Lectures on Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components

John von Neumann (1956)

Presented by: Andrew Leverentz and Anthony Thomas
UC San Diego

November 15, 2018
Summary

- How do we design a reliable system from unreliable components?
- We’ll consider Boolean circuits (“organisms”) built from simple logic gates (“organs”)
- Each “organ” produces an incorrect output with some non-zero probability
- How do we ensure overall system reliability when no single component can be trusted?
- Spoiler alert:
Summary

- How do we design a reliable system from unreliable components?
- We’ll consider Boolean circuits ("organisms") built from simple logic gates ("organs")
- Each “organ” produces an incorrect output with some non-zero probability
- How do we ensure overall system reliability when no single component can be trusted?
- Spoiler alert:

Redundancy to the Rescue

By introducing redundancy to a sufficiently high degree we can drive the probability of error to zero.
Introducing the Model

- An “organism” is a Boolean circuit composed of “organs” which are a generalization of logic gates.
Introducing the Model

- An “organism” is a Boolean circuit composed of “organs” which are a generalization of logic gates.

Organs:
- Take an arbitrary number of Boolean inputs at time $t$
- Produces a single Boolean output at time $t + \delta$

Figure 1: A Simple Organ
More on Organs

- We can design organs which compute basic Boolean functions: and, or, not, etc...
- And organs which maintain state over time:
More on Organs

- We can design organs which compute basic Boolean functions: and, or, not, etc...
- And organs which maintain state over time:

Figure 2: A Simple Memory Organ
Universal Organs

- Some organs are “universal”

- We can represent any Boolean function using only a single universal organ

- NAND is universal: we can construct any Boolean function using NAND gates

- Note: Von Neumann calls NAND the “Scheffer Stroke”

- Examples:
  - NOT(x) = NAND(x,x)
  - AND(x,y) = NOT(NAND(x,y))
  - OR(x,y) = NAND(NOT(x), NOT(y))
Universal Organs

- Some organs are “universal”
  - We can represent any Boolean function using only a single universal organ

Note: Von Neumann calls NAND the “Scheffer Stroke”

Examples:
- \( \text{NOT}(x) = \text{NAND}(x,x) \)
- \( \text{AND}(x,y) = \text{NOT}(%(x), NOT(y)) \)
- \( \text{OR}(x,y) = \text{NAND}(\text{NOT}(x), \text{NOT}(y)) \)
Universal Organs

- Some organs are “universal”
  - We can represent any Boolean function using only a single universal organ
- NAND is universal: we can construct any Boolean function using NAND gates

Figure 3: A NAND Organ
Universal Organs

- Some organs are “universal”
  - We can represent any Boolean function using only a single universal organ
- NAND is universal: we can construct any Boolean function using NAND gates
- Note: Von Neumann calls NAND the “Scheffer Stroke”

Figure 3: A NAND Organ
Universal Organs

- Some organs are “universal”
  - We can represent any Boolean function using only a single universal organ
- NAND is universal: we can construct any Boolean function using NAND gates
- Note: Von Neumann calls NAND the “Scheffer Stroke”
- Examples:
  \[
  \text{NOT}(x) = \text{NAND}(x, x) \\
  \text{AND}(x, y) = \text{NOT}(\text{NAND}(x, y)) \\
  \text{OR}(x, y) = \text{NAND}(\text{NOT}(x), \text{NOT}(y))
  \]

**Figure 3: A NAND Organ**
Introducing an Error Model

- In reality these “organs” will not function perfectly
- We assume that every organ can produce an incorrect output with iid probability $\epsilon$
- Let’s spend some time looking at how error effects a Boolean circuit.
The Role of Error: Example

- Assume $\epsilon < \frac{1}{2}$
- Let’s recall the simple “memory circuit” as an example:

Figure 4: A Simple Memory Organ
Analyzing Error in the Memory Organ

- Suppose the organ is stimulated at time $t$. Then it should continue to be excited in perpetuity.
- But with probability $\epsilon$ it flips its output state.
Analyzing Error in the Memory Organ

- Suppose the organ is stimulated at time $t$. Then it should continue to be excited in perpetuity.
- But with probability $\epsilon$ it flips its output state.
- Then the probability that the organ is stimulated at $t = s + 1$ is given by the recursion:

$$
\rho_{s+1} = (1 - \epsilon)\rho_s + \epsilon(1 - \rho_s)
$$

$$
\rho_{s+1} - \frac{1}{2} = (1 - 2\epsilon)(\rho_s - \frac{1}{2})
$$

$$
\Rightarrow \rho_s - \frac{1}{2} = (1 - 2\epsilon)^s(\rho_0 - \frac{1}{2}) \approx e^{-2\epsilon s}(\rho_0 - \frac{1}{2})
$$

- Then letting $s \to \infty$ the RHS goes to zero so:
Analyzing Error in the Memory Organ

- Suppose the organ is stimulated at time $t$. Then it should continue to be excited in perpetuity.
- But with probability $\epsilon$ it flips its output state.
- Then the probability that the organ is stimulated at $t = s + 1$ is given by the recursion:

$$
\rho_{s+1} = (1 - \epsilon)\rho_s + \epsilon(1 - \rho_s)
$$

$$
\rho_{s+1} - \frac{1}{2} = (1 - 2\epsilon)(\rho_s - \frac{1}{2})
$$

$$
\Rightarrow \rho_s - \frac{1}{2} = (1 - 2\epsilon)^s(\rho_0 - \frac{1}{2}) \approx e^{-2\epsilon s} (\rho_0 - \frac{1}{2})
$$

- Then letting $s \rightarrow \infty$ the RHS goes to zero so:

$$
\lim_{s\rightarrow\infty} e^{-2\epsilon s} (\rho_0 - \frac{1}{2}) = 0 \Rightarrow \rho_s \rightarrow \frac{1}{2}
$$
Interpreting Error

- Stated another way...

**Error Leads to Irrelevance**

After many cycles, the memory of the organ is useless since it has a roughly equal probability of being right or wrong. Allowing error leads to long run irrelevance of an organ.
Interpreting Error

- Stated another way...

<table>
<thead>
<tr>
<th>Error Leads to Irrelevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>After many cycles, the memory of the organ is useless since it has a roughly equal probability of being right or wrong. Allowing error leads to long run irrelevance of an organ.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controlling Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>We need a mechanism for controlling the long run probability of error!</td>
</tr>
</tbody>
</table>
Multiplexing: Controlling Error through Redundancy

- We’ll show that we can reduce the probability of error by carefully introducing redundancy.
- We now consider $N$ iid “copies” of an organ as our atomic unit when constructing organisms.
  - I/O from an organ is now carried on a “bundle” of $N$ lines.
- If the probability of error is not too high, then hopefully most of these lines will be correct.
- For now, we assume the only source of error is in the input - we’ll relax this later.
- Let’s look at a couple examples and then formalize this.
Multiplexing Example

Simple majority gate

Multiplexed majority

Note: With multiplexing, a majority of the output can be correct, even if the inputs are noisy.
Another Example

Simple majority gate

```
0
1
1
1
```

Multiplexed majority

```
0
0
0
1
1
1
1
1
1
1
```

- It’s possible to have an incorrect output due to unlucky combinations of noisy inputs.
Error in Multiplexed Organs

Takeaway: Multiplexing alone is not enough
Error in Multiplexed Organs

**Takeaway: Multiplexing alone is not enough**

- Even if most inputs are correct, we can still end up with incorrect output after many successive gates
Takeaway: Multiplexing alone is not enough

- Even if most inputs are correct, we can still end up with incorrect output after many successive gates
- We can think of this as “signal attenuation”
- We need a way to boost the signal regularly (after each organ)
- We’ll add a new organ called the “restoring organ” to do this
The Restoring Organ

- The majority organ already does this for us!

Figure 5: Restoration for the Majority Organ
Let’s formalize the intuition developed above and do some analysis.
The Restoring Organ

- Let’s formalize the intuition developed above and do some analysis
- Suppose \( \alpha N \) of the incoming lines are stimulated
- We want to estimate the fraction of output lines which are stimulated
- The probability that any majority organ is stimulated is then given by:

\[
\alpha^* = \binom{3}{3} \alpha^3 + \binom{3}{2} \alpha^2 (1 - \alpha) \\
= 3\alpha^2 - 2\alpha^3
\]

- Thus about \( \alpha^* N \) output lines will be stimulated. We’ll look at the distribution of this quantity later...
The Restoring Organ

- If $\alpha < \frac{1}{2} \Rightarrow \alpha^* < \alpha$
- If $\alpha > \frac{1}{2} \Rightarrow \alpha^* > \alpha$

Figure 6: Restoration for the Majority Organ
Wrapping up Restoration

Takeaway

If we chain several such “restoring organs” in succession we can increase the probability the organ functions according to its desired state.
Wrapping up Restoration

Takeaway

If we chain several such “restoring organs” in succession we can increase the probability the organ functions according to its desired state.

A Remaining Problem

So far we’ve assumed the only source of error was elsewhere in the system. What if components of our multiplexed organ have $\epsilon > 0$?
Building Reliable Circuits

Goals

- Design a reliable version of a “universal organ”: a multiplexed NAND gate
- Show mathematically that the probability of overall malfunction is low, even if individual components malfunction
Multiplexed NAND, Version 1: Without Restoration

Construct a multiplexed NAND gate
(similar to the multiplexed majority gate)
Multiplexed NAND, Version 1: Without Restoration

Construct a multiplexed NAND gate
(similar to the multiplexed majority gate)

NOTE: Still susceptible to “signal attenuation”
We’ll add signal restoration later.
Perfectly-functioning multiplexed NAND gate

- Assume we have a multiplexed NAND gate in which the individual NAND gates do not malfunction.
- Only source of error is from noisy inputs.

\[
\zeta \sim \text{Normal}\left(1 - \xi \eta, \xi (1 - \xi) \eta (1 - \eta) \right)
\]

where

- \(\xi\) = proportion “on” in input 1
- \(\eta\) = proportion “on” in input 2
- \(\zeta\) = proportion “on” in output
Perfectly-functioning multiplexed NAND gate

- Assume we have a multiplexed NAND gate in which the individual NAND gates do not malfunction.
- Only source of error is from noisy inputs.
- Using a combinatorial argument and large-$N$ approximations, the output is approximately normally distributed:

\[
\zeta \sim \text{Normal} \left( 1 - \xi \eta, \frac{\xi(1 - \xi)\eta(1 - \eta)}{N} \right)
\]

where

\[
\xi = \text{proportion “on” in input 1}
\]
\[
\eta = \text{proportion “on” in input 2}
\]
\[
\zeta = \text{proportion “on” in output}
\]
Reintroducing Error

- Suppose each gate malfunctions with probability $\epsilon$
Reintroducing Error

- Suppose each gate malfunctions with probability $\epsilon$
- Suppose we have some configuration of input-line activations such that the correct output behavior is
  - $r$ outputs are off
  - $N - r$ outputs are on

Two types of error are approximately normally distributed:

- $(\text{off} \rightarrow \text{on}) \sim \text{Normal}(\epsilon r, \epsilon (1 - \epsilon) r)$
- $(\text{on} \rightarrow \text{off}) \sim \text{Normal}(\epsilon (N - r), \epsilon (1 - \epsilon) (N - r))$

Binomial distribution $\rightarrow$ Normal distribution as $N \rightarrow \infty$
Reintroducing Error

- Suppose each gate malfunctions with probability $\epsilon$
- Suppose we have some configuration of input-line activations such that the correct output behavior is
  - $r$ outputs are off
  - $N - r$ outputs are on
- Two types of error are approximately normally distributed:
  - $(\#\text{ off } \rightarrow \text{ on}) \sim \text{Normal}(\epsilon r, \epsilon(1 - \epsilon)r)$
  - $(\#\text{ on } \rightarrow \text{ off}) \sim \text{Normal}(\epsilon(N - r), \epsilon(1 - \epsilon)(N - r))$

Binomial distribution $\rightarrow$ Normal distribution as $N \rightarrow \infty$
Output activation (with errors)

- Accounting for errors introduced by the multiplexed NAND, the actual output activation will be

\[
\zeta' = \zeta + \frac{(\# \text{ off} \rightarrow \text{on}) - (\# \text{ on} \rightarrow \text{off})}{N}
\]
Output activation (with errors)

- Accounting for errors introduced by the multiplexed NAND, the actual output activation will be

\[ \zeta' = \zeta + \frac{(\# \text{ off} \rightarrow \text{ on}) - (\# \text{ on} \rightarrow \text{ off})}{N} \]

- Since this is a sum of (approximately) normally-distributed variables, we have

\[ \zeta' \sim \text{Normal}\left((1 - \xi \eta) + 2\epsilon(\xi \eta - 1/2), \frac{(1 - 2\epsilon)^2 \xi (1 - \xi) \eta (1 - \eta) + \epsilon(1 - \epsilon)}{N}\right) \]
Multiplexed NAND, Version 2: With Restoration

By adding “restorative” organs, a reliable NAND gate can be implemented as:

Note: For NAND gates, two restorative organs are needed, because each one inverts the result
Behavior of executive & restorative organs

- A “restorative organ” is just like a multiplexed NAND, except the inputs come from a bundle which has been split & scrambled.
  Special case: input proportion 1 = input proportion 2.
Behavior of executive & restorative organs

- A “restorative organ” is just like a multiplexed NAND, except the inputs come from a bundle which has been split & scrambled.
  Special case: input proportion 1 = input proportion 2.
- To describe the behavior of the full multiplexed NAND, we can apply the previous results three times (1 executive organ + 2 restorative organs)
Quantitative results

Given input activations, $\xi$ and $\eta$, what are the distributions of the intermediate activations ($\zeta$ and $\omega$) and the output activation $\psi$?

$$
\begin{align*}
\zeta & \sim \text{Normal}(\mu(\xi, \eta), \sigma^2(\xi, \eta)) \\
\omega & \sim \text{Normal}(\mu(\zeta, \zeta), \sigma^2(\zeta, \zeta)) \\
\psi & \sim \text{Normal}(\mu(\omega, \omega), \sigma^2(\omega, \omega)) \\
\mu(a, b) &= (1 - ab) + 2\epsilon(ab - 1/2) \\
\sigma^2(a, b) &= \frac{(1 - 2\epsilon)^2 ab(1 - a) b(1 - b) + \epsilon(1 - \epsilon)}{N}
\end{align*}
$$
Reliability: A Mathematical Definition

Our compound NAND gate is reliable if there exists a threshold $\Delta$ such that, for input activations $\xi$ and $\eta$, the output activation $\psi$ satisfies the following relations with high probability:

- $\xi \geq 1 - \Delta$ and $\eta \geq 1 - \Delta \Rightarrow \psi \leq \Delta$ [NAND($T, T$) = $F$]
- $\xi \geq 1 - \Delta$ and $\eta \leq \Delta \Rightarrow \psi \geq 1 - \Delta$ [NAND($T, F$) = $T$]
- $\xi \leq \Delta$ and $\eta \geq 1 - \Delta \Rightarrow \psi \geq 1 - \Delta$ [NAND($F, T$) = $T$]
- $\xi \leq \Delta$ and $\eta \leq \Delta \Rightarrow \psi \geq 1 - \Delta$ [NAND($F, F$) = $T$]
Reliability of Multiplexed NAND

If $N$ is sufficiently large, then with high probability

\[
\begin{align*}
\zeta & \approx (1 - \xi \eta) + 2\epsilon(\xi \eta - 1/2) \\
\omega & \approx (1 - \zeta^2) + 2\epsilon(\zeta^2 - 1/2) \\
\psi & \approx (1 - \omega^2) + 2\epsilon(\omega^2 - 1/2)
\end{align*}
\]

Claim: This is sufficient to guarantee the mathematical definition of reliability.
Example: The Effect of Signal Restoration

Numerical results for $\Delta = 0.1$ and $\epsilon = 0.005$ (ignoring variance by taking limit as $N \to \infty$)
Blue: output $\in [0, \Delta]$, Red: output $\in [1 - \Delta, 1]$

Restoration ensures that unambiguous inputs will not lead to ambiguous outputs
References

  http://www.sns.ias.edu/pitp2/2012files/Probabilistic_Logics.pdf