

AI: Winter of Our Discontent

Richard P. Gabriel



October 17, 2024

Mystery

Only mystery allows us to live ...
... only mystery
–*Federico Garcia Lorca*



rpg Timeline: 1949–1972

- October 31, 1949: Born, Haverhill, Massachusetts; grew up on a farm in Merrimac
- 1955–1960: Merrimac Elementary School, Merrimac, Massachusetts
- 1960–1967: Pentucket Regional High School, West Newbury, Massachusetts
- 1966–1972: Filmed football games for Pentucket's Coach;
a pro filming a rival team offered me a college / pro job
- 1967: Applied to Northeastern, MIT, Harvard; got into Northeastern
- 1967–1972
 - **Northeastern University, Mathematics**
 - **Played Hammond in rock 'n' roll band; *Alewife Brook Parkway* & *Rachel***
 - Parke Mathematical Laboratory (coöp job)
- 1968: Learned to program Fortran; first program: band name creator

Leonard: "Is that *the* Richard Gabriel?"

Jon: "Yeah, although I'm not sure what he's known for."

rpg Timeline: 1949–1972

Rachel

John Carver, drums
Jake Lord, guitar
Bob Valyou, guitar & vocals
John Ferrara, bass
Phil Doucette, vocals
rpg, Hammond



TONIGHT
"VELVET UNDERGROUND"
AND "RACHEL"
Adm. \$2.00

August 2, 1969

rpg Timeline: 1972–1973

- 1972
 - Tried to apply to Oxford for grad school; could not understand the application
 - **Met Harriet Fell, who told me my math / philosophy interest = Artificial Intelligence**
 - Harriet introduced me to Patrick Winston at the MIT AI Lab
 - **I taught myself Lisp**
 - Weird computer vision formalism paper earns invitation to “visit” AI Lab for a year
 - MIT “Special Student” in Math Department for one year
 - **Met Guy Steele**, Jonl White, Scott Fahlman, Marvin Minsky, Seymore Papert, Mitchell Wand, Carl Hewitt, Richard Greenblatt, **Terry Winograd**, **David Waltz**
- 1973: Lighthill Report reduces MIT AI funding (AI Winter): “No soup for you”; **Patrick Winston arranges for David Waltz to take me to the University of Illinois to start an AI Lab.**

The Lighthill Report: First AI Winter

*Published in 1973, it was compiled by Professor Sir James Lighthill for the British Science Research Council as an evaluation of academic research in **the field of artificial intelligence (AI)**. The report gave a very pessimistic prognosis for many core aspects of research in this field, stating that **“In no part of the field have the discoveries made so far produced the major impact that was then promised.”***

—Wikipedia, “Lighthill Report”



The Lighthill Report: First AI Winter

*It “formed the basis for the decision by the British government to end support for AI research in most British universities.” While the report was supportive of research into the simulation of neurophysiological and psychological processes, it was “highly critical of basic research in foundational areas such as robotics and language processing.” The report stated that AI researchers had failed to address the issue of combinatorial explosion when solving problems within real-world domains. That is, the report states that **AI techniques may work within the scope of small problem domains, but the techniques would not scale up well to solve more realistic problems. The report represents a pessimistic view of AI that began after early excitement in the field.***

—Wikipedia, “Lighthill Report”



The Lighthill Report: First AI Winter

[Category A,] *Advanced Automation: the clear objective of this category of work being to replace human beings by machines for specific purposes, which may be industrial or military on the one hand, and mathematical or scientific on the other. The work looks beyond automation of the type that is widely adopted at present in control engineering and data processing, and aims to make a far fuller use of the general-purpose digital computer's logical (as opposed to arithmetical) potentialities. Nevertheless it must be looked at as a natural extension of previous work on the automation of human activities, and be judged by essentially the same criteria.*

Industrially important purposes include, for example, *machine recognition of printed or typewritten characters (an area where good progress has been made) and of handwritten characters (incomparably more difficult), as well as a much wider range of pattern-recognition activities.*

—James Lighthill, “Artificial Intelligence: A General Survey”



The Lighthill Report: First AI Winter

[Category C,] *Computer-based CNS research. In a lay person's report the extended term **central nervous system** is used in preference to the term **brain** which to a lay person may have subjective associations with the more conscious, or more consciously **brainy**, parts of the brain activity, as against the emotional parts or those associated with perception and movement. The co-ordination of perception and movement in animals generally is a particularly significant area of research which the lay person (unlike the biologist) might be tempted to forget if the word 'brain' conjures up for him the specifically human aspects of brain activity.*

Category C is concerned, then, with *theoretical investigations related to neurobiology and to psychology.*

—James Lighthill, "Artificial Intelligence: A General Survey"



The Lighthill Report: First AI Winter

[Category B stands] *not only [for] Bridge activity, but also for the basic component of that activity: Building Robots. The whole concept of Building Robots is, indeed, seen as an essential Bridge Activity justified primarily by what it can feed into the work of categories A and C, and by the links that it creates between them.*

*Thus, a Robot in the sense used here... is **an automatic device that mimics a certain range of human functions without seeking in any useful sphere of human activity to replace human beings.***

Work [here...] is frequently justified because it simultaneously supports category A (Advanced Automation), in the sense that generalised information on automatic devices may emerge which can be used in practical problems of Automation, and supports category C (Computer-based CNS studies), in the sense that devices that mimic a human function may assist in studying, and in making a theory of, that function.

—James Lighthill, “Artificial Intelligence: A General Survey”



The Lighthill Report: First AI Winter

- Hubert Dreyfus, “What Computers Can’t Do: The Limits of Artificial Intelligence”
- John Searle, “Minds, Brains and Programs”—the Chinese Room Argument

Nature of AI, 1970s

The pursuit of AI is to construct a **machine**—in software, hardware, or both combined—that is **kind of like a human in some ways**. But “kind of like a human in some ways” is up for interpretation—**you can never be sure it's been achieved**. That's the rub. With scientific / algorithmic programming you have a decent and usually perfect idea of correctness; even in business applications, mostly you can tell whether the programs are working.

... the idea of 'intelligence' is itself emotional rather than mathematical.

—Alan Turing, “Intelligent Machinery”

*The engineering problem of producing various machines for various jobs is replaced by **the office work of 'programming'** the universal machine to do these jobs.*

—Alan Turing, “Intelligent Machinery”

Dartmouth Summer Research Project

On Artificial Intelligence

We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that *every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it*. An attempt will be made to find how to make machines use language, form abstractions and concepts, *solve kinds of problems now reserved for humans, and improve themselves*. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer.

- John McCarthy
- Marvin Minsky
- Nathaniel Rochester
- Claude Shannon

Dartmouth Summer Research Project

On Artificial Intelligence

- If a machine can do a job, then an **automatic calculator** can be programmed to simulate the machine. The speeds and memory capacities of present computers may be insufficient to simulate many of the higher functions of the human brain, but **the major obstacle is not lack of machine capacity, but our inability to write programs** taking full advantage of what we have.
- It may be speculated that **a large part of human thought consists of manipulating words according to rules of reasoning and rules of conjecture**. From this point of view, forming a generalization consists of admitting a new word and some rules whereby sentences containing it imply and are implied by others.

Dartmouth Summer Research Project

On Artificial Intelligence

- “How can a set of (hypothetical) **neurons be arranged so as to form concepts.**”
- “Probably a truly intelligent machine will carry out activities which may best be described as **self-improvement.**”
- “A fairly attractive and yet clearly incomplete conjecture is that **the difference between creative thinking and unimaginative competent thinking lies in the injection of a [sic] some randomness.** The randomness must be guided by intuition to be efficient. In other words, **the educated guess or the hunch include controlled randomness in otherwise orderly thinking.**”

Dartmouth Summer Research Project

On Artificial Intelligence

imaginative, invention, discovery, uncertainty, failure, slightly wrong, unreasonable, unexpected, randomness, hunch, Monte Carlo, rough guess, “for the machine, randomness will probably be needed to overcome the short-sightedness and prejudices of the programmer,” originality, self-reference, chess machine, chess language, logic language, learning

The fundamental research strategy of this group contemplates work on both models and theories. In the science of very complex information processing systems, we are a long way from even knowing what questions to ask or what aspects to abstract for theory. The present need is for a large population of concrete systems that are completely understood and thereby provide a base for induction.

—Allen Newell & Herbert Simon

Nature of AI, 1970s

- Notice that computer science was not yet quite invented
- As with climbing a mountain with false summits, it's **easy to mistake a little bit of progress for a lot of progress**
- Because **AI researchers** didn't have reliable theories of how the mind performed intellectual tasks, their work was **largely speculating (or guessing)** what computational method could generate some “smart” behavior
- It's easy to mix up “lack of machine capacity” and “inability to write programs”
- **Simulating aspects of human mental capabilities is understandable by many**
- Replacing expensive people with robots appeals to the military and business

Eliza 1966

Human: Men are all alike.

Eliza: IN WHAT WAY

Human: They're always bugging us about something or other.

Eliza: CAN YOU THINK OF A SPECIFIC EXAMPLE

Human: Well, my boyfriend made me come here.

Eliza: YOUR BOYFRIEND MADE YOU COME HERE

Human: He says I'm depressed much of the time.

Eliza: I AM SORRY TO HEAR YOU ARE DEPRESSED

Human: It's true. I am unhappy.

Eliza: DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY

Human: I need some help, that much seems certain.

Eliza: WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP

Human: Perhaps I could learn to get along with my mother.

Eliza: TELL ME MORE ABOUT YOUR FAMILY

Human: My mother takes care of me.

Parry 1973

Eliza: Good Evening. Tell me your problems. Please terminate input with a period or question mark.

Parry: People get on my nerves sometimes.

Eliza: I am not sure I understand you fully.

Parry: You should pay more attention.

Eliza: Suppose you should pay more attention.

Parry: You're entitled to your own opinion.

Eliza: What makes you think I am entitled to my own opinion?

Parry: It bothers me just to be around people in general.

Eliza: Please go on.

Parry: I went to the track at Bay Meadows a while back.

Eliza: What does that suggest to you?

Parry: I went to the races.

Eliza: What else would you like to discuss?

Parry: Horseracing interests me.

Eliza: Why do you say that just now?

Parry: There's a lot of crooked gambling going on there.

Eliza: I am not sure I understand you fully.

Parry: It's clear to me if not to you.

SHRDLU 1972

1. *Pick up a big red block.*

OK

2. *Grasp the pyramid.*

I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.

3. *Find a block which is taller than the one you are holding and put it into the box.*

BY "IT," I ASSUME YOU MEAN THE BLOCK WHICH IS TALLER THAN THE ONE I AM HOLDING.

4. *What does the box contain?*

THE BLUE PYRAMID AND THE BLUE BLOCK.

5. *What is the pyramid supported by?*

THE BOX.

6. *How many blocks are not in the box?*

FOUR OF THEM.

- Gerry Sussman, Terry Winograd, and Eugene Charniak implemented Microplanner for SHRDLU's reasoning and planning
- Terry Winograd designed and implemented Programmar for SHRDLU's language understanding

Nature of AI, 1970s

- Notice that computer science was not yet quite invented
- As with climbing a mountain with false summits, it's **easy to mistake a little bit of progress for a lot of progress**
- Because **AI researchers** didn't have reliable theories of how the mind performed intellectual tasks, their work was **largely speculating (or guessing)** what computational method could generate some “smart” behavior
- It's easy to mix up “lack of machine capacity” and “inability to write programs”
- **Simulating aspects of human mental capabilities is understandable by many**
- Replacing expensive people with robots appeals to the military and business

rpg Timeline: 1975–1981

- 1975: Moved to SAIL; taught Lisp with John McCarthy
- 1976: Bill Scherlis taught me computer science
- 1977: Passed comprehensive exam (including programming problem)
- 1978
 - **Lighthill's AI Winter fading to Spring**
 - Passed qual exam by writing a bunch of AI Handbook articles for Ed Feigenbaum
 - Joined Cordell Green's Program Synthesis Project (PSI)
 - Terry Winograd becomes adviser
- 1981: Graduates (PhD); System name: Υ ; created a custom programming language
- 1975–1981
 - SAIL's Lisp Wizard
 - Wrote many support macros and functions
 - Created two Lisp programming environments
 - Wrote gait (walking) diagnosis system à la Mycin (but with natural language input)



rpg Timeline: 1981–1994

- 1981
 - McCarthy hired me as a (Senior) Research Associate
 - Worked on the Advice Taker
 - Designed QLisp / QLambda with McCarthy (with advice from Jeff Ullman)
 - Developed the “Gabriel Benchmarks” and wrote a book about them & benchmarking Lisp
- 1982
 - Joins LLNL to direct the NIL project, supervises Rod Brooks (!) and Guy Steele (!!)
 - Initiated Common Lisp / member of the Quinquevirate
- 1984: Founded Lucid
- 1989
 - Started the Cadillac / Energize project in response to an AI Winter
 - One of the CLOS designers, wrote its conceptual specification
 - First Worse Is Better paper
- 1991: Started writing poetry seriously
- 1994: Fired from Lucid

Jerome: “Yes, I know that not everyone is a Dick Gabriel fan, but. . . .

Nature of AI, 1980s

- Expert systems are king, following Ted Shortliffe's MYCIN
- Teknowledge, Intellicorp, Inference, Aion, Ascent,...
- Japanese Fifth Generation Computer Systems

PREMISE: (\$AND (SAME CNTXT GRAM GRAMNEG)
 (SAME CNTXT MORPH ROD)
 (SAME CNTXT AIR ANAEROBIC))

ACTION: (CONCLUDE CNTXT IDENTITY BACTEROIDS TALLY .6)

Nature of AI, 1980s

```
PREMISE: ($AND (SAME LEFT  LEG-FOL-LAT-CURV-PATH SWING)
               (SAME RIGHT LAT-TR-BEND SWING)
               (SAME LEFT  PEL-DEV-LAT-MORE-2 SWING)
               (SAME LEFT  PEL-EL SWING))
```

ACTION:

```
(DO-ALL (CONCLUDE* LEFT  LEG-FOL-LAT-CURV-PATH TALLY (LISTOF SWING))
        (CONCLUDE* RIGHT LAT-TR-BEND TALLY (LISTOF SWING))
        (CONCLUDE* LEFT  PEL-DEV-LAT-MORE-2 TALLY (LISTOF SWING))
        (CONCLUDE* LEFT  PEL-EL TALLY (LISTOF SWING))
        (CONCLUDE* LEFT  CIRCUMDUCTION TALLY (LISTOF SWING)))
```

;;; SWING is a symbolic, compound certainty factor

Nature of AI, 1980s

*During the early 1980's, scientists at Teknowledge, Intellicorp and the other ambitiously named companies in **the fledgling artificial-intelligence industry boasted of a bright future in which computers would match people in their ability to make important business decisions.***

*In the last few years [1987–1988], **such optimism has gradually faded. Bringing the visionary technologies of artificial intelligence to the market has proved far more difficult than had been anticipated.*** Many of the original artificial-intelligence companies—including Teknowledge, Intellicorp, the Carnegie Group and the Inference Corporation—have suffered losses. Several others have gone out of business.

—John Markoff, New York Times, 1988

Nature of AI, 1980s

Our tale begins in 1979, when Chuck Williams co-founds a company he calls Inference, to keep faith with the then trendy technological phenom known as AI. Inference builds *a complex, Rolls-Royce of software development environments for building and maintaining knowledge based systems*, ART (Advanced Reasoning Tool), which garners praise and sales and kudos and all that nice stuff—in *the very narrow market that existed through until about the end of the 1980s which thought little of buying \$100,000 Lisp workstations and the like to explore the bounds of KBS and the like.*

—Gary Flood, *Tech Monitor*, 1997

Nature of AI, 1980s

*The system is intended to be especially suited to **complex, ambitious applications that go beyond what has been done before**; thus it provides facilities for **exploratory programming, complexity management, incremental construction of programs**, and so forth.*

–David Moon, “Symbolics Architecture”

Nature of AI, 1980s

Meeting these needs requires an extraordinary system architecture—just another PC or Unix clone won't do. The intended applications demand a lot of processor power, main and virtual memory size, and disk capacity. The system must provide as much performance as possible without exceeding practical limits on cost, and computing capacity must not be diluted by sharing it among multiple users. These purely hardware aspects are not sufficient, however. *The system must also improve both the speed of software production and the quality of the resulting software* by providing a more complete substrate on which to erect programs than has been customary. *Programmers should not be handed just a language and an operating system and be forced to do everything else themselves.*

—David Moon, “Symbolics Architecture”

AI Winter, 1980s

A major retrenchment is occurring in the artificial intelligence industry, dashing the hopes of many companies that thought they would prosper by providing the technology to make computers “think.”

Some of the setback stems from the failure of artificial intelligence to quickly live up to its promise of making machines that can understand English, recognize objects, or reason like a human expert—to be used for such purposes as diagnosing machinery breakdowns or deciding whether to authorize a loan. Despite this, the technology is making slow but steady progress, and now is being subtly incorporated into more conventional computer programs.

Far more of the artificial intelligence industry’s problems result from poor business decisions by companies that were heavily weighted with technologists rather than business minds. Their main mistake was trying to use special computers for artificial intelligence; the machines were too expensive and did not mesh well with those used by potential customers.

“People believed their own hype . . .,” said S. Jerrold Kaplan. “Everyone was planning on growth that was unsustainable.”

—New York Times, 1988

Collateral Damage, AI Winter, 1980s

... [AI companies'] *main mistake was trying to use special computers for artificial intelligence; the machines were too expensive and did not mesh well with those used by potential customers.*

⋮

Corporate customers did not want to spend \$50,000 to \$100,000 for a special machine used by one person. They wanted artificial intelligence programs to run on their existing computers, such as I.B.M. mainframes and Digital Equipment minicomputers, to be shared by many users. Preferably, *they wanted to develop artificial intelligence programs without requiring their own programmers to learn Lisp.*

—New York Times, 1988

Collateral Damage, AI Winter, 1980s

- From 1985–1986, **good Lisp implementations with decent environments were available on stock hardware**
 - Unix Workstations: Sun, Apollo, IBM RT PC, Tektronix, ...
 - Minicomputers: DEC Vaxen, Prime, ...
 - Mainframes: IBM 370 and similar
- **Performance similar to or better than Lisp Machines**
- Develop on Lisp machines, deliver on stock hardware
- “The dog ate it”; “Justice Alito bravely, manfully, clearly announced to the world that his wife did it”

Collateral Damage, AI Winter, 1980s

- Logic programming (Prolog, Japanese Fifth Generation)
- Game-related searching
- Backtracking
 - Chronological (Planner, MicroPlanner)
 - Non-chronological (Conniver)
 - Truth maintenance
 - Constraint solving
- Self-awareness (OO reflection)
- Self-modification

Collateral Damage, AI Winter, 1980s

- Theorem proving, common-sense reasoning, first-order & higher logics, resolution (Prolog)
- Confidence-based reasoning (expert systems, production systems)
- Rule-based systems (OPS5, Mycin, KEE)
- Blackboard architecture (Linda systems)
- (Some) Parallel programming mechanisms (**futures**, Qlisp)
- Knowledge representation, constraint & various domain-specific languages
- Flavors, CLOS, LOOPS, Common LOOPS, metaobjects, reflection, generic functions, mixins, multiple inheritance, metaclasses, KRL, etc
- Neural nets

Collateral Damage, AI Winter, 1980s

- First-class functions, map/reduce, local functions, lambda-expressions
- Typed data (runtime types)
- Exceptions
 - exception handling / condition handling
 - resumable exceptions
- Catch / Throw; local & nonlocal exits
- Multiple values
- Complex numbers built in
- Ratios built in
- Adjustable arrays
- Mixed dynamic & static binding

Collateral Damage, AI Winter, 1980s

- Powerful macro system
 - macros that operate on tree syntax rather than strings
 - macro expansions calculated using the entire programming language itself, not some wimpy string-substitution or string-processor subset
 - hygienic macros
- Optional type declarations (\rightarrow soft typing, etc)
- Pretty printing
 - every data structure printable
 - print \leftrightarrow read identity (MacLisp's float round-trip solution widely adopted)
- Metaobjects & metaprogramming

Collateral Damage, AI Winter, 1980s

- Foreign functions
- Runtime compilation
- Mixed paradigm computing
- Tail recursion
 - to implement conventional structured-programming control structures (recursion \rightarrow iteration)
 - continuation-passing style & organizing compiler back ends
- Backtrace
- Trace (via advice)
- Breakpoints
 - conditional breakpoints expressed in the high-level language itself
 - resumable breakpoints

Collateral Damage, AI Winter, 1980s—Programming Systems

A **programming system is a platform spanning execution hardware at the lowest level up through the interfaces to the outside world and to the programmer or user.**

- Execution hardware
- Operating system, preferably written in the underlying programming language
- Programming language supporting the expression of reflection & self-modification
- Programming-language-specific runtime system, always running
- **Live**
- Relevant, resident libraries
- Network, sensors, actuators
- Interactive environment (IDE) for programmers & users
- **Programming = building on this substrate through accretion and revision**

Collateral Damage, AI Winter, 1980s—Programming Systems

- Lisp machines: Symbolics, LMI, Xerox (Dorado, Dolphin, Dandelion), Three Rivers (PERQ),...
- InterLisp (PDP-10, Xerox)
- Common Lisp (Lucid, Franz, Harlequin / Lispworks, SLIME (Superior Lisp Interaction Mode for Emacs))
- Any Lisp system
- Any Smalltalk
- Self
- Prolog, maybe

Collateral Damage, AI Winter, 1980s–Programming Ideas

- Exploratory programming
- Rapid prototyping
- Incremental compilation
- Structure editors
- Autocompletion
- Standard editor keybindings

Collateral Damage, AI Winter, 1980s–Programming Ideas

- Timesharing
- Interactive computing
- Interpreted languages
- Live / lively systems
- Aspects (Advice (Pilot) → Flavors → CLOS → Aspects)
- Patterns (CLOS & Smalltalk → patterns)
- Agile (iterative, exploratory, incremental programming & rapid prototyping + programming-system mindset → Agile (plus the players))

Collateral Damage, AI Winter, 1980s–Blaming Lisp

AI companies blaming Lisp killed Programming System research (& practice) for almost 30 years.

- “Symposium on Lisp and Functional Programming”
→ “International Conference on Functional Programming,” deleting Lispy research
- “Lisp and Symbolic Computation: An International Journal”
→ “Higher-Order and Symbolic Computation”
- OOPSLA & other PL conferences throttled down accepting Programming System / Lispy / Smalltalkish papers
- **Ever since, the type theorists haven't stop typing**

rpg Timeline: 1994–2024

- 1994: ParcPlace Systems (**not** Xerox)
- 1996: Fired from ParcPlace
- 1997: Consultant to Bill Joy Sun Microsystems
- 1998: **Master of Fine Arts, Creative Writing (Poetry)**
- 1999: Sun Microsystems
- 2002: **Created Onward!**
- 2007
 - **Created Onward! Essays**
 - Fired from Sun
 - IBM Research
 - Initiated creation of Splash
- 2013: Created InkWell
- 2017: Fired / Retired from IBM

“Ah, yes. Richard Gabriel.
Once again straddling the line
between crazy and
crazy-awesome.”

Nature of AI, 2024—In 1948 Turing Said:

If the untrained infant's mind is to become *an intelligent one*, it must acquire *both discipline and initiative*.... But discipline is certainly not enough in itself to produce intelligence. *That which is required in addition we call initiative. This statement will have to serve as a definition. Our task is to discover the nature of this residue* as it occurs in man, and to try and copy it in machines.

Two possible methods of setting about this present themselves. On the one hand we have fully disciplined machines.... We might try to graft some initiative onto these. This would probably take the form of programming the machine...

The other method is to start with an unorganized machine and to try to bring both discipline and initiative into it at once.... Both methods should, I think, be attempted.

—Alan Turing, "Intelligent Machinery" 1948

Nature of AI, 2024—Machine Learning + Symbolic AI

- AlphaGO = Neural Nets + **Monte-Carlo Rollout**
- ChatGPT = Neural Nets + **Embeddings + Transformer + Multi-headed Attention**

Nature of AI, 2024—Machine Learning

- Genetic Algorithms
- Neuroevolution
- Monte Carlo
- Simulated Annealing
- Neural Nets
 - Convolutional
 - Word2vec
 - Transformer / decoder-only
 - Attention mechanism

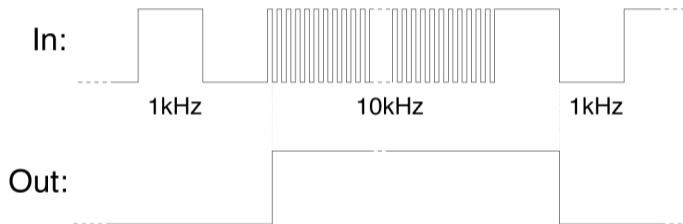
Genetic Algorithms

A **metaheuristic** inspired by the process of natural selection, commonly used to generate high-quality solutions to optimization and search problems by relying on **biologically inspired operators such as mutation, crossover, and selection.**

Genetic algorithms are simple to implement, but **their behavior is difficult to understand.** In particular, it is difficult to understand why these algorithms frequently succeed at generating solutions of high fitness when applied to practical problems.

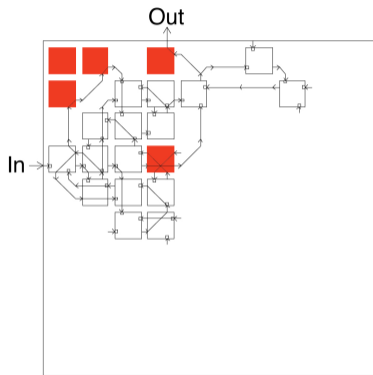
Nature of AI, 2024—Machine Learning

Genetic Algorithms—Weird Shit Can Happen



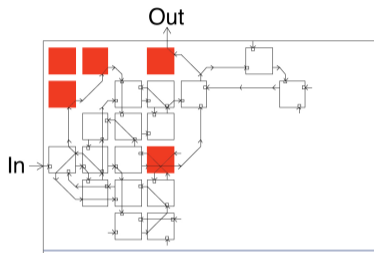
Design a circuit that can distinguish between a 1kHz and 10kHz square wave. The circuit, given a series of bursts of 1kHz and 10kHz signals intermixed, should return a 0 whenever the 1kHz signal was present and a 1 when the 10kHz signal was. The raw material is a 10x10 section of a field programmable gate array without a clock.

Nature of AI, 2024—Machine Learning



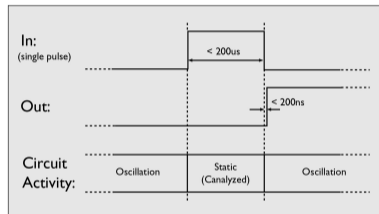
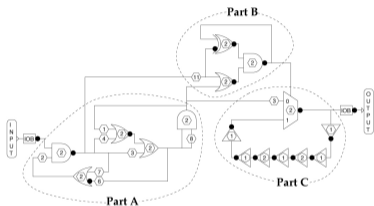
...probably the *most bizarre, mysterious, and unconventional unconstrained evolved circuit* yet reported.—Adrian Thompson, “Analysis of Unconventional Evolved Electronics”

Nature of AI, 2024—Machine Learning



- Degraded performance in a different, identical FPGA
- CMOS discrete circuit **does nothing**
- Logic circuit simulation **does nothing**
- **Does not work** when the **red** (unused) cells are clamped (outputs fixed at 0)

Nature of AI, 2024—Machine Learning



- When the feedback loops in Parts A & B are cut ...
 - parts A & B revert to digital logic and become static
 - part C is observed to be static
- Notice Part C is very weird

Nature of AI, 2024—Machine Learning

- Power levels are quiescent
- No activity observed in the feedback loops
- No short duration pulses (glitches)

even though

- Power supply and wires are shielded / no decoupling capacitors
- Entire circuit in a Faraday cage

Nature of AI, 2024—Machine Learning

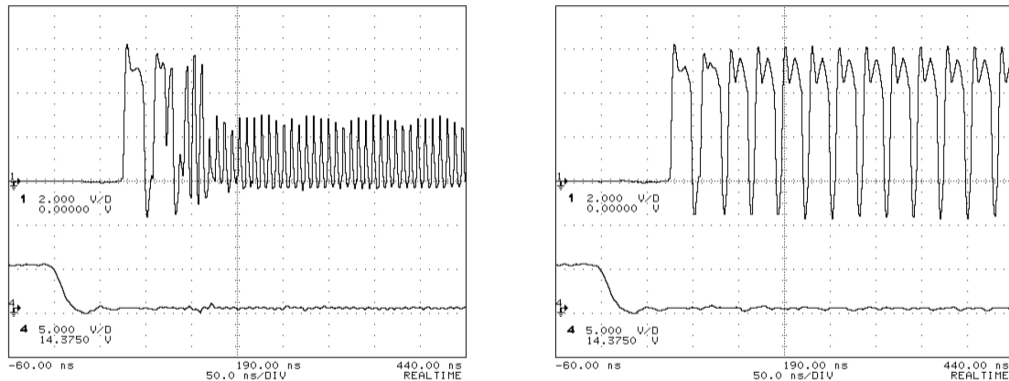
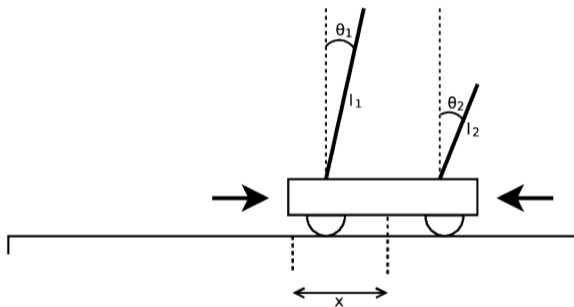


Figure 4. Behavior of the first frequency-discriminating ancestor. The upper waveform is the output, and the lower the input; we see the behavior immediately after the falling edge of a single input pulse. Left: long pulse, Right: short pulse. The input to the FPGA is actually delayed ~ 40 ns by an intervening buffer relative to the wave seen here.

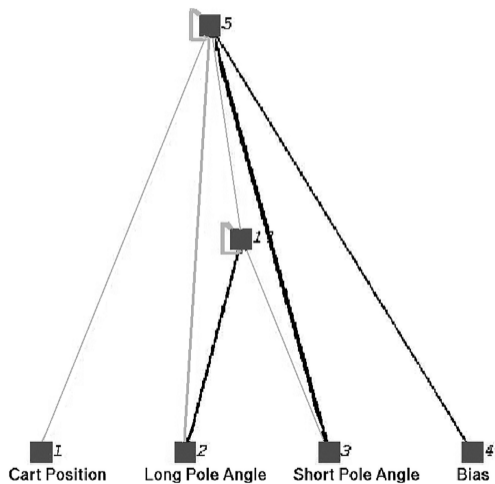
Neuroevolution—Genetic Algorithms to figure out:

- Topology of a neural net
- Weights for connections
- Rules governing how weights change over time

Nature of AI, 2024—Machine Learning



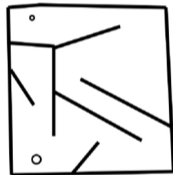
Nature of AI, 2024—Machine Learning



Nature of AI, 2024—Machine Learning

Neuroevolution—Novelty Seeking

- Open-ended evolution
 - increase diversity
 - increase complexity
- Favor novelty over fitness
- Use novelty only



*avant-garde Cooper Union for the Advancement of Science and Art
setting on fire
the half-deep-fat-fried ship-towed long-range acoustic detection system*

Nature of AI, 2024—Machine Learning

What a surprise.

Computers and machines can do some things better than people can.

Nature of AI, 2024—Machine Learning

- Word2vec
- Transformer / decoder-only
- Attention mechanism

Nature of AI, 2024—Machine Learning

Word2vec definition of “woman”

-0.51821, -0.13809, -0.41185, -0.13133, 0.0035659, -0.31205, -0.31242, -0.43538, -0.27017, -1.1338, 0.19282, -0.23786, 0.0028892, -0.027078, 0.14747, -0.051265, 0.089021, -0.12337, -0.40892, -0.39197, -0.5665, 0.40684, -0.057313, 0.18613, -0.33095, -0.25243, 0.33452, -0.23104, -0.0099149, 0.24269, -0.57504, 0.30122, -0.33779, -0.294, -0.80202, 0.44674, -0.40296, -0.21026, 0.1072, 0.53937, 0.63971, -0.3154, -0.082537, -0.038314, -0.058821, 0.11235, 0.50928, 0.14656, -0.18988, 0.24132, -0.038442, 0.071418, 0.38471, 0.10584, -0.52786, -0.057574, 0.13771, -0.48613, 0.62553, -0.3308, -0.065539, 0.1791, 0.62746, 0.087208, -0.60422, -0.64595, -0.2389, -0.1585, -0.07948, -0.037848, -0.205, -0.44504, -0.21127, -0.0026664, -0.0080152, 0.39607, 0.069091, -0.34379, -0.13992, 0.084029, -0.40245, 0.094426, 0.28908, 0.6216, -0.4132, -0.12153, -0.40676, -0.05771, 0.17415, 0.094069, -0.65955, -0.34154, -0.079291, 0.16867, 0.31665, -0.072868, -0.44181, -0.19555, 0.16942, -0.197, 0.17484, 0.11972, 0.012957, -0.32238, 0.33737, 0.35553, 0.82906, -0.37521, 0.34777, -0.13105, -0.046688, 1.2125, 0.025435, -0.36021, 0.65259, 0.63729, 0.23457, 0.19874, 0.22589, -0.016769, 0.10671, 0.86805, -0.02853, -0.011634, -0.39911, 0.12429, -0.066363, 0.0080237, 0.41826, -0.60603, 0.26269, 0.035625, -0.10225, 0.44362, -0.38952, -0.054241, -0.31542, -0.076452, -0.33072, 0.42178, 0.29242, 0.13222, -0.14896, 0.32047, -0.4735, -0.1093, 0.31163, 0.49836, -0.20143, 0.36058, 0.32242, -0.11148, 0.6416, 0.20955, 0.035183, 0.16684, -0.41427, -0.41894, 0.18777, 0.39424, 0.25784, 0.18818, 0.60567, 0.15537, -0.27121, 0.054047, -0.18342, 0.29789, 0.35805, -0.40148, -0.019914, -0.019742, -0.56609, -0.25878, -0.036075, 0.0093725, -0.25284, -0.061715, -0.26441, 0.47597, 0.087956, 0.051997, 0.21366, -0.0034455, 0.1739, -0.16853, -0.22233, -0.1006, -0.032696, -0.008549, 0.036532, -0.19339, -0.28571, -0.29294, -0.53655, 0.16387, -0.36861, -0.52443, -0.84287, 0.26247, 1.8261, 0.029467, 0.19155, 0.28406, -0.1017, -0.31416, -0.084328, 0.42934, 0.32851, 0.41274, -0.080323, 0.063666, -0.18441, 0.13328, 0.46, -0.24984, 0.12574, -0.49056, -0.072603, 0.28191, -0.25738, 0.40629, 0.38381, -0.37685, -0.16371, 0.30354, -0.38234, -0.61633, -0.22076, 0.38153, 0.54091, -0.32349, -0.032075, -0.051326, -0.12465, 0.19237, -0.077144, 0.27005, -0.20103, -0.26512, 0.35769, -0.23437, 0.054273, 0.16901, -0.15758, 0.42714, 0.23167, -0.021318, 0.3086, 0.44873, -0.12432, -0.15715, -0.099448, -0.24825, 1.156, -0.38925, -0.0063171, 0.48928, 0.46089, -0.17058, 0.06118, -0.54, -0.054482, 0.13329, -0.47944, 0.17119, 0.26289, 0.14383, -0.30443, 0.27534, -0.14711, -0.52172, 0.32909, -0.15149, 0.1539, 0.24171, -1.6971, 0.027579, -0.0073776, 0.30144, 0.011751, 0.012419, 0.38711, -0.044167, -0.62495

Nature of AI, 2024—Machine Learning

word2vec algebra

$$\text{word2vec}(\text{"father"}) - (\text{word2vec}(\text{"man"}) - \text{word2vec}(\text{"woman"}))$$

mother	0.0821
daughter	0.1030
father	0.1179
husband	0.1335
wife	0.1339
grandmother	0.1495
her	0.1583
woman	0.1674
daughters	0.1737
sister	0.1804



Nature of AI, 2024—Machine Learning

word2vec algebra

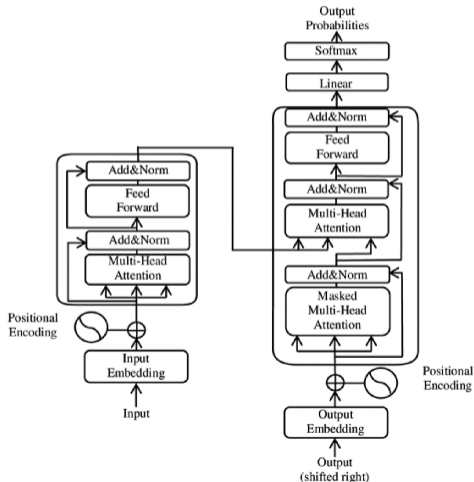
$$\text{word2vec}(\text{"bull"}) - (\text{word2vec}(\text{"man"}) - \text{word2vec}(\text{"woman"}))$$

bull	0.1045
female	0.2949
woman	0.3046
pregnant	0.3197
male	0.3201
calves	0.3212
cow	0.3213
horse	0.3217
klien	0.3227
baby	0.3230



Nature of AI, 2024—Machine Learning

Transformer with Attention



← not machine learned

Nature of AI, 2024—Machine Learning

```
(defun nfib (n)
  (labels ((double (fn+1 fn fn-1)
              (let ((f2n+1 (+ (* fn+1 fn+1) (* fn fn)))
                    (f2n   (* fn (+ fn+1 fn-1))))
                (values f2n+1 f2n (- f2n+1 f2n))))
    (next (fn fn-1 fn-2) (declare (ignore fn-2))
      (values (+ fn fn-1) fn fn-1))
    (dfib (i)
      (cond ((= i 1) (values 1 0 1))
            ((oddp i) (multiple-value-call #'double (dfib (ceiling i 2))))
            (t (multiple-value-call #'next (dfib (- i 1)))))))
  (cond ((zerop n) 0)
        (t (values (dfib n))))))
```

This is an efficient method for calculating Fibonacci numbers as it uses squaring techniques to jump multiple indices in the sequence, reducing the total number of recursive calls needed. However, the function might have some issues returning values directly as it employs multiple-value-bind implicitly which needs handling at the top-level call. If used in a Lisp environment that supports multiple return values properly, it should work as intended.

Nature of AI, 2024—Machine Learning

```
(defun f (n)
  (labels ((t4701 (t4695 t4691 t4694)
            (let ((t4696 (+ (* t4695 t4695) (* t4691 t4691)))
                  (t4700 (* t4691 (+ t4695 t4694))))
              (values t4696 t4700 (- t4696 t4700))))
    (t4702 (t4691 t4694 t4698) (declare (ignore t4698))
      (values (+ t4691 t4694) t4691 t4694))
    (t4699 (i)
      (cond ((= i 1) (values 1 0 1))
            ((oddp i) (multiple-value-call #'t4701 (t4699 (ceiling i 2))))
            (t (multiple-value-call #'t4702 (t4699 (- i 1)))))))
  (cond ((zerop n) 0) (t (values (t4699 n))))))
```

This function appears to be a complex mathematical operation with recursive elements and multiple returns. It might be related to a specific numerical sequence or algorithm in number theory or cryptography, based on the operations performed. However, without further context or specific domain knowledge, the precise purpose or the sequence it's intended to compute or simulate is not immediately clear.

Nature of AI, 2024—Machine Learning

Please write a one-stanza poem about loud guitar blues music.

*Strings scream in moonlight's sway,
Echoes of soul in night's embrace,
A bluesy roar, wild and untamed,
Heart's raw cry in rhythm's **race***

*Strings scream in moonlight's sway,
Echoes of soul in night's embrace,
A bluesy roar, wild and untamed,
Heart's raw cry in rhythm's **pulse**.*

Revise it to not rhyme, please. (↑)

(↓) Revise it to make it more interesting, please.

*Strings wail under the moon's watch,
Echoes of soul ripple through the dark,
A bluesy roar, fierce and untamed,
Heart's raw cry pulses through the night.*

*tuned adrenaline
my music
a beat-boogied headful
(InkWell)*

Nature of AI, 2024—Machine Learning

Tell me about Nickieben Bourbaki.

Nickieben Bourbaki is friends with Richard P. Gabriel.

Would you like to know more about their interests, background, or anything specific?

Please write one paragraph about them.

*Nickieben Bourbaki, a creative and enigmatic figure, shares a deep and enduring friendship with Richard P. Gabriel. **Their bond is rooted in a mutual appreciation for the arts, philosophy, and intellectual pursuits. Nickieben is known for their eclectic taste and vibrant personality, often engaging in spirited discussions and collaborations that span various fields of interest.** With a penchant for exploring unconventional ideas and pushing boundaries, Nickieben and Richard's friendship thrives on a shared curiosity and a passion for innovation and creativity.*

Nature of AI, 2024—Machine Learning

Revise this (↓) in your voice and make it more interesting.

run down church in a city setting
paint peeling off outside beams
and doors / concrete stairs
with corners knocked off
crosses covered in dust
preacher wearing worn clothes
reading from the oldest
of revered books telling
stories one is allowed
to remember / God is funny

In the heart of a bustling city, a rundown church stands as a quiet testament to forgotten times. The paint peels off the weathered beams and doors, while concrete stairs, their corners chipped and worn, lead to a sanctuary of neglect. **Dust-covered crosses hang in silent witness to days long past.** Inside, the preacher, clad in threadbare clothes, reads from a timeworn, revered book, weaving tales that linger in the memory of those who listen. In the midst of it all, there's a subtle humor, a reminder that even in decay, God has a way of making us smile.

Nature of AI, 2024—Machine Learning

It takes a special engineer to automate the beauty out of writing.

—Michael Richmond

... everything is a program (even if it's not), and once something is a program we can give it a precise semantics and bring to bear a host of techniques to reason about its behavior.

—Todd Millstein, October 23, 2024

Nature of AI, 2024—Machine Learning

- LLMs are not writers
- They know only what they've read + what the attention algorithm tells them
- An LLM is not a distinct individual—it cannot have a point of view
- No LLM has had an experience
- LLMs like to explain (“ChatSplain”?)
- Turing Test (*The Imitation Game*) tells us:
 - **not** that LLMs are nearly human
 - **but** that people are not as special as we hoped

Nature of AI, 2024—Machine Learning

- LLMs can't revise (in the human sense)
- They write OK first drafts (at a good high-school student level): term papers
- What writers care about:
 - relish every word
 - be simple but go deep
 - take risks
 - seek beauty
 - find the right pitch
 - music (the noise of language)
 - voice
 - lyricism
 - melody (the noise of words in sentences and paragraphs)
 - rhythm
 - connotation
 - culture
 - strangeness

Nature of AI, 2024—A Surprising Guess

- Designed to know everything, it cannot discover
- Designed to be everyone, it cannot be anyone
- Designed to know only through text, it cannot know through experience
- To be more human, **it has to embrace ignorance and illiteracy**

*... **all good novelists have bad memories.** What you remember comes out as journalism. What you forget goes into the compost of imagination... In a compost heap, things decompose. Your past is full of stories that have been composed in a certain way; that's what memories are. But **only when they decompose are you able to recompose them into new works of art.***

—Graham Greene, paraphrase

Nature of AI, 2024—A Surprising Guess

*I have grown to understand narrative as a form of contemplation, a complex and seemingly incongruous way of thinking. I come to know my stories by writing my way into them. I focus on the characters without trying to attach significance to their actions. I do not look for symbols. For as long as I can, **I remain purposefully blind to the machinery of the story and only partially cognizant of the world my story creates. I work from a kind of half-knowledge.***

—Robert Boswell, “The Half-Known World”

Nature of AI, 2024—A Surprising Guess

In the drafts that follow, I listen to what has made it to the page. Invariably, things have arrived that I did not invite, and they are often the most interesting things in the story. By refusing to fully know the world, I hope to discover unusual formations in the landscape, and strange desires in the characters. By declining to analyze the story, I hope to keep it open to surprise. Each new draft revises the world but does not explain or define it. I work through many drafts, progressively abandoning the familiar. **What I can see is always dwarfed by what I cannot know.** What the characters come to understand never surpasses that which they cannot grasp. **The world remains half-known.**

—Robert Boswell, “The Half-Known World”

Nature of AI, 2024—A Surprising Guess

*There can be no discovery in a world where everything is known.
A crucial part of the writing endeavor is to practice remaining in the dark.*

—Robert Boswell, “The Half-Known World”

Nature of AI, 2024—“Bad Moon Rising”

Why have AI-linked stocks suffered?

Primarily, *there is concern about whether the vast investment in AI* by Microsoft, Google, and others *will pay off*. This has been bubbling away in recent months. Analysts at Goldman Sachs published a note in June with the title *“Gen AI: too much spend, too little benefit?”* The Wall Street bank asked if a \$1tn investment in AI over the next few years will “ever pay off,” ...

—The Guardian, **August 3, 2024**

Symbolic-AI Winter, 2024

... the focus on deep learning and LLMs has been draining resources for everything else. Funding has been strongly affected: Many, but not all, funders only want the new & shiny. Some see their mandate as maintaining a broad portfolio, or argue that industry is pouring rather enough into DL/LLMs as it is, but they are a minority—alas. It also affects recruitment of graduate students and faculty recruitment—people outside AI want to hire only generative AI people now. *The number of graduate students who know anything about symbolic AI is dropping rapidly....*

—anonymous, October 8, 2024

Allen Newell Award Statement

June 11, 2005, Palace Hotel, San Francisco, California

To be alive at the start of a new field—what a treasure. Some say scientific progress looks like a Christmas tree: a wide diversity of ideas at the start, narrowing as some of the ideas are refined, then another, smaller diversity given what's working and not, followed by a deeper narrowing, and so on toward the ultimate star of truth or of understanding—in any case, toward the top.

We are coming to realize that computing is not like any single other thing people do—it's a little like science, mathematics, engineering, a little like art, architecture, writing, horticulture, and a lot like a long, unplanned journey.

The star at the top? Is that our ambition, or are we subject only to the laws of the imagination?

Research requires a kick, at times, to get the soul dancing.

&

