What Could Be Happening In The Frog's Eye

Yuan Yao

A Thesis
in
The Department
of
Computer Science and Software Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Computer Science at
Concordia University
Montreal, Quebec, Canada

September 2008

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-45357-5
Our file Notre référence
ISBN: 978-0-494-45357-5

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

What Could Be Happening In The Frog's Eye

Yuan Yao

Our topic comes out of a juxtaposition of two classic papers: the 1943 paper *A logical calculus of the ideas immanent in nervous activity* by McCulloch and Pitts and the 1959 paper *What the frog's eye tells the frog's brain* by Lettvin, Maturana, McCulloch, and Pitts.

The nervous system is made up of individual signaling elements, the *neurons.* McCulloch and Pitts proposed a simplified model of this system. There every neuron receives a number of signals, which are nonnegative real numbers, at time t; if a weighted sum of these signals exceeds a prescribed threshold, then the neuron sends signal 1 to other neurons at time $t+1$; else it sends signal 0.

Lettvin et al. discovered that the frog's optical nerve transmits to the brain information obtained by preprocessing the data in the retina by four separate operations: sustained contrast detection, net convexity detection, moving edge detection, and net dimming detection.

We speculate about McCulloch-Pitts networks that could implement the fourth of these operations, the detection of sudden dimming. To be biologically plausible, such networks must conform to what is known about frog's anatomy: its retina consists one million photoreceptors, three to four million other neurons, every neuron receives signals from fewer than $100,000$ neurons, and it sends signals to fewer than $10,000$ neurons.

We construct such networks that, within a narrow margin of error, detect dimming from time t to $t+1$ and respond at time $t+4$.

Acknowledgements

First and foremost, I am full of gratitude to my supervisor, Dr. Vašek Chvátal, for his insightful advice and endless support throughout my thesis study. Without his guidance and encouragement, this work cannot be done.

I would like to thank Dr. Markéta Vyskočilová for her professional advice and comments on the biological research part of my thesis. I would like to thank Dr. Ehsan Chiniforooshan for his advice that a single McCulloch-Pitts neuron can receive real numbers and integers at the same time. I would also like to thank Marie-Pierre Cossette for her help during my early studies in frog's anatomy.

Moreover, I would like to thank Dr. David Ford and Dr. Sudhir P. Mudur for their helpful comments on my thesis.

Finally, I would like to thank my family for their constant love and support.

Contents

List of Figures **vii**

1 Introduction **1**

2 Electrophysiology of Neurons **3**

2.1 Neurons 4

2.2 Spikes 5

2.3 Formal Spiking Neuron Models 9

2.3.1 Integrate-and-fire Model 9

2.3.2 Spike Response Model 11

3 McCulloch-Pitts Neural Networks **15**

3.1 The McCulloch-Pitts Neuron 16

3.2 The McCulloch-Pitts Neural Network 17

3.3 Threshold Circuits 19

4 What The Frog's Eye Tells The Frog's Brain **21**

4.1 Retina 21

4.2 Frog's Visual System 22

4.3 Net Dimming Detectors 24

5 Our Definition of Dimming **26**

6 Observation on Dimming Detectors **29**

6.1 Discrete Encodings of Real-valued Inputs 29

6.2 All Dimming Detectors Must be Approximate 30

6.3 Biological Constraints 31

6.4 Photoreceptors Must be Accessed in Groups 32

7 Our Dimming Detector **33**

7.1 The Parameters 33

7.2 The Objective 34

7.3 Stratified Networks and Pipelining 35

7.4 Construction 36

7.5 Analysis 40

7.6 Generalizations 43

  7.6.1 The Objective 43

  7.6.2 Construction 44

  7.6.3 Analysis 47

  7.6.4 Parameters 50

8 Conclusions and Future Research **52**

References **53**

Appendix A Approximate Bit Counting **55**

Appendix B Assumability **60**

List of Figures

1 Texture of the nervous system of man and the vertebrates. By Santiago Ramón y Cajal . . . 3

2 Structure of a single neuron (from website http://webschoolsolutions.com/patts/sytems/nervous) . . . 4

3 Curve of a single spike (from Gerstner and Kistler (2005)) . . . 6

4 Synapse cleft (from http://www.cassiopaea.org) . . . 7

5 One neuron receives signals from two neurons (from Gerstner and Kistler (2005)) . . . 9

6 Schematic diagram of the integrate-and-fire model (from Gerstner and Kistler (2005)) . . . 10

7 Schematic interpretation of the spike response model (from Gerstner and Kistler (2005)) . . . 12

8 Drawing of a section through the optic tectum of a sparrow (Santiago Ramon y Cajal, c.1900) . . . 21

9 Diagram of frog's visual System . . . 22

10 Condition 1 is satisfied but the image does not dim . . . 27

11 Overall structure of the approximate network . . . 39

12 Condition 1 unit fires or not based on the relative positions between $\alpha(t)$ and $\alpha(t+1)$ . . . 41

13 Dimming section detector fires or not based on the relative positions between $\alpha_i(t)$ and $\alpha_i(t+1)$ . . . 42

14 Condition 1 unit fires or not based on the relative positions between $\alpha(t)$ and $\alpha(t+1)$ . . . 49

15 Dimming section detector fires or not based on the relative positions between $\alpha_i(t)$ and $\alpha_i(t+1)$ 50

1 Introduction

Ramón y Cajal (1852 – 1934) was a Spanish histologist, physician, and Nobel laureate. He is considered to be one of the founders of neuroscience (http://en.wikipedia.org/wiki/Santiago_Ram%C3%B3n_y_Cajal). In particular, he developed some of the key conceptual insights and much of the early empirical support for the *neuron doctrine* — the principle that the nervous system is made up of individual signaling elements, the neurons, which contact one another only at specialized points of interaction, called synapse (Kandel 1991).

The signals are represented by complex changes in the electrical and chemical properties of the neurons. We survey these processes in Section 2.

The only reason we include this material in the thesis is to highlight the contrast between the complexity of signaling in the central nervous system and the simplicity of its model presented in Section 3.

The model was proposed in 1943 by Warren S. McCulloch, a neuroscientist, and Walter Pitts, a logician, in a paper entitled *A logical calculus of the ideas immanent in nervous activity.* This paper had a great influence on the development of *artificial neural networks* (mathematical or computational models of the networks of neurons that are present in the central nervous system).

Study of artificial neural networks can be divided into two distinct categories:

1. there are researchers who aim to construct better computer hardware and software by imitating the brain;
2. there are researchers who aim to better understand the brain by studying its mathematical models.

This thesis belongs to the second category.

Our topic comes out of a juxtaposition of the McCulloch-Pitts paper with another classic paper: *What the frog's eye tells the frog's brain*, published in 1959 by Jerome Lettvin, Humberto Maturana, Warren McCulloch, and Walter Pitts. These four authors found by experiments that, rather than transmitting raw data from the photoreceptors to the brain, the frog's optical nerve transmits to the brain information obtained by preprocessing the data in the retina by four separate operations: 1) sustained contrast detection, 2) net convexity detection, 3) moving edge detection, and 4) net dimming detection. Each operation has its result transmitted by a particular group of fibers. We quote from this paper in Section 4.

If the McCulloch-Pitts model is accurate, then the frog's retina must contain four neural networks built along the lines suggested by McCulloch and Pitts and computing the four functions reported by Lettvin et al.. Along these lines, we build a neural network which approximates one of the four functions reported by Lettvin et al..

Coaching the intuitive concept of dimming in mathematical terms is not completely straightforward. In Section 5 we propose a definition of dimming intended to capture the intuitive meaning and to safeguard against the interference of noise.

After making a few preliminary observations on dimming detector in Section 6, we construct in Section 7 a McCulloch-Pitts network that detects dimming as defined in Section 5. The size of the network is consistent with available data concerning the frog's retina.

2 Electrophysiology of Neurons

This section is adapted from Gerstner and Kistler (2005)

The elementary processing units in the central nervous system are *neurons*, which are connected to each other in an intricate pattern. A tiny portion of such a network of neurons is sketched in Figure 1 by Santiago Ramón y Cajal.

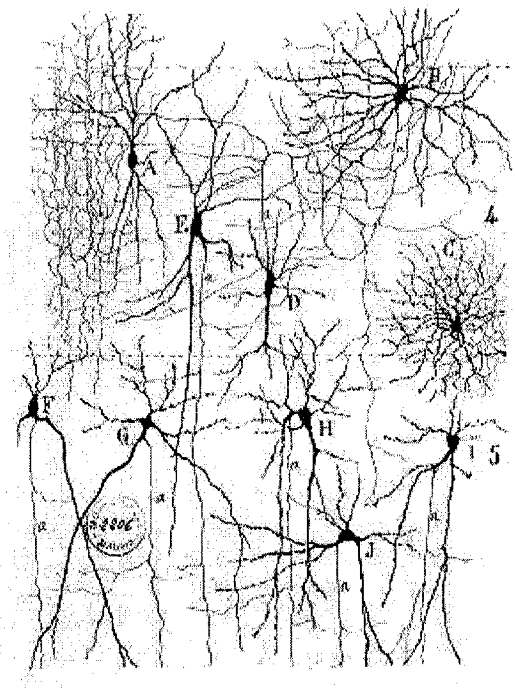

Figure 1: Texture of the nervous system of man and the vertebrates. By Santiago Ramón y Cajal

Figure 1 gives a glimpse of the network of the neurons in the cortex. In reality, cortical neurons and their connections are packed into a dense network with more than 10^4 cell bodies per cubic millimeter.

2.1 Neurons

This section is adapted from Kandel (1991) and Gerstner and Kistler (2005).

A typical neuron can be divided into three functionally distinct parts, called *dendrites*, *soma*, and *axon.* They are shown in Figure 2.

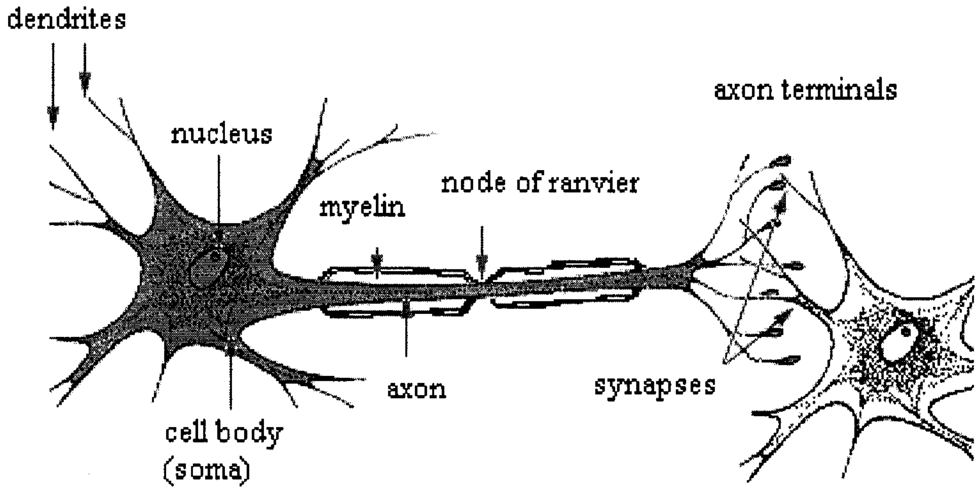

Figure 2: Structure of a single neuron (from website http://webschoolsolutions.com/patts/sytems/nervous)

Typically, the dendrites are the "input devices" which collect signals from other neurons and transmit them to the soma; the soma is the "central processor" that produces a signal or not depending on the input; the axon is the "output device" which delivers the signal generated by the neuron to other neurons. The site between two contact neurons is called *synapse.* When a neuron sends signals across a synapse, it is common to refer to the sending neuron as the *pre-synaptic cell* and to the receiving neuron as the *post-synaptic cell.* The tiny gap between pre- and post-synaptic cell membranes is called the *synaptic cleft.* Kandel (1991) says, talking about human central nervous system,

> The branches of the axon of one neuron may form synapses with as many as 1000 other

> neurons

on page 19, and

> A spinal motor cell, whose dendrites are moderate in both number and extent, receives about $10,000$ contacts — $2,000$ on the cell body and 8000 on the dendrites. The larger dendritic tree of Purkinje cell of the cerebellum receives approximately $150,000$ contacts!

on page 22.

The neurons of the brain can be classified functionally into three major groups: afferent, motor, and interneuronal. Afferent or sensory neurons carry information into the nervous system. Motor neurons carry commands to muscles and glands. Interneurons constitute by far the largest class and consist of all the remaining cells in the nervous system that are not specifically sensory or motor.

2.2 Spikes

This section is adapted from Parberry (1994) and Gerstner and Kistler (2005).

The neuron is bounded by a lipid membrane which is permeable to water, but impermeable to ions and various water-soluble molecules. This structure will maintain the concentration of internal ions different from the concentration of ions in the environment between neurons. The concentrations of sodium and potassium ions are particularly important. Due to the different concentrations of ions across the membrane, there is a difference of potential between the interior of the cell and its surroundings. We call this difference the *membrane potential.* A typical neuron is at rest with a constant *resting membrane potential* of about $-65mV$ to $-70mV$.

At the point of axon joining the soma, when the potential exceeds a threshold value of $-55mV$, the membrane will open sodium channels immediately and allow sodium to enter the membrane by osmotic pressure. The polarity of the membrane in that area reverses and rises to approximately $+60mV$. Then the potassium channels open, allowing potassium to leave the membrane also under osmotic pressure. Then the membrane potential decreases to about $-90mv$ and then returns gradually back to the resting state. The net results of the short electrical pulses are called *action potentials.* Observed action potentials are called *spikes.* The spikes have an amplitude of about $100mV$ and typically a duration of $1-2ms$. The curve of a spike is shown in Figure 3.

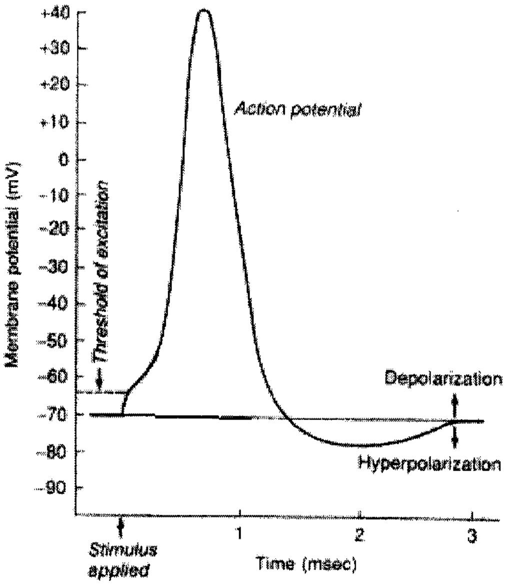

Figure 3: Curve of a single spike (from Gerstner and Kistler (2005))

The resulting action potential travels down the axon and arrives at a synapse. It triggers the release of neurotransmitters from the pre-synaptic terminal into the synaptic cleft. As soon as neurotransmitter molecules have reached the postsynaptic side, they will be detected by specialized receptors in the postsynaptic cell membrane and open special channels, so that ions can exchange between the interior of the cell and its surroundings. The ion exchanging alters the postsynaptic membrane potential. Thus, the chemical signal is translated into an electrical response. Figure 4 shows the chemical functions happening in synapse cleft.

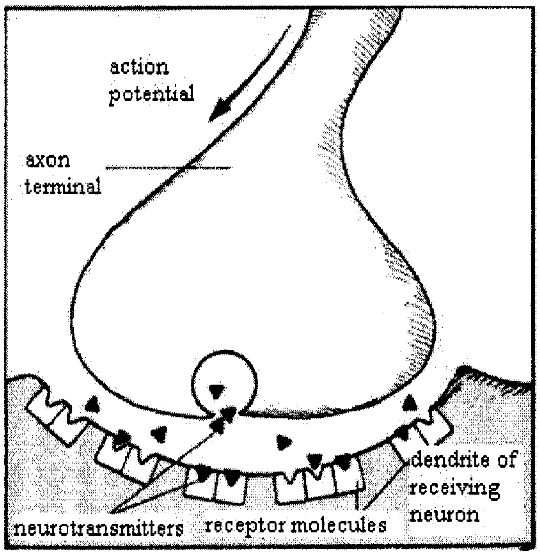

Figure 4: Synapse cleft (from http://www.cassiopaea.org)

A chain of spikes generated by a single neuron is called a *spike train* - a sequence of stereotyped events which occur at regular or irregular intervals. The form of a single spike does not carry any information. Rather, it is the number and the timing of spikes which matter. Spikes in a spike train are usually well separated. Even with very strong input, it is impossible to excite a second spike during or immediately after a first one. The minimal distance between two spikes defines the *absolute refractory period* of the neuron where it is difficult, but not impossible, to excite a spike.

An input at an excitatory synapse reducing the negative polarization of the membrane is called *depolarizing*; an input increasing the negative polarization of the membrane even further is called *hyperpolarizing.* We let $u(t)$ denote the membrane potential of a neuron at time t and u_{rest} denote the rest membrane potential

Consider a spike from a presynaptic neuron j which arrives at neuron i at time $t = 0$. Then,

$$u_i(t) = u_{rest}, \text{ whenever } t < 0,$$

and we define *post-synaptic potential* $\epsilon_{ij}(t)$ as:

$$u_i(t) - u_{rest} = \epsilon_{ij}(t) \text{ whenever } t \geq 0.$$

If the post-synaptic potential is positive, the synapse is said to be *excitatory*; if the post-synaptic potential is negative, the synapse is *inhibitory*.

Next, consider two pre-synaptic neurons $j = 1, 2$, which both send spikes to the post-synaptic neuron i. Neuron $j = 1$ fires spikes at $t_1^{(1)}$, $t_1^{(2)}$, $\ldots$; similarly neuron $j = 2$ fires at $t_2^{(1)}$, $t_2^{(2)}$, $\ldots$. Each spike respectively evokes a post-synaptic potential ϵ_{i1} or ϵ_{i2}. post-synaptic potentials have amplitudes in the range of $1mV$. Before the post-synaptic neuron i fires, the total change of the potential is approximately the sum of the individual post-synaptic potentials,

$$u_i(t) = \sum_j \sum_f \epsilon_{ij}(t - t_j^{(f)}) + u_{rest}.$$

This is illustrated in Figure 5.

As soon as $u_i(t)$ reaches a critical value θ, this situation changes and neuron i now exhibits a spike with an amplitude of about $100mV$. This spike will propagate along the axon of neuron i to the synapses of other neurons. After the pulse the membrane potential does not directly return to the resting potential, but passes through a phase of hyper-polarizing below the resting value. The critical value for spike initiation is about $20 - 30mV$ above the resting potential. In most neurons, about $20 - 50$ pre-synaptic spikes have arrive within a short time window before post-synaptic spikes

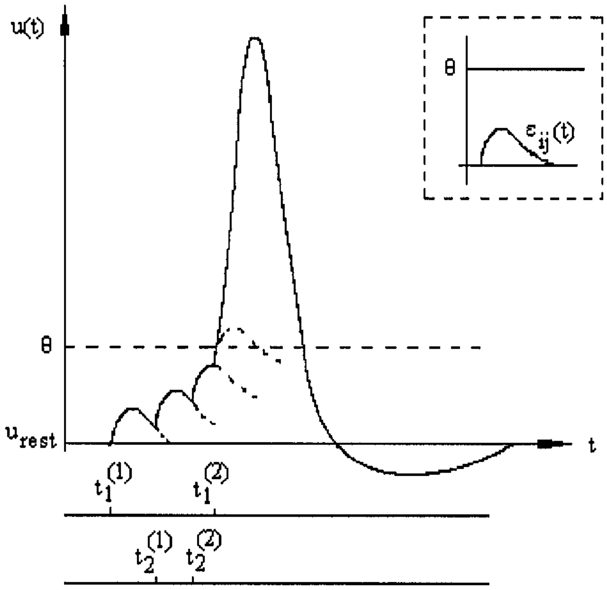

Figure 5: One neuron receives signals from two neurons (from Gerstner and Kistler (2005))

are triggered.

2.3 Formal Spiking Neuron Models

This section is adapted from Gerstner and Kistler (2005).

2.3.1 Integrate-and-fire Model

The basic circuit of an *integrate-and-fire model* is shown in Figure 6. The circuit consists of a capacitor C in parallel with a resistor R driven by a current I. The current I can be split into two components, I_R and I_C. I_R is the resistive current which passes through the linear resistor R. According to Ohm's law, we have $I_R = u/R$ where u is the voltage across the resistor. I_C is a capacitive current which charges the capacitor C. According to the definition of the capacity, $C = Q/u$ (where Q is

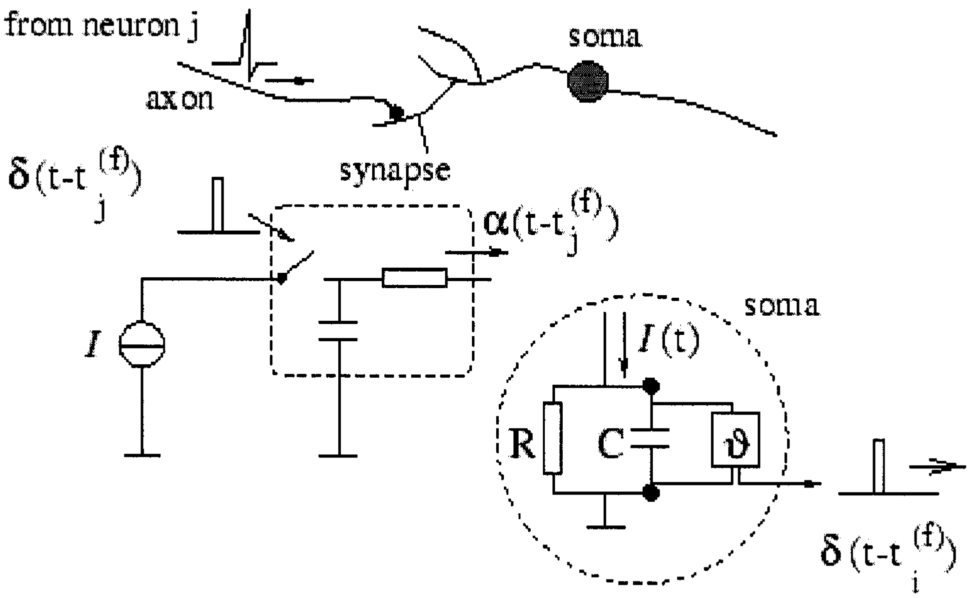

Figure 6: Schematic diagram of the integrate-and-fire model (from Gerstner and Kistler (2005))

the charge and u the voltage across the capacitor), we have $I_C = C\frac{du}{dt}$. Thus, at time t,

$$I(t) = \frac{u(t)}{R} + C\frac{du(t)}{dt} \tag{1}$$

Multiply (1) by R and introduce the time constant τ_m of the *leaky integrator* where $\tau_m = RC$. Then we have

$$\tau_m \frac{du(t)}{dt} = -u(t) + RI(t) \tag{2}$$

We refer to $u(t)$ as the membrane potential and to τ_m as the membrane time constant of the neuron.

In integrate-and-fire models the form of an action potential is not described explicitly. Spikes are formal events characterized by a *firing time*, $t^{(f)}$, which is is defined by a threshold criterion

$$t^{(f)} : u(t^{(f)}) = \vartheta$$

Immediately after $t^{(f)}$, the potential is reset to a new value u_{rest}

$$\lim_{t \rightarrow t^{(f)}; t > t^{(f)}} u(t) \quad = \quad u_{rest} \tag{3}$$

where $u_{rest} < \vartheta$. For $t > t^{(f)}$ the dynamics are again given by (2) until the next threshold crossing occurs. The combination of leaky integration and reset defines the *leaky integrate-and-fire model.*

$$\begin{cases} \tau_m \dfrac{du}{dt} &= -u(t) + RI(t) \\ \lim\limits_{t \rightarrow t^{(f)}; t > t^{(f)}} u(t) &= u_{rest} \end{cases} \tag{4}$$

In its general version, the integrate-and-fire neuron may also incorporate an *absolute refractory* period. When u reaches the threshold at time $t^{(f)}$, the dynamics (2) is interrupted during an absolute refractory time Δ_{abs} and the integration is restarted at time $t^{(f)} + \Delta_{abs}$ with the new initial condition u_{rest}.

2.3.2 Spike Response Model

The *spike response model* is a generalization of the leaky integrate-and-fire model. The state of neuron i is described by its membrane potential u_i. Before spikes arrive, u_i is at its resting value, u_{rest} where $u_{rest} = 0$. Each incoming spike will perturb u_i and it takes some time before u_i returns to zero. The function ϵ_{ij} describes the perturbation on u_i to the incoming spike from neuron j. This perturbation can be enhanced if new incoming spike arrives before u_i reaches zero. After the summation of the effects of several incoming spikes, if u_i reaches the threshold θ, then an output spike is triggered. The form of the action potential and the after-potential is described by a function η. Let us suppose neuron i has fired its last spike at time $\hat{t}_i$. After firing the evolution of u_i is given by

$$\begin{aligned} u_i(t) &= \eta(t-\hat{t}_i) + \sum_j w_{ij} \sum_f \epsilon_{ij}(t-\hat{t}_i, t-t_j^{(f)}) \\ &\quad + \int_0^\infty \kappa(t-\hat{t}_i, s) I^{ext}(t-s)\,ds \end{aligned} \tag{5}$$

where $t_j^{(f)}$ are spikes of presynaptic neurons j, w_{ij} is the synaptic efficacy, and $s = t-t_j^{(f)}$. The last term accounts for an external driving current I^{ext}. The two sums run over all presynaptic neurons j and all firing times $t_j^{(f)} < t$ of neuron j. We emphasize that all terms depend on $t-\hat{t}_i$, the time since the last output spike. Fig 7 shows the time course $u_i(t)$ of the membrane potential of neuron i as a function of time t.

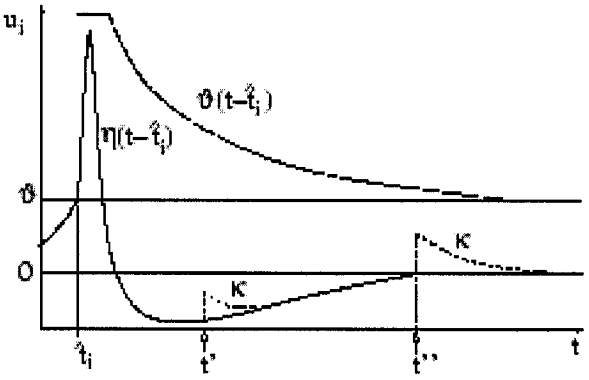

Figure 7: Schematic interpretation of the spike response model (from Gerstner and Kistler (2005))

According to (5), we can see that all terms depend on $t-\hat{t}_i$, which is the time since the last output spike. In a spike response model, the threshold θ is not fixed but may also depend on $t - \hat{t}_i$,

$$\theta \rightarrow \theta(t - \hat{t}_i)$$

During an absolute refractory period Δ^{abs}, we may set θ to a large and positive value to avoid firing and let it relax back to its equilibrium value for $t > \hat{t}_i + \Delta^{abs}$. As mentioned before $\hat{t}_i$ is the last firing time,

$$\hat{t}_i = max\left\{t_i^{(f)} < t\right\}$$

.

The functions η, κ and ϵ_{ij} are *response kernels* which describe the effect of spike emission and spike reception on the variable u_i. A spike of neuron i has been initiated at $\hat{t}_i$. The kernel $\eta(t - \hat{t}_i)$ for $t > \hat{t}_i$ describes the standard form of an action potential of neuron i. Here we need to emphasis that the exact shape of the potential carries no information because the shape is always the same. What matters is whether there is the event 'spike' or not. This event is fully characterized by the firing time $t_i^{(f)}$.

The kernel $\kappa(t - \hat{t}_i, s)$ is the linear response of the membrane potential to an input current. It describes the time course of a deviation of the membrane potential from its resting value that is caused by a short current pulse ("impulse response"). The response depends on the time that has passed since the last output spike at $\hat{t}_i$. Immediately after $\hat{t}_i$, many ion channels are open. The resistance of the membrane is reduced. Therefore, the voltage response to an input current pulse decays back to zero more rapidly than in a neuron that has been inactive. A reduced or shorter response is one of the signatures of neuronal refractoriness. This form of refractory effect is taken care of by making the kernel κ depend, via its first argument, on the time difference $t - \hat{t}_i$. In Figure 7, we can see that response to a first input pulse at

t' is shorter and less pronounced than that to a second one at t''.

The kernel $\epsilon_{ij}(t - \hat{t}_i, s)$ is a function of s which is equal to $(t - t_j^{(f)})$. ϵ_{ij} can be interpreted as the time course of a postsynaptic potential evoked by the firing of a presynaptic neuron j at time $t_j^{(f)}$. It models either an excitatory or inhibitory postsynaptic potential. Similarly as for the kernel κ, the exact shape of the postsynaptic potential depends on the time $t - \hat{t}_i$ that has passed since the last spike of the postsynaptic neuron i. In particular, if neuron i has been active immediately before the arrival of a presynaptic action potential, the postsynaptic neuron is in a state of refractoriness. In this case, the response to an input spike is smaller than that of an 'unprimed' neuron. The first argument of $\epsilon_{ij}(t - \hat{t}_i, s)$ accounts for the dependence upon the last firing time of the postsynaptic neuron.

3 McCulloch-Pitts Neural Networks

In 1943, Warren S. McCulloch, a neuroscientist, and Walter Pitts, a logician, published an important paper entitled *A logical calculus of the ideas immanent in nervous activity.* There they tried to describe how the brain produces highly complex patterns by using many basic neurons that are connected together. They gave a highly simplified model of a neuron which is widely called "McCulloch-Pitts neuron". This model captures key features of the biological neuron with the following physical assumptions:

> 1. The activity of the neuron is an "all-or-one" process.
> 2. A certain fixed number of synapses must be excited within the period of latent addition in order to excite a neuron at any time, and this number is independent of previous activity and position on the neuron.
> 3. The only significant delay within the nervous system is synaptic delay.
> 4. The activity of any inhibitory synapse absolutely prevents excitation of the neuron at that time.
> 5. The structure of the net does not change with time.

This was a seminal paper in the development of *artificial neural networks.*

Artificial neural networks are mathematical or computational models of the networks of neurons that are present in the central nervous system. Their study can be divided into two distinct categories:

1. there are researchers who aim to construct better computer hardware and software by imitating the brain;
2. there are researchers who aim to better understand the brain by studying its mathematical models.

This thesis belongs to the second category.

Much of the work that belongs to the first category concerns self-adjusting artificial neural networks that are capable of learning. Their origin is the book *The organization of behavior: a neuropsychological theory* published in 1949 by Donald Hebb. He said:

> Let us assume that the persistence or repetition of a reverberatory activity (or "trace") tends to induce lasting cellular changes that add to its stability.... When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.

This theory is commonly evoked to explain some types of associative learning in which simultaneous activation of cells leads to a pronounced increased in synaptic strength. In this thesis, we are not concerned with such neural networks: our networks are static and do not learn.

Siu et al. (1995) in their section 1.7 give an account of the major developments in neural network research since the pioneering work of McCulloch and Pitts.

3.1 The McCulloch-Pitts Neuron

We use **R** to denote the set of all real numbers and **N** to denote the set of all nonnegative integers.

A *McCulloch-Pitts neuron* is a threshold function

$$f : \mathbf{R}^m \rightarrow \{0,1\}$$

defined, for some constants $w_1, w_2, \ldots, w_m$ (mnemonic for "weights") and θ (mnemonic for "threshold") by

$$f(x_1, x_2, \ldots, x_m) = \begin{cases} 1 & \text{if } \sum w_i x_i \geq \theta, \\ 0 & \text{otherwise.} \end{cases}$$

We think of the function as a neuron with signals $x_1, x_2, \ldots, x_m$ received at the synapses; positive weights correspond to excitatory synapses and negative weights correspond to inhibitory synapses; $f(x_1, x_2, \ldots, x_m) = 1$ means that the neuron is firing.

3.2 The McCulloch-Pitts Neural Network

(This section was written by my advisor Vašek Chvátal.)

The *McCulloch-Pitts neural network* consists of some number n of McCulloch-Pitts neurons.

$$f_j : \mathbf{R}^{m+n} \rightarrow \{0,1\} \qquad (j = 1, 2, \ldots, n).$$

Given any sequence $x_1, x_2, \ldots, x_m$ of functions

$$x_i : \mathbf{N} \rightarrow \mathbf{R} \qquad (i = 1, 2, \ldots, m),$$

it computes a sequence $y_1, y_2, \ldots, y_n$ of functions

$$y_j : \mathbf{N} \rightarrow \{0,1\} \qquad (j = 1, 2, \ldots, n)$$

as follows: for all j, we set $y_j(0) = 0$ and, for all positive integers t,

$$y_j(t) = f_j(x_1(t-1), \ldots, x_m(t-1), y_1(t-1), \ldots, y_n(t-1)).$$

We think of variable t as marking discrete time; each of the n neurons may receive its signals from any of the m inputs (it is linked to input i if and only if its w_i is nonzero) and from any of the n neurons (it is linked to neuron j if and only if its w_{m+j} is nonzero). The real numbers $x_i(t)$ measure the inputs received by the network from the outside environment at time t; the bits $y_j(t)$ tell us which neurons are firing at time t; firing or not firing of a neuron at time t depends on the signals sent to it at time $t-1$ (and received, with the "synaptic delay", at time t). Some of the n neurons may be designated as output neurons.

With this neural network, we associate a directed graph on $m+n$ vertices. These vertices are $x_1, x_2, \ldots, x_m$ and $y_1, y_2, \ldots, y_n$; if f_k is defined by

$$f_k(x_1, \ldots, x_m, y_1, \ldots, y_n) = \begin{cases} 1 & \text{if } \sum a_{ik}x_i + \sum b_{jk}y_j \geq \theta_k, \\ 0 & \text{if otherwise,} \end{cases}$$

then there is a directed edge from x_i to y_k if and only if $a_{ik} \neq 0$ and there is a directed edge from y_j to y_k if and only if $b_{jk} \neq 0$. If this graph contains no directed cycles, then the network is called a *feed-forward network.*

The number of neurons in the network is called the *size* of the network. The *level* of a neuron in a feed-forward network is defined as the maximum number of neurons on any path from an input to this neuron (which is counted as one of the neurons on the path); the level of vertices $x_1, x_2, \ldots, x_m$ are defined as zero. The *depth* of a feed-forward network is defined to be the maximum number of neurons on any path from an input to an output.

3.3 Threshold Circuits

A McCulloch-Pitts neuron is also called a *threshold gate.* A threshold gate computes a linear threshold function $f : \mathbf{R}^n \rightarrow \{0,1\}$ such that

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} w_i x_i \geq \theta \\ 0 & otherwise \end{cases}$$

A feed-forward McCulloch-Pitts neural network is also called a *threshold circuit.* A threshold circuit consists of some number n of threshold gates,

$$f_j : \mathbf{R}^{m+j-1} \rightarrow \{0,1\} \qquad (j = 1, 2, \ldots, n).$$

Given any sequence of real numbers $x_1, x_2, \ldots, x_m$, it computes a sequence of bits $y_1, y_2, \ldots, y_n$ as follows:

$$\begin{aligned} y_1 &= f_1(x_1, \ldots, x_m) \text{ and} \\ y_j &= f_j(x_1, \ldots, x_m, y_1, \ldots, y_{j-1}), \text{ for all } j = 2, 3, \ldots n. \end{aligned}$$

The threshold gate which computes y_1 may receive its signals from any of the m inputs (it is linked to input i if and only if its w_i is nonzero); the threshold gate j with $j \geq 2$ may receive its signals from any of the m inputs (it is linked to input i if and only if its w_i is nonzero) and from any of the $j-1$ threshold gates which compute $y_1, y_2, \ldots, y_{j-1}$ (it is linked to the threshold gate k with $1 \leq k \leq j-1$ if and only if its w_{m+k} is nonzero). Some of the n threshold gates may be designated as output gates.

With this threshold circuit, we associate a directed graph on $m+n$ vertices. These vertices are $x_1, x_2, \ldots, x_m$ and $y_1, y_2, \ldots, y_n$; if f_1 is defined by

$$f_1(x_1, \ldots, x_m) = \begin{cases} 1 & \text{if} \quad \sum_{i=1}^{m} a_{i1} x_i \geq \theta_1, \\ 0 & \text{if} \quad \text{otherwise}, \end{cases}$$

then there is a directed edge from x_i to y_1 if and only if $a_{i1} \neq 0$; if f_k with $2 \leq k \leq n$ is defined by

$$f_k(x_1, \ldots, x_m, y_1, \ldots, y_{k-1}) = \begin{cases} 1 & \text{if} \quad \sum_{i=1}^{m} a_{ik} x_i + \sum_{j=1}^{k-1} b_{jk} y_j \geq \theta_k, \\ 0 & \text{if} \quad \text{otherwise}, \end{cases}$$

then there is a directed edge from x_i to y_k if and only if $a_{ik} \neq 0$ and there is a directed edge from y_j to y_k if and only if $b_{jk} \neq 0$.

4 What The Frog's Eye Tells The Frog's Brain

4.1 Retina

The retina is a thin layer of neural cells that lines the back of the eyeball of vertebrates and some cephalopods (http://en.wikipedia.org/wiki/Retina), such as squids and octopi. It has three layers of nerve cells. Figure 8 by Ramón y Cajal shows the structure of a sparrow's retina.

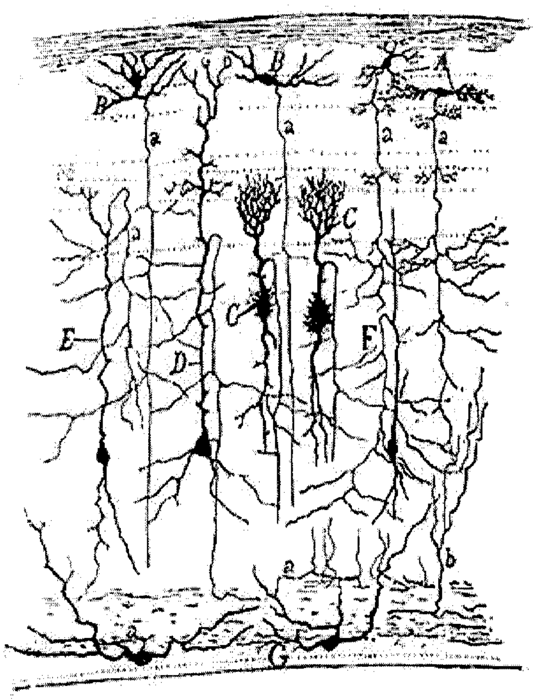

Figure 8: Drawing of a section through the optic tectum of a sparrow (Santiago Ramon y Cajal, c.1900)

The outermost layer of retina contains *photoreceptor cells* (mainly rods and cones) which respond to light. The innermost layer contains ganglion cells. The optic nerve carries the ganglion cell axons to the brain. Between photoreceptor cells and ganglion cells, there are bipolar cells, horizontal cells, and amacrine cells. Photoreceptors, bipolar, and horizontal cells make synaptic connections with each other in the *outer plexiform layer.* The bipolar, amacrine, and ganglion cells make contact in the *inner plexiform layer.* Bipolar cells bridge the two layers.

4.2 Frog's Visual System

Frog's retina has unique features: the uniformity of its retina, the normal lack of eye and head movements, and the relative simplicity of the connection of his eye to his brain. Maturana (1959) found that, in the frog, there are about one million receptors (rods and cones), 2.5 to 3.5 million connecting neurons (bipolar cells, horizontal cells, and amacrine cells), and half a million ganglion cells. Figure 9 shows the diagram of frog's visual system.

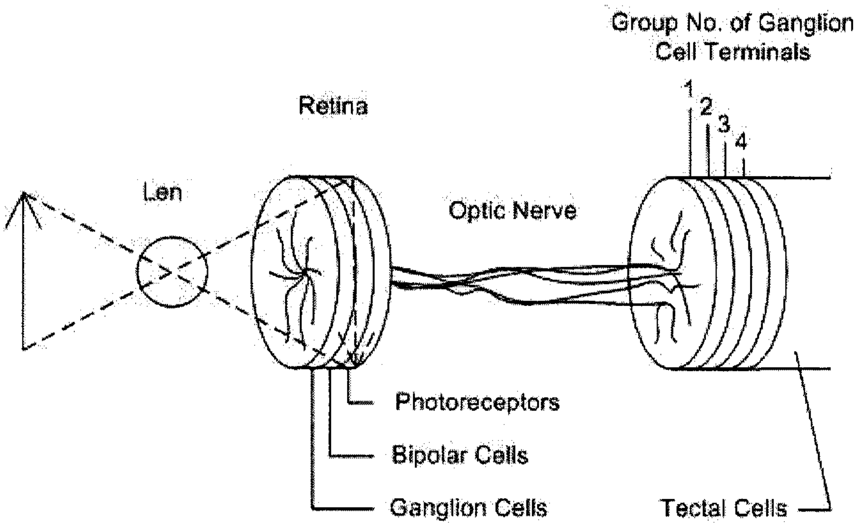

Figure 9: Diagram of frog's visual System

In 1959, Jerome Lettvin, Humberto Maturana, Warren McCulloch, and Walter Pitts published an important paper entitled *What the frog's eye tells the frog's brain.* They found by experiments that, rather than transmitting raw data from the photoreceptors to the brain, the frog's optical nerve transmits to the brain information obtained by preprocessing the data in the retina by four separate operations: 1) sustained contrast detection, 2) net convexity detection, 3) moving edge detection, and 4) net dimming detection. Each operation has its result transmitted by a particular group of fibers. Here is what Lettvin et al. say:

> 1. ***Sustained Contrast Detectors***
>
>   An unmyelinated axon of this group does not respond when the general illumination is turned on or off. If the sharp edge of an object either lighter or darker than the background moves into the field and stops, it discharges promptly and continues discharging no matter what the shape of the edge or whether the object is smaller or larger than the receptive field. ...
>
> 2. ***Net Convexity Detectors***
>
>   These fibers form the other subdivision of the unmyelinated population. ... The fiber will not respond to the straight edge of a dark object moving through its receptive field. ... Usually a fiber will respond indefinitely only to objects which have moved into the field and then lie wholly or almost wholly interior to the receptive field. The discharge is greater the greater the convexity, or positive curvature, of the boundary of the dark object until the object becomes as small as about one half the width of the receptive field. ... Objects lighter than the background produce almost no response unless they have enough relief to cast a slight shadow at the edge. ...
>
> 3. ***Moving-Edge Detectors***
>
>   These fibers are myelinated .... Such a fiber responds to any distinguishable edge moving through its receptive field, whether black against white or the other way around. ... The response to moving objects is much greater than to changes in total illumination ....
>
> 4. ***Net Dimming Detectors***
>
>   ... They are myelinated and the fastest conducting afferents, clocked at at 10 meters per second. One such fiber responds to sudden reduction of illumination

> by a prolonged and regular discharge. ...

The fibers mentioned here are axons of ganglion cells. In the frog's retina, there are half a million ganglion cells and one million photoreceptors. There is a synaptic path from a photoreceptor to many ganglion cells and a ganglion cell receives paths from many thousand photoreceptors. These paths are formed by connecting neurons (bipolars, horizontals, and amacrines), of which there are 2.5 to 3.5 million. Hartline (1938) first used the term *receptive field* for the region of retina whose photoreceptors are connected to a particular ganglion cell by synaptic paths.

4.3 Net Dimming Detectors

Our initial plan was to construct a McCulloch-Pitts network simulating a ganglion cell whose axon belongs to the fourth group of fibers (and its pre-synaptic ancestors of this ganglion, which extend though the layer of interconnecting neurons all the way to the photoreceptors). About these fibers, Lettvin et al. wrote

> If the spot of light disappears from the [the receptive field]...the discharge begins immediately, decreases slowly in frequency, and lasts a long time. It can be abolished promptly by turning the spot of light on again. ...
>
> ...the responses occurs ...at both center and periphery of that field and that effect of removing light from the periphery adds to the effect of a reduction of light at the center, with a weight decreasing with distance.

in their section I (Introduction) and

> ...Darkening of a spot produces less response when it is in the periphery of the field than when it is at the center. The effect of a moving object is directly related to its size and relative darkness. The response is prolonged if a large dark object stops within the field. it is almost independent of illumination, actually increasing as the light gets dimmer. ...If the general lighting is sharply dimmed, but not turned off entirely, the

> consequent prolonged response is diminished or abolished after a dark object passes through the receptive field. ...one turns of the light and sets up a prolonged response. Then the amount of light which must be restored to interrupt the response gets less and less the longer one waits. ...If we darken the general lighting only by a factor of 100, we also get a prolonged discharge. However, if we turn of the light completely a few seconds after the 100/1 dimming and then turn it back on to the same dim level, the discharge is increased by the second dimming and is completely or almost completely abolished by the relighting. The effect of moving a dark object through the field after dimming is to impose a second dimming pulse followed by brightening as the object passes

in their section III (Findings). This seems very complicated. We have settled for a simple variation on this complex theme. This will be described in our next section.

5 Our Definition of Dimming

In the frog's retina, each receptive field is a two-dimensional region filled with a mosaic of photoreceptors. In our abstraction, a receptive field means simply a set of elements with we call, interchangeably, photoreceptors or *pixels.*

We reserve letter m for the number of photoreceptors in our receptive field. By an *image*, we mean a sequence

$$\xi_1, \xi_2, \ldots, \xi_m$$

of real numbers such that each ξ_i measures the light on the i-th photoreceptor on a continuous scale between the black 0 and white 1.

In everyday language, "dimming" may refer to a protracted process. However, Lettvin et al., talk about "sudden reduction of illumination" and "sudden darkening". Accordingly, we will examine changes in the image from time t, when the image is

$$\xi_1(t), \xi_2(t), \ldots, \xi_m(t) \tag{6}$$

to the next moment $t+1$, when the image is

$$\xi_1(t+1), \xi_2(t+1), \ldots, \xi_m(t+1). \tag{7}$$

If a pixel i darkens from time t to time $t+1$, i.e., $\xi_i(t) > \xi_i(t+1)$, then we call this pixel a *dimming* pixel. If a pixel lightens from time t to time $t+1$, i.e., $\xi_i(t) < \xi_i(t+1)$, then we call this pixel a *lightening* pixel.

We declare that an image is dimming from time t to time $t+1$ if all of the following two conditions are satisfied.

1. The average brightness in the image at time $t+1$ is less then the average

brightness sum at time t, in the sense that

$$\frac{1}{m}\sum_{i=1}^{m}\xi_i(t+1) < \frac{1}{m}\sum_{i=1}^{m}\xi_i(t).$$

2. The total number of dimming pixels is greater than some prescribed proportion, say 10%, of the total number of the non-black pixels.

The two-dimensional geometry of the retina is completely irrelevant to this definition: all that matters are the two sequences (6) and (7).

Condition 1 seems obvious. Condition 2 is used to filter out a certain kind of noise. For example, take $m = 25$. If two pixels dim by 0.25 and the remaining 23 pixels lighten by 0.01, then Condition 1 is satisfied and yet we do not want to say the image is dimming: here, only two pixels darken while the whole rest of the background actually lightens up. This example is illustrated in Figure 10, where the 25 pixels are arranged in a 5×5 array. To filter out this kind of noise, we could insist that at least 10% of the total number of pixels be dimming. However, such a simple policy could get frogs in trouble because it disregards the fact that totally black pixels cannot darken any more. For instance, if a large predator completely blocks the light against a background where 22 out of the 25 pixels are already black, then we would miss the appearance of this shadow. The adjusted policy, where the total number of pixels is replaced by the total number of non-black pixels, eliminates noise without exposing

0.50	0.50	0.50	0.50	0.50
0.50	0.50	0.50	0.50	0.50
0.50	0.50	0.50	0.50	0.50
0.50	0.50	0.50	0.50	0.50
0.50	0.50	0.50	0.50	0.50

t

$\Longrightarrow$

0.51	0.51	0.51	0.51	0.51
0.51	0.51	0.51	0.51	0.51
0.51	0.51	0.25	0.25	0.51
0.51	0.51	0.51	0.51	0.51
0.51	0.51	0.51	0.51	0.51

$t+1$

Figure 10: Condition 1 is satisfied but the image does not dim

frogs to life-threatening danger.

6 Observation on Dimming Detectors

6.1 Discrete Encodings of Real-valued Inputs

In all McCulloch-Pitts networks, the inputs (which are the real numbers) have a different data type from the internal variables (which are bits). The fact that the input data type differs from the data type of the internal variables reflects biological reality. The photoreceptors (rods and cones) respond to light not by firing action potentials, but by graded changes in membrane potential. Many of them synapse on a single bipolar cell, which reacts to the combination of these graded inputs by the usual zero-one output (firing or not firing). Accordingly, the initial layer of our dimming detector must encode the real number input ξ (measuring light on a continuous scale from black 0 to white 1) as a sequence of bits. To do this, we coerce ξ into a discrete scale of k *gray-scale degrees* by setting $x = \lfloor \xi k \rfloor$, and then we encode x as a sequence of bits.

Two encoding schemes spring to mind:

- *unary encoding*, where each integer x in the range $[0, k]$ is encoded by a sequence of k bits and the first x bits are set at 1;
- *binary encoding*, where each integer x in the range $[0, k]$ is encoded by a sequence of $\lfloor \log_2 k \rfloor + 1$ bits as usual.

The binary encoding is more efficient (by a factor of three when k is twelve, a factor of ten when k is 60), but it does not seem plausible that the frog's eye works with binary encodings. For instance, if k is 50, then input 0.31 (coerced into gray-scale degree 15) will be encoded into 01111, and input 0.32 (coerced into gray-scale degree 16) will be encoded into 10000. The relatively minor change in illumination, 1% of the whole range from the black to white, would force all five encoding neurons to flip their on/off firing status. For this reason, we reject binary encoding as absurd.

6.2 All Dimming Detectors Must be Approximate

Let m be the number of photoreceptors in a receptive field and let each $\xi_i(t)$ measure the light on the i-th photoreceptor on time t on a continuous scale between the black 0 and white 1. We would like to construct McCulloch-Pitts neural networks which, given a sequence $\xi_1, \xi_2, \ldots, \xi_m$ of functions

$$\xi_i : \mathbf{N} \to \mathbf{R} \qquad (i = 1, 2, \ldots, m),$$

compute a single function

$$y : \mathbf{N} \to \{0, 1\} \qquad (j = 1, 2, \ldots, n)$$

such that $y(t+\Delta) = 1$ for some positive integer constant Δ if and only if the receptive field dimmed (in the sense defined in Section 5) from time t to time $t+1$.

Unfortunately this cannot be done. Since McCulloch-Pitts networks do not work with real numbers but only with their discrete approximations, they cannot test conditions such that

$$\sum_{i=1}^{m} \xi_i(t+1) < \sum_{i=1}^{m} \xi_i(t).$$

They can only perform approximations of such tests. For example, they could compute a bit $c_1(t+\Delta)$ such that

- $c_1(t+\Delta) = 1$ whenever $\sum_{i=1}^{m} \xi_i(t+1) \leq \sum_{i=1}^{m} \xi_i(t) - 0.02m$,

- $c_1(t+\Delta) = 0$ whenever $\sum_{i=1}^{m} \xi_i(t+1) > \sum_{i=1}^{m} \xi_i(t)$.

(If $\sum_{i=1}^{m} \xi_i(t) - 0.02 < \sum_{i=1}^{m} \xi_i(t+1) \leq \sum_{i=1}^{m} \xi_i(t)$, then $c_1(t+\Delta)$ may be 1 or 0.) Similarly, the non-black pixels in Condition 2 may be replaced by very dark pixels in the sense of $\xi < 0.01$; a pixel might be declared dimming if $\xi_i(t+1) \leq \xi_i(t) - 0.02$ and lightening if $\xi_i(t+1) > \xi_i(t)$.

The dimming detector we are going to construct uses such an approximation of our definition of dimming.

6.3 Biological Constraints

We will not pretend that our networks are models of the net dimming detectors studied by Lettvin et al.: they react to a stimulus that is far simpler than the complex stimuli described by Lettvin et al. and they react by firing a single spike rather than a slowly vanishing sequence of spikes. Even so, we would like to keep the parameters of these networks consistent with our knowledge of frog's neuroanatomy.

- *The number of inputs.* In their experiments, Lettvin et al. confronted the frog's eye with an aluminum hemisphere that was used as a background for various visual stimuli. They found that the receptive field of the dimming detectors is about 15° large. This area covers about $0.7\% \sim 0.9\%$ of the hemisphere: since there are one million photoreceptors of retina, we conclude that the number of photoreceptors in the receptive field of a single dimming detector is between $7,000$ and $9,000$.

- *The size.* Except for the single output (corresponding to a ganglion cell), each neuron in our network corresponds to one of the connecting neurons (bipolars, horizontals, and amacrines) in the frog's retina. In Section 4.2, we noted that in the frog, there are about one million receptors, 2.5 to 3.5 million connecting neurons, and half a million ganglion cells.

- *The fan-in.* We know that, in humans, a spinal motor cell receives about 10,000 contacts and a Purkinje cell of the cerebellum receives approximately 150,000 contacts. It seems reasonable to conclude that 10,000 is an upper bound (possibly a very generous one) on the number of pre-synaptic contact received by a neuron in frog's retina.
- *The fan-out.* We know that, in humans, the branches of the axon of one neuron may form synapses with as many as 1000 other neurons. It seems reasonable to conclude that 1,000 is an upper bound on the number of post-synaptic contact sent by a neuron in frog's retina.

6.4 Photoreceptors Must be Accessed in Groups

The real-valued data

$$\xi_1(t), \xi_2(t), \ldots, \xi_M(t)$$

that come from the M photoreceptors in the retina must be encoded by bits and these bits must be stored for confrontation with the encoding of

$$\xi_1(t+1), \xi_2(t+1), \ldots, \xi_M(t+1).$$

When k bits are used to encode each real number, we need at least k neurons to encode the number. Since the retina contains fewer than four million neurons other than the photoreceptors and the ganglion cells, it can handle at most $4 \times 10^6/k$ real numbers at a time; since M is about 10^6, this means that the M photoreceptors cannot be handled one by one. The only way we see of getting around this problem is to have each ganglion cell access the photoreceptors not individually but in groups.

7 Our Dimming Detector

We split the entire set of 10^6 photoreceptors into $2,500$ sections of size 400; each receptive field consisting of $8,000$ photoreceptors will be split into 20 sections.

Each section (which may belong to many overlapping receptive fields) is allotted its own $k + 2$ neurons. When $k = 100$, the total number of neurons allotted to all the $2,500$ sections comes to slightly more than a quarter of a million. (These neurons correspond to bipolar and horizontal cells in the outer plexiform layer.) Subtracting this total from the available pool of 2.5 to 3.5 million neurons, we are left with 2.25 to 3.25 million neurons, which comes to between 4 and 7 neurons per ganglion cell. (These neurons correspond to bipolar and amacrine cells in the inner plexiform layer.) We are going to construct a dimming detector which uses, in addition to the neurons allotted to the sections of the retina in its receptive field, three extra neurons (the ganglion cell itself being the fourth one).

The way of grouping photoreceptors into sections is not unique. In this section, we first demonstrate a network based on the above arrangement. At the end of this section, we give a generic parameterized network in Section 7.6.

7.1 The Parameters

Our inputs are 8000 functions

$$\xi_j : \mathbf{N} \to \mathbf{R} \quad (j = 1, 2, \ldots 8000)$$

measuring the intensity of the light on the photoreceptors. These 8000 receptors form 20 sections in the retina, with 400 receptors in each section. We choose the numbering of the photoreceptors so that the i-th group consists of photoreceptors

$$400(i-1)+j \quad \text{with} \quad j=1,2,\ldots 400.$$

The *average brightness* $\alpha(t)$ of the receptive field at time t is defined by

$$\alpha(t)=\frac{1}{8000}\sum_{k=1}^{8000}\xi_k(t)$$

The *average brightness* $\alpha_i(t)$ of the i-th section of receptive field at time t is defined by

$$\alpha_i(t)=\frac{1}{400}\sum_{j=1}^{400}\xi_{400(i-1)+j}(t).$$

Let us note that $\alpha(t)$ is the average of

$$\alpha_1(t),\alpha_2(t),\ldots,\alpha_{20}(t).$$

7.2 The Objective

We want to construct a McCulloch-Pitts network which, given the 8000 input functions ξ_j, computes an output function

$$y:\ \mathbf{N}\rightarrow\mathbf{R}$$

such that, for some positive integer Δ and for some positive numbers ε_1, ε_2, and ε_3, the output value $y(t+\Delta)$ obeys the following rules.

CASE 1: $\alpha_i(t)\geq\varepsilon_1$ *for more than ten of the twenty subscripts* i.

In this case, we insist that

- if $\alpha(t+1)\leq\alpha(t)-\varepsilon_2$

and $\alpha_i(t+1) \leq \alpha_i(t) - \varepsilon_3$ for at least three subscripts i, then $y(t+\Delta) = 1$;

- if $\alpha(t+1) \geq \alpha(t)$, then $y(t+\Delta) = 0$;
- if $\alpha_i(t+1) \leq \alpha_i(t) - \varepsilon_3$ for at most two subscripts i, then $y(t+\Delta) = 0$.

CASE 2: $\alpha_i(t) \geq \varepsilon_1$ *for at most ten of the twenty subscripts* i.

In this case, we insist that

- if $\alpha(t+1) \leq \alpha(t) - \varepsilon_2$ and $\alpha_i(t+1) \leq \alpha_i(t) - \varepsilon_3$ for at least two subscripts i, then $y(t+\Delta) = 1$;
- if $\alpha(t+1) \geq \alpha(t)$, then $y(t+\Delta) = 0$;
- if $\alpha_i(t+1) \leq \alpha_i(t) - \varepsilon_3$ for at most one subscript i, then $y(t+\Delta) = 0$.

Under all other circumstances, not covered by this list of six, $y(t+\Delta)$ may be 1 or 0, and we do not care which it is.

7.3 Stratified Networks and Pipelining

We say that a McCulloch-Pitts network is *stratified* if, for each of its directed edges $\{u, v\}$, the level of v is one plus the level of u. When a network is layered, its computations can be *pipelined* in the sense that processing input

$$\xi_1(t), \xi_2(t), \ldots, \xi_m(t)$$

does not interfere with the processing of input

$$\xi_1(t+1), \xi_2(t+1), \ldots, \xi_m(t+1).$$

The dimming detector that we are going to construct will be stratified.

7.4 Construction

The network begins with twenty *encoding units.* For each section $i = 1, 2, \ldots, 20$, encoding unit E_i receives 400 real numbers

$$\xi_{400(i-1)+j}(t)$$

with $j = 1, 2, \ldots 400$, after a unit delay, outputs 100 bits, $e_{i\sigma}(t+1)$ with $\sigma = 1, 2, \ldots, 100$, such that

$$\sum_{\sigma=1}^{100} e_{i\sigma}(t+1) = \left\lfloor \frac{1}{4} \sum_{j=1}^{400} \xi_{400(i-1)+j}(t) \right\rfloor .$$

This unit consists of 100 neurons, one for each gray-scale degree $\sigma = 1, 2, \ldots, 100$, working in parallel. The σ-th of these neurons is defined by the threshold function

$$e_{i\sigma} = \begin{cases} 1 & \text{if } \sum_{j=1}^{400} \xi_{400(i-1)+j} \geq 4\sigma, \\ 0 & \text{otherwise.} \end{cases}$$

Next come twenty *dimming section detectors* and twenty *dark section detectors.* Each of these detectors is a single neuron.

For each section $i = 1, 2, \ldots, 20$, the dimming section detector receives 400 real numbers $\xi_{400(i-1)+j}(t+1)$ with $j = 1, 2, \ldots, 400$ from the i-th section photoreceptors and 100 bits $e_{i\sigma}(t+1)$ with $\sigma = 1, 2, \ldots, 100$ from the i-th encoding unit E_i; after a unit delay, it outputs a bit $\delta_i(t+2)$. This neuron is defined by the threshold function

$$\delta_i = \begin{cases} 1 & \text{if } 4\sum_{\sigma=1}^{100} e_{i\sigma} - \sum_{j=1}^{400} \xi_{400(i-1)+j} \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

For each section $i = 1, 2, \ldots, 20$, the dark section detector receives one bit $e_{i1}(t+1)$ and, after a unit delay, outputs a bit $\nu_i(t+2)$. This neuron is defined by the threshold function

$$\nu_i = \begin{cases} 1 & \text{if } -e_{i1} \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

The rest of the network consists of four neurons: *Condition* 1 *neuron*, *Condition* 2 *neuron*, *memory neuron*, and *conjunction neuron*.

Condition 1 neuron receives 2000 bits $e_{i\sigma}(t+1)$ with $i = 1, 2, \ldots, 20$ and $\sigma = 1, 2, \ldots, 100$ from all encoding units and 8000 real numbers $\xi_j(t+1)$ with $i = 1, 2, \ldots, 8000$ from all photoreceptors. After a unit delay, it outputs a bit $c_1(t+2)$.

This neuron is defined by the threshold function

$$c_1 = \begin{cases} 1 & \text{if } 4\sum_{i=1}^{20}\sum_{\sigma=1}^{100} e_{i\sigma} - \sum_{j=1}^{8000} \xi_j \geq 20, \\ 0 & \text{otherwise.} \end{cases}$$

Condition 2 neuron receives 20 bits $\delta_i(t+2)$ with $i = 1, 2, \ldots, 20$ from all dimming section detectors and 20 bits $\nu_i(t+2)$ with $i = 1, 2, \ldots, 20$ from all dark section detectors. After a unit delay, it outputs a bit $c_2(t+3)$. This neuron is defined by the threshold function

$$c_2 = \begin{cases} 1 & \text{if } \sum_{i=1}^{20} \delta_i + 0.1 \sum_{i=1}^{20} \nu_i \geq 3, \\ 0 & \text{otherwise.} \end{cases}$$

The single memory neuron receives one bit $c_1(t+2)$ from Condition 1 neuron and, after a unit delay, outputs one bit $c_1'(t+3)$. This neuron is defined by the threshold function

$$c_1' = \begin{cases} 1 & \text{if } c_1 \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Conjunction neuron receives one bit $c_1'(t+3)$ from the memory neuron and one bit $c_2(t+3)$ and, after a unit delay, outputs a bit $y(t+4)$ as the ultimate output of the network. This neuron is defined by the threshold function

$$y = \begin{cases} 1 & \text{if } c_1' + c_2 \geq 2, \\ 0 & \text{otherwise.} \end{cases}$$

Figure 11 shows the structure of the network. In the figure, there are two kinds of bold arrows, dashed and solid. Both kinds of arrows represent a group of lines which propagate a group of data; the dashed arrows propagate real numbers while the solid arrows propagate bits.

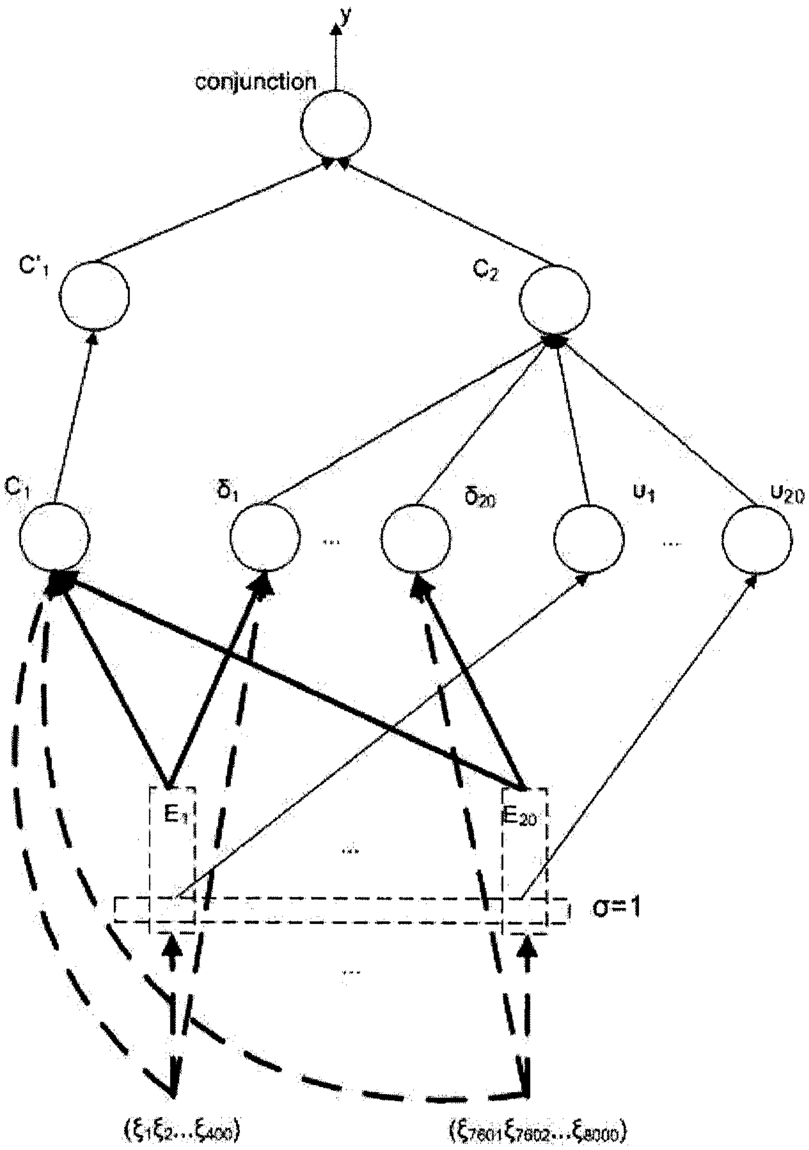

Figure 11: Overall structure of the approximate network

7.5 Analysis

Let us define

$$x_i(t) = \sum_{\sigma=1}^{100} e_{i\sigma}(t+1).$$

In this notation, we have

- $\sum_{\sigma=1}^{100} e_{i\sigma}(t+1) = x_i(t)$,
- $\delta_i(t+2) = 1$ if and only if $x_i(t) - 100\alpha_i(t+1) \geq 0.25$,
- $\nu_i(t+3) = 1$ if and only if $x_i(t) = 0$,
- $c_1'(t+3) = c_1(t+2) = 1$ if and only if $\sum_{i=1}^{20} x_i(t) - 2000\alpha_i(t+1) \geq 5$

Since

$$x_i(t) = \lfloor 100\alpha_i(t) \rfloor.$$

we have

$$100\alpha_i(t) - 1 \; < \; x_i(t) \; \leq \; 100\alpha_i(t),$$

and so

$$2000\alpha(t) - 20 \; < \; \sum_{i=1}^{20} x_i(t) \; \leq \; 2000\alpha(t).$$

If $c_1(t+3) = 1$, then

$$\sum_{i=1}^{20} x_i(t) \geq 2000\alpha(t+1) + 5,$$

and so

$$2000\alpha(t) \geq \sum_{i=1}^{20} x_i(t) \geq 2000\alpha(t+1) + 5 > 2000\alpha(t+1).$$

If $c_1(t+3) = 0$, then

$$\sum_{i=1}^{20} x_i(t) < 2000\alpha(t+1) + 5,$$

and so

$$2000\alpha(t) < \sum_{i=1}^{20} x_i(t) + 20 < 2000\alpha(t+1) + 25.$$

To summarize,

$$\begin{aligned} \alpha(t+1) \leq \alpha(t) - 0.0125 &\Rightarrow c_1(t+3) = 1 \Rightarrow \alpha(t+1) < \alpha(t), \\ \alpha(t+1) \geq \alpha(t) &\Rightarrow c_1(t+3) = 0 \Rightarrow \alpha(t+1) > \alpha(t) - 0.0125 \end{aligned}$$

Figure 12 shows the relative positions between $\alpha(t)$ and $\alpha(t+1)$ that trigger Condition 1 unit to fire or block Condition 1 to fire.

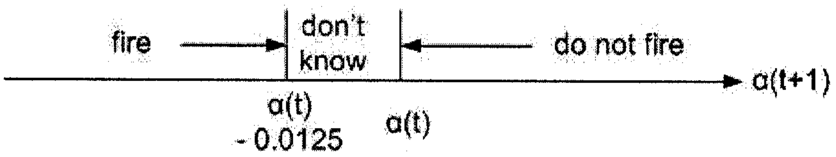

Figure 12: Condition 1 unit fires or not based on the relative positions between $\alpha(t)$ and $\alpha(t+1)$

In addition, we have

- $c_2(t+3) = 1$ if and only if

$$\sum_{i=1}^{20} \delta_i(t+2) \geq 3$$

or

$$\sum_{i=1}^{20} \delta_1(t+2) = 2 \text{ and } \sum_{i=1}^{20} \nu_1(t+2) \geq 10$$

or

$$\sum_{i=1}^{20} \delta_1(t+2) = 1 \text{ and } \sum_{i=1}^{20} \nu_1(t+2) = 20.$$

If $\delta_i(t+2) = 1$, then

$$x_i(t) \geq 100\alpha_i(t+1) + 0.25,$$

and so

$$100\alpha_i(t) \geq x_i(t) \geq 100\alpha_i(t+1) + 0.25 > 100\alpha_i(t+1).$$

If $\delta_i(t+2) = 0$, then

$$x_i(t) \ < \ 100\alpha_i(t+1) + 0.25,$$

and so

$$100\alpha_i(t) < x_i(t) + 1 < 100\alpha_i(t+1) + 1.25.$$

To summarize,

$$\begin{aligned} \alpha_i(t+1) \ &\leq \ \alpha_i(t) - 0.0125 \ \Rightarrow \ \delta_i(t+2) = 1 \ \Rightarrow \alpha_i(t+1) \ < \ \alpha_i(t), \\ \alpha_i(t+1) \ &\geq \ \alpha_i(t) \ \Rightarrow \ \delta_i(t+3) = 0 \ \Rightarrow \alpha_i(t+1) \ > \ \alpha_i(t) - 0.0125 \end{aligned}$$

Figure 13 shows the relative positions between $\alpha_i(t)$ and $\alpha_i(t+1)$ that trigger the dimming section detector to fire or block the dimming section detector to fire.

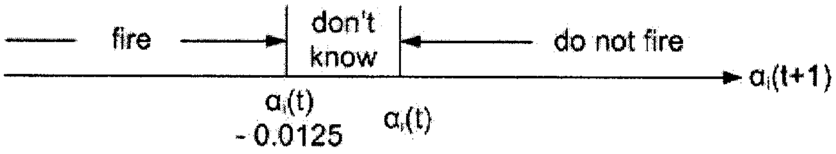

Figure 13: Dimming section detector fires or not based on the relative positions between $\alpha_i(t)$ and $\alpha_i(t+1)$

Note also that,

$$\nu_i(t+3) = 1 \ \Leftrightarrow \ x_i(t) = 0 \ \Leftrightarrow \ \lfloor 100\alpha_i(t) \rfloor = 0 \ \Leftrightarrow \ \alpha_i(t) < 0.01,$$

$$\nu_i(t+3) = 0 \;\Leftrightarrow\; x_i(t) \geq 1 \;\Leftrightarrow\; \lfloor 100\alpha_i(t) \rfloor \geq 1 \;\Leftrightarrow\; \alpha_i(t) \geq 0.01.$$

We conclude that the construction described in Section 7.4 accomplishes the objective set in Section 7.2 with $\Delta = 4$, $\varepsilon_1 = 0.01$ and $\varepsilon_2 = \varepsilon_3 = 0.0125$.

7.6 Generalizations

Let us use k to represent the gray-scale degrees, s to represent the number of photoreceptors in a section, and n to represent the number of sections in a receptive field.

7.6.1 The Objective

In Section 7.1, we defined the average brightness $\alpha(t)$ of receptive field at time t and the average brightness $\alpha_i(t)$ of the i-th section of receptive field at time t with $s = 400$ and $n = 20$. The generalized definitions of $\alpha(t)$ and $\alpha_i(t)$ are

$$\alpha(t) = \frac{1}{sn} \sum_{l=1}^{sn} \xi_l(t)$$

$$\alpha_i(t) = \frac{1}{s} \sum_{j=1}^{s} \xi_{s(i-1)+j}(t).$$

The generalized McCulloch-Pitts network, given the sn input functions ξ_j, computes an output function

$$y : \mathbf{N} \rightarrow \mathbf{R}$$

such that, for some positive integer Δ and for some positive numbers ε_1, ε_2, and ε_3, the output value $y(t+\Delta)$ obeys the following rules:

- if $\alpha(t+1) \geq \alpha(t)$, then $y(t+\Delta) = 0$;

- if $\alpha(t+1) \leq \alpha(t) - \varepsilon_2$,

  and if $\alpha_i(t+1) \leq \alpha_i(t) - \varepsilon_3$ for more than 10% of subscripts i such that $\alpha_i(t) \geq \varepsilon_1$, then $y(t+\Delta) = 1$;

- if $\alpha_i(t+1) \leq \alpha_i(t) - \varepsilon_3$ for at most 10% of subscripts i such that $\alpha_i(t) \geq \varepsilon_1$, then $y(t+\Delta) = 0$;

Under all other circumstances, not covered by this list of three, $y(t+\Delta)$ may be 1 or 0, and we do not care which it is.

7.6.2 Construction

The network begins with n encoding units. For each section $i = 1, 2, \ldots, n$, the encoding unit E_i receives s real numbers $\xi_{s(i-1)+j}(t)$ with $j = 1, 2, \ldots, s$ and, after a unit delay, outputs k bits, $e_{i\sigma}(t+1)$ with $\sigma = 1, 2, \ldots, k$, such that

$$\sum_{\sigma=1}^{k} e_{i\sigma}(t+1) = \left\lfloor \frac{k}{s} \sum_{j=1}^{s} \xi_{s(i-1)+j}(t) \right\rfloor .$$

This unit consists of k neurons and the σ-th neuron is defined by the threshold function

$$e_{i\sigma} = \begin{cases} 1 & \text{if } \frac{1}{s} \sum_{j=1}^{s} \xi_{s(i-1)+j} \geq \frac{1}{k}\sigma, \\ 0 & \text{otherwise.} \end{cases}$$

Next come n dimming section detectors and n dark section detectors. Each of these detectors is a single neuron.

For each section $i = 1, 2, \ldots, n$, the dimming section detector receives s real numbers $\xi_{s(i-1)+j}(t+1)$ with $j = 1, 2, \ldots, s$ from the photoreceptors in i-th section and k bits $e_{i\sigma}(t+1)$ with $\sigma = 1, 2, \ldots, k$ from the i-th encoding unit E_i; after a unit delay, it outputs a bit $\delta_i(t+2)$. This neuron is is defined by the threshold function

$$\delta_i = \begin{cases} 1 & \text{if } \dfrac{1}{k}\displaystyle\sum_{\sigma=1}^{k} e_{i\sigma} - \dfrac{1}{s}\displaystyle\sum_{j=1}^{s} \xi_{s(i-1)+j} \geq \dfrac{1}{s}, \\ 0 & \text{otherwise.} \end{cases}$$

For each section $i = 1, 2, \ldots, n$, the dark section detector receives one bit $e_{i\sigma}(t+1)$ and, after a unit delay, outputs a bit $\nu_i(t+2)$. This neuron is defined by the threshold function

$$\nu_i = \begin{cases} 1 & \text{if } -e_{i1} \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

The rest of the network consists of four neurons: Condition 1 unit, Condition 2 unit, memory neuron, and conjunction unit.

Condition 1 unit receives kn bits $e_{i\sigma}(t+1)$ with $i = 1, 2, \ldots, n$ and $\sigma = 1, 2, \ldots, k$ from all encoding units and sn real numbers $\xi_j(t+1)$ with $i = 1, 2, \ldots, sn$ from all photoreceptors. After a unit delay, it outputs a bit $c_1(t+2)$. This neuron is defined by the threshold function

$$c_1 = \begin{cases} 1 & \text{if } \dfrac{1}{kn}\displaystyle\sum_{i=1}^{n}\sum_{\sigma=1}^{k} e_{i\sigma} - \dfrac{1}{sn}\displaystyle\sum_{j=1}^{sn} \xi_j \geq \dfrac{1}{s}, \\ 0 & \text{otherwise.} \end{cases}$$

Condition 2 neuron receives n bits $\delta_i(t+2)$ with $i = 1, 2, \ldots, n$ from all dimming section detectors and n bits $\nu_i(t+2)$ with $i = 1, 2, \ldots, n$ from all dark section detectors. After a unit delay, it outputs a bit $c_2(t+3)$. Condition 2 unit is defined by the threshold function

$$c_2 = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} \delta_i + 0.1 \sum_{i=1}^{n} \nu_i \geq 1 + 0.1n, \\ 0 & \text{otherwise.} \end{cases}$$

The single memory neuron receives one bit $c_1(t+2)$ from Condition 1 neuron and, after a unit delay, outputs one bit $c_1'(t+3)$. This neuron is defined by the threshold function

$$c_1' = \begin{cases} 1 & \text{if } c_1 \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

Conjunction neuron receives one bit $c_1'(t+3)$ from the memory neuron and one bit $c_2(t+3)$ and, after a unit delay, outputs a bit $y(t+4)$ as the ultimate output of the network. This neuron is defined by the threshold function

$$y = \begin{cases} 1 & \text{if } c_1' + c_2 \geq 2, \\ 0 & \text{otherwise.} \end{cases}$$

7.6.3 Analysis

In Section 7.5, with $k = 100$, $x_i(t)$ is defined as an integer in range $[0, 100]$ which is encoded by $e_{i1}(t), e_{i2}(t), \ldots, e_{i100}(t)$. The generalized definition of $x_i(t)$ is

$$x_i(t) = \sum_{\sigma=1}^{k} e_{i\sigma}(t+1).$$

In this notation, we have

- $\sum_{\sigma=1}^{k} e_{i\sigma}(t+1) = x_i(t)$,
- $\delta_i(t+2) = 1$ if and only if

  $$\frac{1}{k}x_i(t) - \alpha_i(t+1) \geq \frac{1}{s},$$

  which implies

  $$x_i(t) - k\alpha_i(t+1) \geq \frac{k}{s},$$

- $\nu_i(t+2) = 1$ if and only if $x_i(t) = 0$,
- $c_1'(t+3) = c_1(t+2) = 1$ if and only if

  $$\frac{1}{kn}\sum_{i=1}^{n} x_i(t) - \alpha_i(t+1) \geq \frac{1}{s}$$

  which implies

  $$\sum_{i=1}^{n} x_i(t) - kn\alpha_i(t+1) \geq \frac{kn}{s}.$$

Since

$$x_i(t) = \lfloor k\alpha_i(t) \rfloor.$$

we have

$$k\alpha_i(t) - 1 \; < \; x_i(t) \; \leq \; k\alpha_i(t),$$

and so

$$kn\alpha(t) - n \; < \; \sum_{i=1}^{n} x_i(t) \; \leq \; kn\alpha(t).$$

If $c_1(t+2) = 1$, then

$$\sum_{i=1}^{n} x_i(t) \geq kn\alpha(t+1) + \frac{kn}{s},$$

and so

$$kn\alpha(t) \geq \sum_{i=1}^{n} x_i(t) \geq kn\alpha(t+1) + \frac{kn}{s} > kn\alpha(t+1).$$

If $c_1(t+2) = 0$, then

$$\sum_{i=1}^{n} x_i(t) < kn\alpha(t+1) + \frac{kn}{s},$$

and so

$$kn\alpha(t) < \sum_{i=1}^{n} x_i(t) + n < kn\alpha(t+1) + \frac{kn}{s} + n.$$

To summarize,

$$\alpha(t+1) \; \leq \; \alpha(t) - \left(\frac{1}{k} + \frac{1}{s}\right) \;\; \Rightarrow \;\; c_1(t+2) = 1 \;\; \Rightarrow \;\; \alpha(t+1) \; < \; \alpha(t),$$

$$\alpha(t+1) \; \geq \; \alpha(t) \;\; \Rightarrow \;\; c_1(t+2) = 0 \;\; \Rightarrow \;\; \alpha(t+1) \; > \; \alpha(t) - \left(\frac{1}{k} + \frac{1}{s}\right)$$

Figure 14 shows the relative positions between $\alpha(t)$ and $\alpha(t+1)$ that trigger Condition 1 unit to fire or block Condition 1 to fire.

In addition, we have

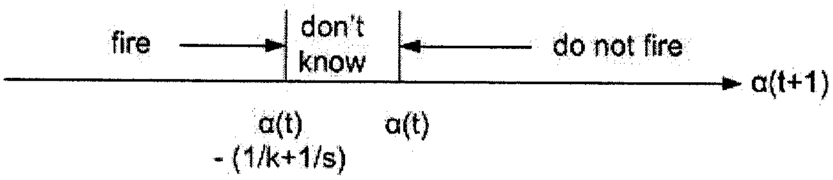

Figure 14: Condition 1 unit fires or not based on the relative positions between $\alpha(t)$ and $\alpha(t+1)$

- $c_2(t+3) = 1$ if and only if

$$\sum_{i=1}^{n} \delta_i(t+2) \geq 0.1 \left(n - \sum_{i=1}^{n} \nu_1(t+2) \right) + 1.$$

If $\delta_i(t+2) = 1$, then

$$x_i(t) \geq k\alpha_i(t+1) + \frac{k}{s},$$

and so

$$k\alpha_i(t) \geq x_i(t) \geq k\alpha_i(t+1) + \frac{k}{s} > k\alpha_i(t+1).$$

If $\delta_i(t+2) = 0$, then

$$x_i(t) \ < \ k\alpha_i(t+1) + \frac{k}{s},$$

and so

$$k\alpha_i(t) < x_i(t) + 1 < k\alpha_i(t+1) + \frac{k}{s} + 1.$$

To summarize,

$$\begin{aligned} \alpha_i(t+1) \ \leq \ \alpha_i(t) - \left(\frac{1}{k} + \frac{1}{s}\right) \ &\Rightarrow \ \delta_i(t+2) = 1 \ \Rightarrow \alpha_i(t+1) \ < \ \alpha_i(t), \\ \alpha_i(t+1) \ \geq \ \alpha_i(t) \ &\Rightarrow \ \delta_i(t+2) = 0 \ \Rightarrow \alpha_i(t+1) \ > \ \alpha_i(t) - \left(\frac{1}{k} + \frac{1}{s}\right) \end{aligned}$$

Figure 15 shows the relative positions between $\alpha_i(t)$ and $\alpha_i(t+1)$ that trigger the dimming section detector to fire or block the dimming section detector to fire.

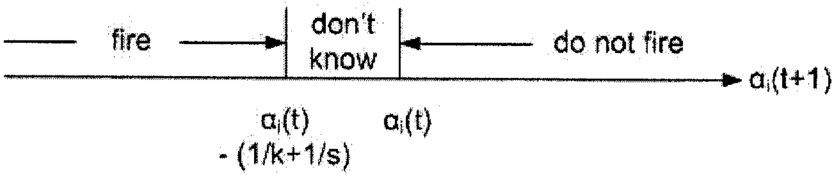

Figure 15: Dimming section detector fires or not based on the relative positions between $\alpha_i(t)$ and $\alpha_i(t+1)$

Note also that,

$$\nu_i(t+3) = 1 \Leftrightarrow x_i(t) = 0 \Leftrightarrow \lfloor k\alpha_i(t) \rfloor = 0 \Leftrightarrow \alpha_i(t) < \frac{1}{k},$$

$$\nu_i(t+3) = 0 \Leftrightarrow x_i(t) \geq 1 \Leftrightarrow \lfloor k\alpha_i(t) \rfloor \geq 1 \Leftrightarrow \alpha_i(t) \geq \frac{1}{k}.$$

We conclude that the construction describe in Section 7.6.2 accomplishes the objective set in Section 7.6.1 with $\Delta = 4$, $\varepsilon_1 = 1/k$ and $\varepsilon_2 = \varepsilon_3 = 1/k + 1/s$.

7.6.4 Parameters

Parameters k, s, and n are all positive integers such that

$$7000 \leq sn \leq 9000, \text{ and } \frac{k}{s} \ll 3.5.$$

In addition, one of the purposes of Condition 2 unit is to filter out the noise sections, which means that n cannot be too small (s cannot be too big). We require that $n \geq 10$. Moreover, in section 7.6.3, we analyzed that

- if $\alpha(t) - (1/k + 1/s) < \alpha(t+1) < \alpha(t)$, then $c1(t+2)$ can be either 0 or 1;

- if $\alpha_i(t) - (1/k + 1/s) < \alpha_i(t+1) < \alpha_i(t)$, then $\delta_i(t+2)$ can be either 0 or 1 with $i = 1, 2, \ldots, n$.

We need to make the error term $(1/k + 1/s)$ as small as possible. We can accept an error term smaller than 0.02.

8 Conclusions and Future Research

It was relief to find, after several failed attempts (some details of these attempts are given in Appendix A and Appendix B), a McCulloch-Pitts network which approximates the net dimming detectors of Lettvin et al. (1959) and fits (even though just barely) the constraints reported by Maturana (1959). Here are some of the conclusions that we have come to in the course of this work:

- Biological McCulloch-Pitts networks must approximate the inputs that their afferent neurons provide on continuous scale by numbers on discrete scale and they must encode the discrete approximations by sequences of bits. In this context, binary encodings seem absurd.
- Biological McCulloch-Pitts networks do not carry out exact computations, but (like the network we construct in Section 7) work within a small margin of error.
- Simple arithmetic suggests that, rather than being accessed individually, photoreceptors in the frog's retina are accessed in groups. It seems that the average number of photoreceptors in such a groups is on the order of hundreds and the number of these groups is on the order of thousands. Each receptive field is a set of such groups and each group may belong to many overlapping receptive fields.

There are two obvious directions for future research. One is to refine the work presented in this thesis to construct McCulloch-Pitts networks that simulate the net dimming detectors reported by Lettvin et al. The other direction is to construct McCulloch-Pitts networks that simulate the behavior of other ganglion cells.

Much more is known today about receptive fields than was known in 1959. Tessier-Lavigne (1991) is a survey of some of this progress.

References

Anthony, M. (2001). *Discrete mathematics of neural networks: Selected topics.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

Chow, C. K. (1961). Boolean functions realizable with single threshold devices. *Proceedings of the IRE*, *49*, 370–371.

Chvâtal, V., & Hammer, P. L. (1977). Aggregation of inequalities in integer programming. *Annals of Discrete Mathematics*, *1*, 145–162.

Elgot, C. C. (1961). Truth functions realizable by single threshold organs. *IEEE symposium on switching circuit theory and Logical design*, 225–245.

Gerstner, W., & Kistler, W. (2005). *Spiking neuron models single neurons, populations, plasticity.* Cambridge University Press, Cambridge, UK.

Hartline, H. K. (1938). The response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *American Journal of Physiology*, *121*, 400–415.

Hebb, D. (1949). *The organization of behavior: a neuropsychological theory.* Wiley-Interscience, New York, NY, USA.

Kandel, E. (1991). *Nerve cells and behavior*, Chap. 2 of Principles of Neural Science, pp. 18–33. Appleton and Lange, East Norwalk, CT, USA.

Lettvin, J. Y., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. (1959). What the frog's eye tells the frog's brain. *Proceedings of the IRE*, *47*, 1940–1951.

Maturana, H., Lettvin, J., McCulloch, W., & Pitts, W. (1960). Anatomy and physiology of vision in the frog. *The Journal of General Physiology*, *43*, 129–175.

Maturana, H. R. (1959). Number of fibres in the optic nerve and the number of ganglion cells in the retina of anurans. *Nature*, 183–1406.

McCulloch, W. S., & Pitts, W. H. (1943). A logical calculus of the ideas immanent in neural nets. *Bulletin of Mathematical Biophysics*, *5*, 115–133.

Parberry, I. (1994). *Circuit complexity and neural networks.* The MIT Press, Cambridge, MA, USA.

Purves, D., Angustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (Eds.). (2004). *Neuroscience.* Sinauer Associates, Inc., Sunderland, MA, USA.

Sui, K.-Y., Roychowdhury, V., & Kailath, T. (1995). *Discrete neural computation: A theorical foundation.* Prentice Hall PTR, Englewood Cliffs, NJ, USA.

Tessier-Lavigne, M. (1991). *Phototransduction and information processing in the retina*, Chap. 28 of Principles of Neural Science, pp. 400–420. Appleton and Lange, East Norwalk, CT, USA.

Appendix A Approximate Bit Counting

By a (c, ε)-*approximate bit counter* (where c and ε are positive constants) with n inputs and t outputs, we mean a McCulloch-Pitts network which, given input bits $x_1, x_2, \ldots, x_n$, computes output bits $y_1, y_2, \ldots, y_t$ such that

$$\frac{1}{t}\sum_{j=1}^{t} y_j \le \frac{c}{n}\sum_{i=1}^{n} x_i < \frac{1}{t}\sum_{j=1}^{t} y_j + \varepsilon. \tag{8}$$

We are interested in (c, ε)-approximate bit counters with huge n and small t, because they allow us to approximately compare sums of n bits: when the counter transforms inputs $x_1, x_2, \ldots, x_n$ into outputs $y_1, y_2, \ldots, y_t$ and transforms $x'_1, x'_2, \ldots, x'_n$ into outputs $y'_1, y'_2, \ldots, y'_t$, we can be sure that

$$\begin{aligned}
\frac{1}{t}\sum_{j=1}^{t} y'_j > \frac{1}{t}\sum_{j=1}^{t} y_j + \varepsilon \ &\Rightarrow\ \frac{1}{n}\sum_{i=1}^{n} x'_i > \frac{1}{n}\sum_{i=1}^{n} x_i,\\
\frac{1}{t}\sum_{j=1}^{t} y'_j < \frac{1}{t}\sum_{j=1}^{t} y_j - \varepsilon \ &\Rightarrow\ \frac{1}{n}\sum_{i=1}^{n} x'_i < \frac{1}{n}\sum_{i=1}^{n} x_i.
\end{aligned}$$

Theorem 1 *For every choice of positive integers d, s, t, n such that $n = s^d t$, there is a $(1, d/t)$-approximate bit counter with n inputs, t outputs, depth d, size*

$$(s^d + s^{d-1} + \cdots + 1)t,$$

and each neuron having fan-in st and fan-out t.

Proof. We are going to construct networks which satisfy

$$s^{d-1}\sum_{j=1}^{t} y_j \le \sum_{i=1}^{n} x_i < s^{d-1}(\sum_{j=1}^{t} y_j + d).$$

Let us describe the construction in a recursive way. When $d = 1$, the network consists of t neurons in parallel; its i-th neuron outputs

$$y_j = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} x_i \geq js, \\ 0 & \text{otherwise.} \end{cases}$$

When $d > 1$, we split the n input bits into s groups, each group having precisely $s^{d-1}t$ input bits. For each $r = 1, 2, \ldots, s$, let G_r denote a set of subscripts i such that input bit x_i belongs to the r-th group. Recursively, we construct s networks such that the r-th network, given input bits x_i $(i \in G_r)$, computes output bits $z_{(r-1)t+j}$ $(j = 1, 2, \ldots, t)$ such that

$$s^{d-1} \sum_{j=1}^{t} z_{(r-1)t+j} \leq \sum_{i \in G_r} x_i < s^{d-1}(\sum_{j=1}^{t} z_{(r-1)t+j} + (d-1)); \tag{9}$$

Then we add a copy of the network built when $d = 1$. This copy, given input bits z_i $(i = 1, 2, \ldots, st)$, computes output bits y_j $(j = 1, 2, \ldots, t)$ such that

$$s \sum_{j=1}^{t} y_j \leq \sum_{i=1}^{st} z_i < s(\sum_{j=1}^{t} y_j + 1). \tag{10}$$

From (9) and (10), we get

$$\sum_{i=1}^{n} x_i = \sum_{r=1}^{s} \sum_{i \in G_r} x_i \geq s^{d-1}(\sum_{r=1}^{s} \sum_{j=1}^{t} z_{(r-1)t+j}) \geq s^{d-1} \sum_{i=1}^{st} z_i = s^d \sum_{j=1}^{t} y_j$$

and

$$\begin{aligned}
\sum_{i=1}^{n} x_i = \sum_{r=1}^{s} \sum_{i \in G_r} x_i &< \sum_{r=1}^{s} (s^{d-1}(\sum_{j=1}^{t} z_{(r-1)t+j} + (d-1))) \\
&= s^{d-1} \sum_{r=1}^{s} \sum_{j=1}^{t} z_{(r-1)t+j} + s(d-1)s^{d-1}) \\
&= s^{d-1} \sum_{r=1}^{st} z_r + (d-1)s^d \\
&< s^{d-1}(s(\sum_{j=1}^{t} y_j + 1)) + (d-1)s^d \\
&= s^d(\sum_{j=1}^{t} y_j + d),
\end{aligned}$$

which proves (8) □

Theorem 2 *For every choice of positive integers n and t such that $n \geq 100t$, there is a $(c, 2/t)$-approximate bit counter with n inputs, t outputs, and $0.8 \leq c \leq 1$;*

- *this network has depth two,*
- *every neuron has fan-in at most $\sqrt{1.25nt}$,*
- *every neuron has fan-out t.*

Proof. For every real number x such that $x \geq 100$, the interval $[x, 1.25x]$ includes at least one square of an integer. In particular, the interval $[n/t, 1.25n/t]$ includes the square of some integer s. Let us write $N = s^2t$ and note that

$$n \leq N \leq 1.25n.$$

By Theorem 1, there is a $(1, 2/t)$-approximate bit counter with N inputs and t outputs, where the depth is two, every neuron has fan-in s, and every neuron has fan-out t. This counter provides an $(n/N, 2/t)$-approximate bit counter with input bits $x_1, x_2, \ldots, x_n$: given input bits $x_1, x_2, \ldots, x_n$, we set $x_i = 0$ whenever $n \leq i \leq N$ and observe that

$$\frac{1}{N}\sum_{i=1}^{N} x_i = \frac{n/N}{n}\sum_{i=1}^{n} x_i.$$

□

Theorem 3 *For every choice of positive integers* n *and* t *such that* $n \geq 2400t$, *there is a* $(c, 3/t)$*-approximate bit counter with* n *inputs,* t *outputs, and* $0.8 \leq c \leq 1$;

- *this network has depth three,*
- *every neuron has fan-in at most* $\sqrt[3]{1.25nt^2}$,
- *every neuron has fan-out* t.

Proof. For every real number x such that $x \geq 2400$, the interval $[x, 1.25x]$ includes at least one cube of an integer. In particular, the interval $[n/t, 1.25n/t]$ includes the cube of some integer s. Let us write $N = s^3 t$ and note that

$$n \leq N \leq 1.25n.$$

By Theorem 1, there is a $(1, 3/t)$-approximate bit counter with N inputs and t outputs, where the depth is three, every neuron has fan-in st, and every neuron has fan-out t. This counter provides an $(n/N, 3/t)$-approximate bit counter with input bits $x_1, x_2, \ldots, x_n$: given input bits $x_1, x_2, \ldots, x_n$, we set $x_i = 0$ whenever $n \leq i \leq N$ and observe that

$$\frac{1}{N}\sum_{i=1}^{N} x_i = \frac{n/N}{n}\sum_{i=1}^{n} x_i.$$

□

Before we came up with the idea of that photoreceptors must be accessed in groups, there were two critical issues we need to work out: one was that the size of the network exceeded the the total number of neurons in the retina; the other one was that the fan-in of a McCulloch-Pitts neuron exceeded the upper bound of the total number of pre-synaptic cells a single neuron can have.

The approximate bit counter was devised by my supervisor, Dr. Vašek Chvátal, to count a huge number (more than $100,000$) of bits (this task cannot be done by a single neuron).

After we divided the photoreceptors in groups, the number of inputs are lowered below the constraint number ($10,000$). Therefore, we can use a single neuron to accomplish the counting.

Appendix B Assumability

Definition 1 *For a positive integer* m*, a Boolean function* f *is called* m-***summable*** *if there are (not necessarily distinct) zero-one vectors* $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_k}$ *and* $\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_k}$ *such that* $1 \leq k \leq m$ *and*

$$\sum_{i=1}^{k} \mathbf{x_i} = \sum_{i=1}^{k} \mathbf{y_i}, \quad \text{and} \quad f(\mathbf{x_i}) = 1, \quad f(\mathbf{y_i}) = 0 \quad \text{for all } i = 1, 2, \ldots, k,$$

If f *is not* m*-summable, then* f *is called* m-***assumable****.*

Theorem 4 *(Elgot, 1961)(Chow, 1961) A Boolean function* f *is a linear threshold function if and only if it is* m*-assumable for every* m*.*

Proof. The "only if" part. Let $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_k}$ and $\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_k}$ be zero-one vectors such that $f(\mathbf{x_i}) = 1$ for all i and $f(\mathbf{y_j}) = 0$ for all j. If f is a threshold function, represented by weight vector $\mathbf{w}$ and and threshold θ, then we have $\mathbf{w}^T\mathbf{x_i} \geq \theta$ for all i and $\mathbf{w}^T\mathbf{y_j} < \theta$ for all j. Therefore, we have

$$\mathbf{w}^T \sum_{i=1}^{k} \mathbf{x_i} \geq k\theta > \mathbf{w}^T \sum_{i=j}^{k} \mathbf{y_j},$$

which implies that

$$\sum_{i=1}^{k} \mathbf{x_i} \neq \sum_{j=1}^{k} \mathbf{y_j}.$$

The "only if" part. We will use the following two theorems:

(i) A system:

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i \quad \text{for all } i = 1, 2, ..., m \tag{11}$$

is unsolvable if and only if there are non-negative numbers $u_1, u_2, \ldots, u_m$ such that

$$
\begin{aligned}
&u_i \geq 0 \quad \text{for all} \quad i = 1, 2, \ldots, m, \\
&\sum_{i=1}^{m} a_{ij} u_i = 0 \quad \text{for all} \quad j = 1, 2, \ldots, n, \\
&\sum_{i=1}^{m} b_i u_i < 0.
\end{aligned}
$$

(ii) If a system $\mathbf{A}\mathbf{x} = 0, \mathbf{x} \geq \mathbf{0}, \mathbf{x} \neq \mathbf{0}$ is solvable and if $\mathbf{A}$ has rational entries, then there exists a solution $\mathbf{x}$ with integer entries.

We want to prove that if f is m-assumable, then f is a threshold function. This is equivalent to proving that if f is not a threshold function, then f is summable.

Let $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_s}$ and $\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_t}$ be vectors in $\{0,1\}^n$ such that $f(\mathbf{x_i}) = 1$ for all i and $f(\mathbf{y_j}) = 0$ for all j. If f is not a threshold function, then the system

$$
\begin{aligned}
&\mathbf{w}^T \mathbf{x_i} \geq \theta, \text{for all } i = 1, 2, \ldots, s \\
&\mathbf{w}^T \mathbf{y_j} < \theta, \text{for all } j = 1, 2, \ldots, t
\end{aligned}
\tag{12}
$$

(in variables $w_1, w_2, \ldots, w_n$ and θ) is unsolvable. By scaling, system (12) is unsolvable if and only if system

$$
\begin{aligned}
-\mathbf{w}^T \mathbf{x_i} + \theta &\leq -1, \quad \text{for all } i = 1, 2, \ldots, s \\
\mathbf{w}^T \mathbf{y_j} - \theta &\leq 0, \qquad \text{for all } j = 1, 2, \ldots, t
\end{aligned}
\tag{13}
$$

is unsolvable.

By theorem (i), if system (13) is unsolvable, then there exist nonnegative numbers $\lambda_1, \lambda_2, \dots, \lambda_s$ and $\mu_1, \mu_2, \dots, \mu_t$ such that

$$\begin{aligned}
&\sum_{i=1}^{s} \lambda_i x_{ir} - \sum_{j=1}^{t} \mu_j y_{jr} = 0 \quad \text{for all} \quad r = 1, 2, \dots, n, \\
&\sum_{i=1}^{s} \lambda_i + \sum_{j=1}^{t} (-\mu_j) = 0, \\
&\lambda_1, \lambda_2, \dots, \lambda_s, \mu_1, \mu_2, \dots, \mu_t \geq 0.
\end{aligned} \tag{14}$$

By theorem (ii), there exist integers $\lambda_1, \lambda_2, \dots, \lambda_s$ and $\mu_1, \mu_2, \dots, \mu_t$ with the same properties.

Writing $\mathbf{x_i}^T = (x_{i1}, x_{i2}, \dots, x_{in})$ for all i and $\mathbf{y_j}^T = (y_{j1}, y_{j2}, \dots, y_{jn})$ for all j, we record system (14) as

$$\left.\begin{aligned}
\sum_{i=1}^{s} \lambda_i \mathbf{x_i} &= \sum_{j=1}^{t} \mu_j \mathbf{y_j}, \\
\sum_{i=1}^{s} \lambda_i &= \sum_{j=1}^{t} \mu_j.
\end{aligned}\right\} \tag{15}$$

With $k = \sum_{i=1}^{s} \lambda_i$, system (15) can be spread out as:

$$\begin{aligned}
&\overbrace{\underbrace{x_1 + \cdots + x_1}_{\lambda_1} + \underbrace{x_2 + \cdots + x_2}_{\lambda_2} + \cdots + \underbrace{x_s + \cdots + x_s}_{\lambda_s}}^{k} \\
= &\overbrace{\underbrace{y_1 + \cdots + y_1}_{\mu_1} + \underbrace{y_2 + \cdots + y_2}_{\mu_2} + \cdots + \underbrace{y_s + \cdots + y_s}_{\mu_t}}^{k},
\end{aligned} \tag{16}$$

From Equation (16), we can see that f is m-summable for all m such that $m \geq k$. □

Anthony (2001) and Chvátal and Hammer (1977) gave proofs for this theorem from two different angles. According to their ideas, Dr. Chvátal helped me to rewrite the proof trying to make it easier to be understood.

Before we stabilized our net work model, we had attempted to use Boolean circuits to implement the dimming detector. The critical issue of the Boolean circuit is that the depth of the network is very large which means there is a huge delay in the circuit.

Threshold gates can be applied in Boolean circuits to efficiently reduce the depth of the circuits. The theorem in this appendix provides mathematical criteria to determine which Boolean circuit can be replaced by a threshold gate.