The Third AI Summer
AAAI Robert S. Engelmore Memorial Lecture
Director’s Cut

Henry Kautz
Director, Information & Intelligent Systems, National Science Foundation
Professor, University of Rochester
Greatest Hits?
Director’s Cut

• Numerous corrections!
• Hidden slides revealed for the first time!
• Three entirely new slides!
Big Picture?

History?

Dangers?

Future?
Three in One

Part I: History of AI
Part II: AI for Bad
Part III: Future of AI
Part I: History of AI
The Cartoon History of AI

Symbolic AI

Neural Networks
The Cartoon History of AI

Symbolic AI

Neural Networks

Deep Learning
AI Summers and Winters

- Enthusiasm, funding, and approaches to AI have been cyclical
- Enduring insights have blossomed in each summer
Overview

• Overview of a few of the enduring insights from each AI summer
• First Summer: Irrational Exuberance (1948 – 1966)
• First Winter (1967 – 1977)
• Second Summer: Knowledge is Power (1978 – 1987)
• Second Winter (1988 – 2011)
• Third Summer (2012 – ?)
• Why there might not be a third winter!
First Summer: Irrational Exuberance (1948 – 1966)

William Walter Grey’s Tortoises
Insight 1: The Tortoise’s Legacy

1A: Artificial Neural Networks
- McCulloch & Pitts neuron (1943)
- Perceptron: Single-layer learning (Rosenblatt 1958)
- Backpropagation: Multi-layer learning (Werbos 1974; Rummelhart, Hinton, Williams 1986)
- Parameter sharing: Convolutional networks (Fukushima 1980; LeCun 1989)

1B: Behavior = Agent + Environment
- Self-organizing machines (Turing 1948)
- Dynamic programming (Bellman 1957)
- Temporal difference learning (Witten 1977; Sutton & Barto 1981)
- Self-play (Samuels 1959; Tesauro 1992)
Insight 2: Declarative Knowledge Representation

• Logic Theorist: Mathematical theorem proving (Newell, Shaw, & Simon 1956)
• Advice Taker: Commonsense knowledge (McCarthy 1958)
• Semantic networks: Graph representations (Richens 1958)
• KL-ONE: Object-oriented logic (Brachman 1977)
• Bayesian networks: Probabilistic graph representations (Pearl 1985)
• Probabilistic relational models (Friedman et al. 1999)
• Knowledge graphs (Google 2012)
Insight 3: Efficient Combinatorial Search

• Game-tree search (Samuels 1952)
• Model finding (Davis, Putnam, Logemann, & Loveland 1962)
• A* (Hart, Nilsson, & Raphael 1968)
• Means-ends analysis (Newell, Shaw, & Simon 1960; Nilsson et al. 1969)
• Hierarchical planning (Tate 1975)
• Conflict-driven learning (Marquis-Silva & Sakallah 1996)
• Planning as satisfiability (Kautz & Selman 1992)
• Local search for SAT (Selman & Kautz 1993)
• Optimization with parity constraints (Ermon et al. 2013)
First Winter (1967 – 1977)

• DARPA funding cuts
  • Early failure of practical speech recognition and autonomous tank project
  • Mansfield Amendment (1969) mandated that DARPA stop funding basic undirected research

• United Kingdom: Lighthill Report (1973)
  • Criticized failure of AI outside of toy micro-worlds
  • Claimed AI could never tame combinatorial explosion of real-world domains
  • Complete dismantling of AI research in UK
Second Summer: Knowledge is Power (19[67]8 – 1987)

MYCIN (Shortliffe & Buchannan 1976)
RULE035

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

ACTION:  (CONCLUDE CNTXT IDENTITY BACTEROIDES TALLY .6)

IF:   1) The gram stain of the organism is gramneg, and
      2) The morphology of the organism is rod, and
      3) The aerobicity of the organism is anaerobic

THEN:  There is suggestive evidence (.6) that the identity
       of the organism is bacteroides

MYCIN (Shortliffe & Buchannan 1976)
Insight: Modeling Expert Knowledge

Knowledge Engineering
• Dendral (Feigenbaum et al. 1968)
• MYCIN (Shortliffe & Buchannan 1975)
• XCON (Dermott 1978)

Knowledge Induction
• Decision tree learning (Quinlan 1986)
• Inductive logic programming (Muggleton 1991)
• Decision-theoretic expert systems (Heckerman, Horvitz, Nathwani 1992)

• Knowledge-engineered expert systems proved costly to maintain
• Collapse of market for specialized AI workstations
• Failure of Japan’s Fifth-Generation AI effort based on Prolog hardware and software
Third Summer: Deep Learning (2012 – ?)

China’s Sputnik Moment
Third Summer: Deep Learning (2012 – ?)
Insight 1: Hierarchical Representation Learning

• While knowledge representation was all about capturing hierarchies...

• Traditional machine learning considered only two levels of representation: feature and class

• Deep learning’s innovation was to learn hierarchical representations

Image by Honglak Lee et al. (2011)
Insight 2: Similarity

- Deep learning representations directly support **concept similarity**
- Reasoning about similarity is **vital** for most real-world domains
- Not captured by probability!

![Image by David Rozado (2019)]
Why Winter Might Not Return

• The World’s Unreasonable Predictability

• *Truly unreasonable* – who would think that “statistical monkeys at typewriters” could produce paragraphs of coherent prose?
  
  • Prompt: In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.
  
  • GPT-2: ... While examining these bizarre creatures the scientists discovered that the creatures also spoke some fairly regular English. Pérez stated, “We can see, for example, that they have a common ‘language,’ something like a dialect or dialectic.” Dr. Pérez believes that the unicorns may have originated in Argentina, where the animals were believed to be descendants of a lost race of people who lived there before the arrival of humans in those parts of South America.
Non-Human Intelligence

• Deep learning can create a simulacra of human intelligence, *without being anything like* human intelligence

• Compare with *octopus intelligence*
  • Could a creature with a *lifespan of 3 years*, which *cannibalizes* the male after mating, and spawns *70,000* offspring ever really understand *you*?

• But an octopus’ brain is more like a human brain than is any deep learning systems!
It Just Works

• Despite being utterly unlike human intelligence, today’s AI can drive powerful applications for good or for evil
• AI supremacy == world supremacy?
• Growing number of workshops and conference tracks on AI for Good
• So let’s turn to ...
Part II: AI for Bad
Overview

• Dangerous applications are enabled by third summer AI
• Threats most prominent in the news might not be the worst!
• Three examples:
  • Face recognition
  • Fake news
  • Autonomous weapons
In the News: Keeping Your Face Private

Facebook to Pay $550M to Settle Class Action Case Over Facial Recognition

The settlement in a case over the social network’s Tag Suggestions feature is the latest financial blow the company has taken over its handling of user privacy.

San Francisco Bans Facial Recognition Technology

Attendees interacting with a facial recognition demonstration at this year’s CES in Las Vegas. Joe Buglewicz for The New York Times

By Kate Conger, Richard Fausset and Serge F. Kovaleski

May 14, 2019

SAN FRANCISCO — San Francisco, long at the heart of the technology revolution, took a stand against potential abuse on Tuesday by banning the use of facial recognition software by the police and other agencies.
In the News: Keeping Your Face Private

• Face Recognition == Threat

• But your mobile phone apps are giving your location away – no AI needed!

• A. Sadilek, H. Kautz & J. P. Bigham (2012), Finding Your Friends and Following Them to Where You Are

• L. Liao, D. J. Patterson, D. Fox, and H. Kautz (2007), Learning and Inferring Transportation Routines

Greater Threat: Keeping Your Mind Private

Physical Activity  Online Activity
Predictive Model
Mood  Beliefs  Intentions

Can be used for GOOD
Detecting Low Self-Esteem in Youths from Web Search Data (Zaman, Acharyya, Kautz, & Silenzio 2019)
Greater Threat: Keeping Your Mind Private

But this is about BAD...
Greater Threat: Keeping Your Mind Private

- Totalitarian states are inefficient when they target entire demographic groups for repression
  - Cost
  - Radicalization
  - International pressure

- AI enables precise micro-targeting of wrong-thinkers

Xinjiang Re-Education Camp
In the News: Fake News

Artificial intelligence (AI)

New AI fake text generator may be too dangerous to release, say creators

The Elon Musk-backed nonprofit company OpenAI declines to release research publicly for fear of misuse

▲ The AI wrote a new passage of fiction set in China after being fed the opening line of Nineteen Eighty-Four by George Orwell (pictured). Photograph: Mondadori/Getty Images
But Fake News is Old News
Greater Threat: Fake Friends

1 in 5 millennials are lonely and have ‘no friends’: survey
By Hannah Frischberg
August 2, 2019 | 5:19pm

Suicide is Gen Z's second-leading cause of death, and it's a worse epidemic than anything millennials faced at that age

- The suicide rate for people aged 10 to 24 increased by 56% between 2007 to 2017, according to new data from the CDC.

- For children aged 10 to 14, the suicide rate tripled between 2007 to 2017 after years of decline.

- Suicide had increased among millennials, but the data suggests Gen Z might be most at risk for mental illness.
Greater Threat: Fake Friends
In the News: Autonomous Weapons

Defense Innovation Board AI Ethical Guidelines

• Human beings should exercise appropriate levels of judgment and remain responsible for the development, deployment, use and outcomes of DOD AI systems.

• Does this mean human-in the loop?

Modular Advanced Armed Robotic System
Quentic North America
Ethical Military Use of AI != Human in the Loop

Larry Lewis

- Senior Advisor for the State Department on Civilian Protection, Obama administration
- Member US Delegation for UN Deliberations on Lethal Autonomous Weapons Systems
- Artificial intelligence may make weapons systems and the future of war relatively less risky for civilians than it is today

Killer robots reconsidered: Could AI weapons actually cut collateral damage?

By Larry Lewis, January 10, 2020

The Bulletin of the Atomic Scientists, 10 Jan 2020
Greater Threat: Autonomous Weapons of Environmental War: Deep Sea Mining

History’s Largest Mining Operation About to Begin

It’s underwater—and the consequences are unimaginable.

AI is also a big part of the defense of oceans!

The Mariana Trench

Woods Hole “Sentry” UAV

Story by Wil S. Hylton

JANUARY/FEBRUARY 2020 ISSUE

SCIENCE
Part III: Future of AI
The Marcus-LeCun Twitter War

• *Editor’s note: artistic license has been applied to the quotations*

• Marcus: AI systems need to perform symbolic reasoning!
• LeCun: Yes, been working on that.
Overview

• Violent agreement on the need to bring together the neural and symbolic traditions

• How do neuro and symbol systems interact?
  • symbolic Neuro symbolic
  • Symbolic[Neuro]
  • Neuro ; Symbolic
  • Neuro: Symbolic $\rightarrow$ Neuro
  • Neuro_{Symbolic}
  • My favorite: Neuro[Symbolic]
symbolic **Neuro** symbolic

• Deep learning SOP (Standard Operating Procedure)

![Deep learning SOP diagram](https://blog.aylien.com/)

Image from [https://blog.aylien.com/](https://blog.aylien.com/)
Symbolic[Neuro]

- Neural pattern recognition subroutine within a symbolic problem solver
- AlphaGo (2016)
- AlphaZero (2017)
Symbolic[Neuro]

Image from https://deepdrive.berkeley.edu/
Neuro ; Symbolic

- Cascade from neural network into symbolic reasoner
- Neuro-Symbolic Concept-Learner (Mao, Gan, Kohli, Tenenbaum, Wu 2019)
Training compiles away symbolic rules
- Rule A \( \rightarrow \) B becomes an input-output training pair (A, B)
- Deep Learning for Symbolic Mathematics (Lample & Charton 2020)
- But no guarantee of correctness

input: \( \int x^n \, dx \)  
output: \( (1/n+1) \, x^{n+1} \)
• Symbolic rules are used as templates for structures within the neural network
  • Tensor Product Representations (Smolensky et al. 2016)
  • Logic Tensor Networks (Serafina & Garcez 2016)
• Demonstrated for encoding abstraction and part-of hierarchies
• Open: What about true combinatorial reasoning by cases?
My Favorite: Neuro[Symbolic]

- Imbed true symbolic reasoning inside a neural engine
- Enable super-human and super-neuro combinatorial reasoning
  - For deliberative, Type 2 reasoning
  - Rare in ordinary animal life
  - Common in “business AI”
- Interface at the Attention Schema
  - Internal model of the system’s state of attention
  - Not the same as attention itself!
Neuro-Symbolic Interface

• Proposal: When attention to concepts is very high, they are decoded to symbolic entities in an attention schema

• Appearance of a goal in the attention schema signals that deliberative symbolic reasoning should be initiated
  • I’m not just perceiving or remembering, I’m consciously thinking!

• Efficient combinational search can then be performed over the entities in the attention schema
Symbolic Reasoning Engine

Go_To(Cheese)

Decoder

Encoder

Neural Network

Forward(1)
Turn left
Forward (3)
...

52
Discussion

• Why not maintain a (possibly reduced) vector representation in the attention schema?
  • Recurrent Independent Mechanisms: Bengio, Schölkopf et al.

• In order to leverage efficient constraint-satisfaction algorithms
  • 200 terabyte proof of the Boolean Pythagorean Triples Problem (Huele et al. 2016)

Marijn Heule (2016)
Discussion

• The proposed symbolic solver is not differentiable, so it won’t support gradient descent-based learning
• You wouldn’t want to back propagate through a trillion inference steps anyway.
• The neural network can learn from inputs/outputs of SRS using non-gradient based optimization
• This scheme does not handle non-logical reasoning
• Could be extended to probabilistic reasoning, but probably not similarity-based reasoning
• Goal is expert reasoning, not commonsense reasoning: a step toward superintelligence, not human intelligence
Summary

• Each AI summer has led to enduring scientific insights
• Today’s third summer is different
  • It might not be followed by a winter
  • It enables powerful applications for good and bad
• The next steps in AI are tighter symbolic-neuro integration
Thank You

If I have seen further it is by standing on the sholders of Giants.
- Isaac Newton, February 5, 1675, in a letter to Robert Hook.
  Hooke never replied.