering is by set inclusion.

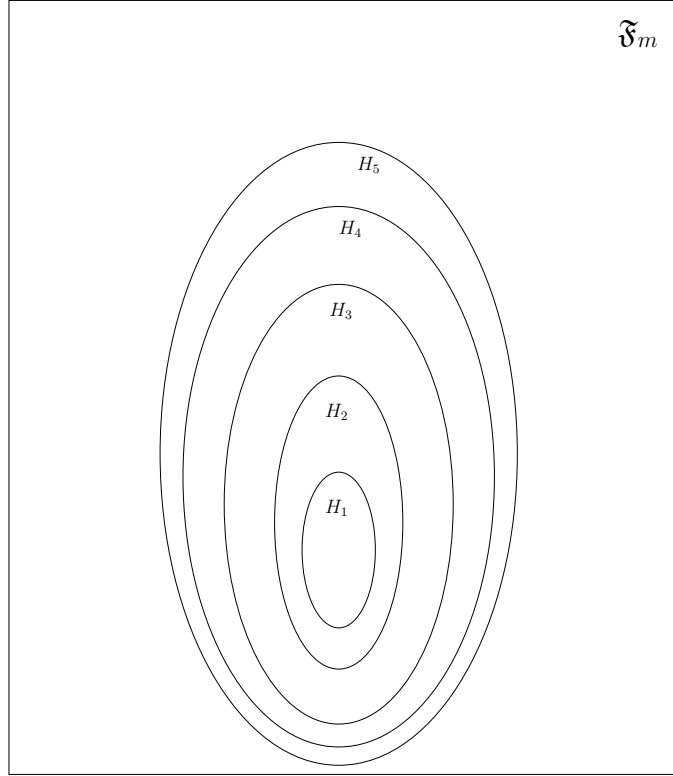


Figure 7-3. Venn Diagram illustrating a transformation hierarchy.

The above definitions allow us to create a variety of hierarchies based on specific properties of transformations. If we can then say something about the hierarchy classes of specific architectures in each hierarchy, it enables us to get a better understanding of various aspects of the transformations effected by networks with this architecture.

Even if we cannot identify the hierarchy class of a given architecture in a hierarchy, proving bounds² on them might give us some insight. Also, as alluded to before, these bounds can be used to establish that one set of networks is more complex than another.

A set in a hierarchy is an “upper bound” on a hierarchy class if it contains the hierarchy class as a subset. Likewise, a set in a hierarchy is a “lower bound” on a hierarchy class if the hierarchy class contains the set as a subset.

² Again, formally, the bounds are with respect to the partial ordering induced by set inclusion.

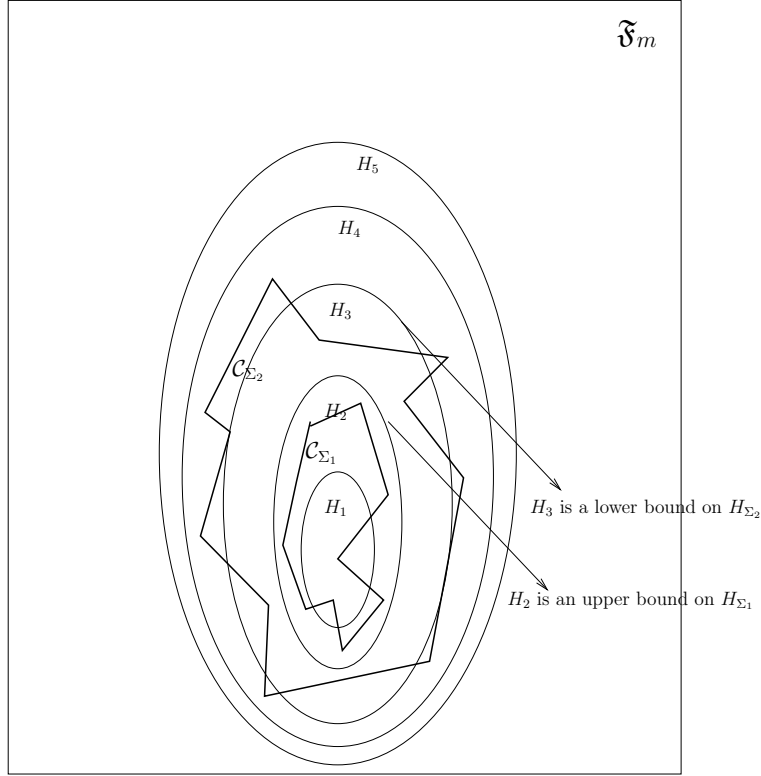


Figure 7-4. Venn Diagram demonstrating how upper bounds and lower bounds on hierarchy classes in a transformation hierarchy can be used to establish complexity results.

Bounds on hierarchy classes of specific sets of networks can be used to establish complexity results. If there are two sets of networks, the first contained in the second and if an upper bound on the hierarchy class of the first set is “smaller” than a lower bound on the hierarchy class of the second (with respect to the same hierarchy), then the second set is more complex than the first. Figure 7-4 gives a picture. Note that this is just a sufficient condition, not a necessary one, for one set to be more complex than the other. Depending on the hierarchy in question, it is possible that both sets have the same hierarchy class, yet one is more complex than the other. The next lemma formalizes the above observations.

Lemma 9. *Let Σ_1 and Σ_2 be two sets of acyclic networks, each comprising networks of order m , with $\Sigma_1 \subseteq \Sigma_2$. Furthermore, let H_{Σ_1} and H_{Σ_2} be the corresponding hierarchy classes in a transformation hierarchy $\mathcal{H} = \langle H_1, H_2, \dots, H_i, \dots, \mathfrak{F}_m \rangle$ in \mathfrak{F}_m . Moreover, let*

H_u be an upper bound on H_{Σ_1} and H_l be a lower bound on H_{Σ_2} . If $H_u \subset H_l$, then Σ_2 is more complex than Σ_1 .

Proof. Let \mathcal{C}_{Σ_1} and \mathcal{C}_{Σ_2} be the complexity classes of Σ_1 and Σ_2 respectively. By hypothesis, $\mathcal{C}_{\Sigma_1} \subseteq H_u$ and $H_l \subseteq \mathcal{C}_{\Sigma_2}$. Since, $H_u \subset H_l$, we have $\mathcal{C}_{\Sigma_1} \subset \mathcal{C}_{\Sigma_2}$. From Lemma 8, it now follows that Σ_2 is more complex than Σ_1 . \square

Indeed, this suggests an economical way to prove complexity results, since the upper bounds and lower bounds could apply to several sets of networks.

In the next section, we apply these notions to explicitly construct a transformation hierarchy and prove some lower bounds for some architectures according to the hierarchy.

7.2 Lower Bounds on the Hierarchy Classes of some Architectures in a certain Transformation Hierarchy

In this section, we construct a specific sequence of subsets of \mathfrak{F}_m and show that they constitute a transformation hierarchy. Next, we establish some lower bounds on the hierarchy classes of some architectures in this hierarchy. To prove that a certain set in a hierarchy is a lower bound on a hierarchy class, it suffices to show a transformation that is not in the set, yet is in the complexity class in question.

We start off by defining a class of transformations parameterized by a positive integer. For the sake of exposition, we will start by defining a “First-order Transformation” which we will then generalize to a “ k^{th} -order Transformation”. We will then show that for all $j \geq i$ every i^{th} order transformation is also a j^{th} order transformation.

Intuitively, a first-order transformation has the flavor of an SRM₀ neuron model, in that each synapse has a “kernel” function such that effects of inputs spikes according to this kernel are summed over all input spikes across all synapses. The transformation prescribes an output spike if and only if this sum equals a certain “threshold”, which is a positive number.

Definition 21 (First-order Transformation). A transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be a “First-order Transformation” if there exists a $\tau \in \mathbb{R}^+$ and functions $f_j : \mathbb{R} \rightarrow \mathbb{R}$, for $1 \leq j \leq m$, so that for every $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$, we have $\Xi_t \mathcal{T}(\chi) = \langle t \rangle$ if and only if we have $\sum_{j=1}^m \sum_{i=1}^{l_j} f_j(x_j^i) = \tau$, where $\Xi_{[0,\infty)} \sigma_t(\chi) = \langle \vec{x}_1, \dots, \vec{x}_m \rangle$ with $\vec{x}_j = \langle x_j^1, x_j^2, \dots, x_j^{l_j} \rangle$, for $1 \leq j \leq m$.

Informally, a k^{th} -order transformation is a generalization of a first-order transformation with higher-dimensional kernel functions. Thus, a second-order transformation, for example, has functions that take every pair of spikes and add up their “effects”, in addition to “first-order” effects.

Definition 22 (k^{th} -order Transformation). A transformation $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$ is said to be a “ k^{th} -order Transformation” if there exists a $\tau \in \mathbb{R}^+$ and functions $f_{j_1} : \mathbb{R} \rightarrow \mathbb{R}$, $f_{j_1 j_2} : \mathbb{R}^2 \rightarrow \mathbb{R}, \dots, f_{j_1 j_2 \dots j_k} : \mathbb{R}^k \rightarrow \mathbb{R}$, with $1 \leq j_p \leq m$, where $1 \leq p \leq k$, so that for every $\chi \in \mathcal{F}_m$ and $t \in \mathbb{R}$, we have $\Xi_t \mathcal{T}(\chi) = \langle t \rangle$ if and only if we have $\sum_{j_1=1}^m \sum_{i_1=1}^{l_{j_1}} f_{j_1}(x_{j_1}^{i_1}) + \sum_{j_1=1}^m \sum_{j_2=1}^m \sum_{i_1=1}^{l_{j_1}} \sum_{i_2=1}^{l_{j_2}} f_{j_1 j_2}(x_{j_1}^{i_1}, x_{j_2}^{i_2}) + \dots + \sum_{j_1=1}^m \dots \sum_{j_k=1}^m \sum_{i_1=1}^{l_{j_1}} \dots \sum_{i_k=1}^{l_{j_k}} f_{j_1 \dots j_k}(x_{j_1}^{i_1}, \dots, x_{j_k}^{i_k}) = \tau$, where $\Xi_{[0,\infty)} \sigma_t(\chi) = \langle \vec{x}_1, \dots, \vec{x}_m \rangle$ with $\vec{x}_j = \langle x_j^1, x_j^2, \dots, x_j^{l_j} \rangle$, for $1 \leq j \leq m$.

For all $j \geq i$, every i^{th} -order transformation is also a j^{th} -order transformation. This can be seen by setting the value of all the functions whose domain has dimensionality greater than i to be zero everywhere. Therefore, for all $j \geq i$, the set of all i^{th} -order transformations is a subset of the set of all j^{th} -order transformations. This naturally induces a transformation hierarchy in \mathfrak{F}_m

Proposition 6. For $k \in \mathbb{Z}^+$, let $\mathcal{O}_m^k \subseteq \mathfrak{F}_m$ be the set of all k^{th} -order transformations of the form $\mathcal{T} : \mathcal{F}_m \rightarrow \mathcal{S}$. Then, $\mathcal{O}_m = \langle \mathcal{O}_m^1, \mathcal{O}_m^2, \dots, \mathfrak{F}_m \rangle$ is a transformation hierarchy in \mathfrak{F}_m .

For certain acyclic network architectures, we now establish some lower bounds on their hierarchy classes in the above-mentioned transformation hierarchy.

Theorem 4. Let Σ_2 be the set the set of all networks with the architecture of the network in Figure 7-5. Then \mathcal{O}_2^1 is a lower bound on the hierarchy class of Σ_2 in \mathcal{O}_2 .

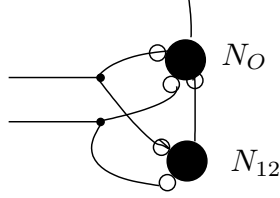


Figure 7-5. Diagram depicting architecture of networks in Σ_2 .

Proof. We prove that \mathcal{O}_2^1 is a lower bound on the hierarchy class of Σ_2 in \mathcal{O}_2 by showing a transformation that a network in Σ_2 can effect, but which lies outside \mathcal{O}_2^1 .

For the sake of brevity, we describe the salient responses of the neurons from which it is straightforward to construct an SRM_0 model for them. For the sake of contradiction assume that the transformation effected by the network is a first-order transformation. In Figure 7-5, neuron N_{12} is an inhibitory neuron and the neuron N_0 is an excitatory neuron. Both the input spike trains provide excitatory input to both neurons. We assume, for the sake of contradiction, that the potential of the output neuron N_0 can be written down as a first-order transformations. The argument is made on two input spikes that occurred t_1 and t_2 seconds ago in the first and second input spike train respectively. Consider the values of the functions $f_1(t_1)$ and $f_2(t_2)$, where $f_1(\cdot)$ and $f_2(\cdot)$ are component functions of the putative first-order function in the transformation. The neuron N_0 is set up so that it produces a spike now, if a spike happens either at t_1 alone or t_2 alone. Since the transformation is first-order, this gives us two equations $f_1(t_1) = \tau$ and $f_2(t_2) = \tau$. When there is a spike both at positions t_1 and t_2 , N_0 would reach threshold earlier. However, the occurrence of both these spikes causes N_{12} to spike. This, in turn, causes an inhibitory effect on the membrane potential of N_0 , which compensates for the extra excitation, so that it spikes exactly once, now. We therefore have the equation $f_1(t_1) + f_2(t_2) = \tau$. These three equations (in the two variables $f_1(t_1)$ and $f_2(t_2)$) form an inconsistent system of linear equations, and therefore $f_1(t_1)$ and $f_2(t_2)$ do not exist, contradicting our hypothesis. Therefore the transformation induced by the current network is not a

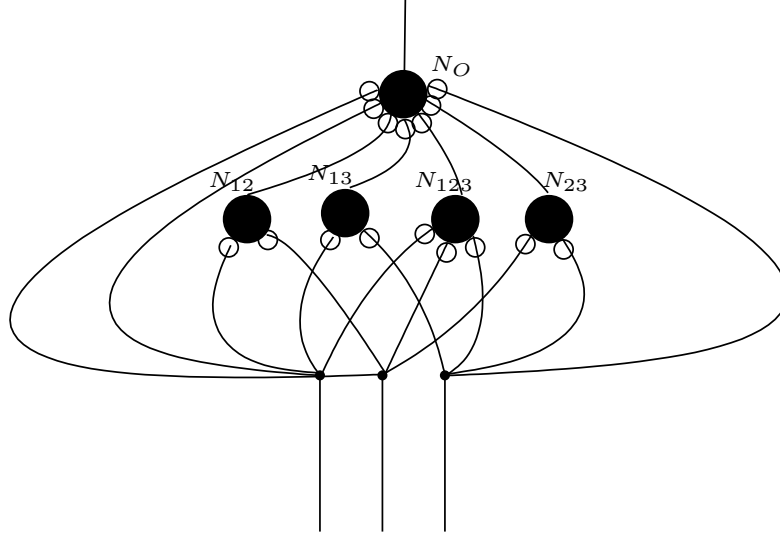


Figure 7-6. Diagram depicting architecture of networks in Σ_3 .

first-order transformation. Thus, \mathcal{O}_2^1 is a lower bound on the hierarchy class of Σ_2 in \mathcal{O}_2 . □

Next, we apply a similar strategy to derive a lower bound on the hierarchy class of another network architecture. It will then be clear how one can generalize the present technique.

Theorem 5. *Let Σ_3 be the set the set of all networks with the architecture of the network in Figure 7-6. Then \mathcal{O}_3^2 is a lower bound on the hierarchy class of Σ_3 in \mathcal{O}_3 .*

Proof. As before, we prove this by showing a transformation that a network in Σ_3 can effect, but which lies outside \mathcal{O}_3^2 . The argument is made on three spike positions t_1 , t_2 and t_3 in the past, and the values of the functions $f_1(t_1)$, $f_2(t_2)$, $f_3(t_3)$, $f_{12}(t_1, t_2)$, $f_{23}(t_2, t_3)$ and $f_{31}(t_3, t_1)$, which are component functions of the putative second-order transformation. Again, N_0 is set up so it spikes on each of the three individual spikes occurring alone. This gives us the equations $f_1(t_1) = \tau$, $f_2(t_2) = \tau$ and $f_3(t_3) = \tau$. N_{12} works exactly as in the previous example, and so do N_{23} and N_{31} , so as to make the output neuron spike whenever every pair of spikes occur. This gives us the equations $f_1(t_1) + f_2(t_2) + f_{12}(t_1, t_2) = \tau$, $f_2(t_2) + f_3(t_3) + f_{23}(t_2, t_3) = \tau$ and $f_3(t_3) + f_1(t_1) + f_{31}(t_3, t_1) = \tau$. Now, when spikes

occur simultaneously at all three times, the inhibition provided by N_{12} , N_{23} and N_{31} causes the membrane potential of N_0 to always stay below threshold. The neuron N_{123} now provides enough excitation to N_0 , in order to make it spike now. This gives us the equation $f_1(t_1) + f_2(t_2) + f_3(t_3) + f_{12}(t_1, t_2) + f_{23}(t_2, t_3) + f_{31}(t_3, t_1) = \tau$. It is straightforward to verify that this system of 7 equations in 6 variables is inconsistent. Therefore, \mathcal{O}_3^2 is a lower bound on the hierarchy class of Σ_3 in \mathcal{O}_3 . \square

It is straightforward to obtain similar results for higher-order transformations with this technique. While in the previous chapter, we, in effect, upper-bound the complexity class of the set of all transformations effected by an acyclic network, much work remains to be done in this direction.

CHAPTER 8 CONCLUSION AND FUTURE WORK

8.1 Contributions

We conclude by summarizing our contributions in this dissertation. We began by finding that acyclic networks cannot, in general, be viewed as spike-train to spike-train transformations. This pessimistic outlook was then surmounted by the observation that under spiking regimes usually observed in the brain, this view of acyclic networks as spike-train to spike-train transformations is in fact well-founded. Indeed, we rigorously formulated precise conditions which we proved from first principles. This significantly clarifies and formalizes the informal notion often advanced in the literature that networks of neurons are just entities that transform spike trains to spike trains. We then directed our attention to the question of, if network architecture constrains the type of transformations that these networks can do. To this end, we defined notions of relative complexity and derived a useful technical lemma that significantly eases the task of answering the aforementioned questions. Finally, we proved complexity results involving some classes of architectures that suggest the feasibility of this approach. Next, we showed that the current abstract model cannot reveal complexity results for networks beyond depth two. We established this by proving that every acyclic network has an equivalent acyclic network of depth two, so that both the networks effect exactly the same transformation. We then developed some additional theoretical tools by studying the space of all spike-train to spike-train transformations. We defined notions of transformational hierarchies in this space and related the set of transformations spanned by some network architectures to sets in a hierarchy.

Not only are the results themselves of theoretical significance, we believe that in these proofs lie the seeds of computation in the brain; that is, this forces us to directly address questions about which properties of neurons are crucial for which properties of networks and how local properties of neurons constrain global network behavior.

The current work also assumes added importance in light of recent efforts such as the Human Connectome Project, funded by the NIH, to map the neuronal circuitry of the human brain. Since also measuring the dynamical properties of each of the individual neurons appears to be out of reach of current experimental technology, the need for theoretical techniques agnostic to dynamical properties that can help analyze and interpret this data cannot be overemphasized. The current work offers promise of progress towards such a goal.

8.2 Directions for Future Work

Due to the results of Chapter 6, the first order of things is to explore new axioms to be added to the current abstract model, so that complexity results beyond those for networks of depth two emerge. We have several candidate axioms in mind. For example, the present model does not assume that the membrane potential changes smoothly with time, even though this would be a biologically reasonable assumption to make. The other option is to bring noise into the picture. This would also necessitate changing the definitions, for one would then have to make probabilistic statements rather than deterministic ones that we have considered here.

Once this is done and we have a large body of complexity results in our arsenal, we would like to investigate to what extent these results can afford us insights into the larger class of recurrent networks. Here the problem appears much harder and therefore more interesting. In particular, we are currently exploring ways of asking similar questions of recurrent networks too.

And, finally, once we have a rough sense of what the right questions are and what the answers might be like, we seek to return to square one, and begin once again with a view to treating biological phenomena which are not presently treated. Synaptic plasticity, for example would be a prime candidate to treat in the new model.

8.3 Epilogue

When we began work on this dissertation, the idea was to start with a highly emaciated abstract model of a neuron and see how much structure one could extract out of a small number of axioms. This was not just an exercise in mathematical vanity, but a consequence of the large variety of neurons and their multifarious abilities; indeed, the more one assumes about neurons, the smaller the class of biological neurons one covers. The initial attack led to the counterexample that even a single neuron cannot be consistently viewed as a spike-train to spike-train transformation. While this seemed to be a setback, we were fortunate to hit upon the idea that led to the Gap Lemma that showed that by considering a restricted subset of spike trains, which luckily was also biologically well-motivated, we could in fact redeem the intuitive view of neurons and indeed acyclic networks as a spike-train to spike-train transformations. That a need for such a biologically well-founded assumption emerged from the theory was very encouraging. This gave us a lot of faith in the existence of much structure which we then sought to uncover. In hindsight, although we show that there are limits to what we could infer from this emaciated model, that we could show so much structure, in our opinion, shows great promise in this type of axiomatic approach to questions about networks of neurons. While the questions we sought to answer here were relatively narrow, this experience leads us to believe that such an approach could be of value in asking other types of questions as well. We therefore end with a note of optimism and hope that the reader will join us in unearthing the marvels that lie ahead.

REFERENCES

- Banerjee, A. (2001a). On the phase-space dynamics of systems of spiking neurons. I: Model and experiments. *Neural Comp.*, *13*(1), 161–193.
- Banerjee, A. (2001b). *The Phase-Space Dynamics of Systems of Spiking Neurons*. Ph.D. thesis, Rutgers University.
- Banerjee, A. (2006). On the sensitive dependence on initial conditions of the dynamics of networks of spiking neurons. *J. Comput. Neurosci.*, *20*(3), 321–348.
- Bear, M. F., Connors, B. W., & Paradiso, M. A. (2007). *Neuroscience: Exploring the brain*. Baltimore, MD: Lippincott Williams & Wilkins, 3rd ed.
- Bohte, S., Kok, J., & La Poutre, H. (2002). Error-backpropagation in temporally encoded networks of spiking neurons. *Neurocomputing*, *48*(1-4), 17–37.
- Dayan, P., & Abbott, L. F. (2005). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: MIT Press.
- Gerstner, W., & Kistler, W. (2002). *Spiking neuron models: Single neurons, populations, plasticity*. New York, NY: Cambridge University Press.
- Goldman, D. E. (1943). Potential, Impedance and Rectification in Membranes. *J. Gen. Physiol.*, *27*, 37–60.
- Gutig, R., & Sompolinsky, H. (2006). The tempotron: a neuron that learns spike timing-based decisions. *Nat. Neurosci.*, *9*(3), 420–428.
- Hodgkin, A., & Huxley, A. (1952a). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (Lond.)*, *117*, 500–544.
- Hodgkin, A., & Huxley, A. (1952b). Currents carried by sodium and potassium ions through the membrane of the giant axon of *Loligo*. *J. Physiol. (Lond.)*, *116*, 449–472.
- Hodgkin, A., & Huxley, A. (1952c). The components of membrane conductance in the giant axon of *Loligo*. *J. Physiol. (Lond.)*, *116*, 473–496.
- Hodgkin, A., & Huxley, A. (1952d). The dual effect of membrane potential on sodium conductance in the giant axon of *Loligo*. *J. Physiol. (Lond.)*, *116*, 497–506.
- Hodgkin, A., A.L. Huxley, & Katz, B. (1952). Measurement of current-voltage relations in the membrane of the giant axon of *Loligo*. *J. Physiol. (Lond.)*, *116*, 424–448.
- Koch, C. (2004). *Biophysics of Computation: Information Processing in Single Neurons (Computational Neuroscience Series)*. New York, NY: Oxford University Press.
- Maass, W. (1995). Vapnik-Chervonenkis dimension of neural nets. In M. Arbib (Ed.) *The handbook of brain theory and neural networks*, (pp. 1000–1003). Cambridge, MA: MIT Press.

- Maass, W. (1996). Lower bounds for the computational power of networks of spiking neurons. *Neural Comp.*, 8(1), 1–40.
- Minsky, M., & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.
- Nirenberg, S., Carcieri, S., Jacobs, A., & Latham, P. (2001). Retinal ganglion cells act largely as independent encoders. *Nature*, 411(6838), 698–701.
- Poirazi, P., Brannon, T., & Mel, B. (2003). Pyramidal neuron as two-layer neural network. *Neuron*, 37(6), 989–999.
- Rall, W. (1960). Membrane potential transients and membrane time constant of motoneurons. *Exp. Neurol.*, 2(5), 503 – 532.
- Rieke, F., Warland, D., van Steveninck, R., & Bialek, W. (1997). *Spikes: exploring the neural code*. Cambridge, MA: MIT Press.
- Rosenblatt, F. (1988). *The perception: a probabilistic model for information storage and organization in the brain*, (pp. 89–114). Cambridge, MA: MIT Press.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). *Learning internal representations by error propagation*, (pp. 318–362). Cambridge, MA: MIT Press.
- Shepherd, G. (2004). *The synaptic organization of the brain*. New York, NY: Oxford University Press.
- Squire, L. R., Bloom, F. E., McConnell, S. K., Roberts, J. L., Spitzer, N. C., & Zigmond, M. J. (2003). *Fundamental neuroscience*. London: Academic Press, 2nd ed.
- Strehler, B., & Lestienne, R. (1986). Evidence on precise time-coded symbols and memory of patterns in monkey cortical neuronal spike trains. *Proc. Nat. Acad. Sci. USA*, 83(24), 9812.
- Tuckwell, H. C. (1988). *Introduction to Theoretical Neurobiology*. Cambridge studies in mathematical biology, 8. New York, NY: Cambridge University Press.

BIOGRAPHICAL SKETCH

Venkatakrishnan Ramaswamy was born in Bombay, India in 1982. After going through schooling in seven different schools, he received his undergraduate degree in Information Technology from the University of Madras, India in 2004, graduating First Class with Distinction. He graduates with a Ph.D. from the University of Florida in May 2011.