# BROAD AI

### now and later

Michael Witbrock, PhD University of Auckland Broad Al Lab @witbrock Aristotle (384–322 BCE) Organon

## **ROOTS OF AI**

Nullus ho

ub

tern ?

53

albus eft.

albus .

non ca

eff homo alba

hocauténett

peccllario Man

uniusaffirmati.

tet negate nega

mauit, & ab co

at in Signa a

Relating un vie

crates albras, no

Deregulis oppositarum propositioni. Cap,6.

Subcontratiz.

Constatis

discours.

Conter.

contra.

OTs homo

albus eft ?

Subalteritz

Aligs ha albus eft.

物性能引发

**ROOTS OF AI** 

H

Santiago Ramón y Cajal (1852 -1934) Cerebral Cortex

E

### WHAT'S AI

- OLD definition: Al is everything we don't yet know how program
- Now some things that people can't do:
  - unique capabilities (e.g. Style transfer)
  - superhuman performance (some areas of speech, vision, games, some QA, etc)
- Current AI Systems can be divided by their kind of capability:
  - Skilled (Image recognition, Game Playing (Chess, Atari, Go, DoTA), Driving)
  - Attentive (Trading: Aidyia; Senior Care: CareMedia, Driving)
  - Knowledgeable, (Google Now, Siri, Watson, Cortana)
  - High IQ (Cyc, Soar, Wolfram Alpha)

### GOFAI

- Thought is symbol manipulation
- Large numbers of precisely defined symbols (terms)
- Based on mathematical logic

(implies

(and

(isa ?INST1 LegalAgreement) (agreeingAgents ?INST1 ?INST2)) (isa ?INST2 LegalAgent))

 Problems solved by searching for transformations of symbolic representations that lead to a solution





# Slow Development

4 <sup>th</sup> Century BCE	Early Development of Logic (Aristotle)
19 <sup>th</sup> Century	Complex mechanical computers (Babbage); Early understanding of neural basis of thought (Ramón y Cajal)
1949	Hebbian (correlational) Learning (Hebb)
1950	Imitation Game (Turing)
1959	Search and ML for Checkers/Draughts (Samuel)
1966	ELIZA – simple chatbot (Weizenbaum)
1971	SHRDLU – Graphical Chatbot with world model (Winograd)
1972	SHAKEY – Autonomous Robot with planning (Rosen et al.)
1984	Cyc – Large-scale logic-based KB and reasoner
1986	Backprop – basis for modern NN's (Rumelhart <i>et al</i> ) NAVLAB 1 – self driving vehicle (Thorpe <i>et al</i> )
1997	Deep Blue – defeats world chess champion (Hsu, Campbell et al.)

### Thinking Quickly (System I)

# Thinking Slowly (System II)

Done well by animals and people Massively parallel algorithms Done poorly until now by computers Not impressive to ordinary people Human Superpower c.f. other animals Serial and slow Done poorly by most people Impressive (prizes, high pay)

"Sir, an animal's reasoning is like a dog's walking on his hind legs. It is not done well; but you are surprised to find it done at all." - apologies to Samuel

Achieved on computers by highpower, low density, slow simulation of vastly different neural hardware Fundamental design principle of computers Computer superpower c.f. human

Recurrent Deep Learning

Deep Reasoning

### MACHINE LEARNING

- Meaning is implicit in the data
- Thought is the transformation of learned representations

#### PANDARUS:

Alas, I think he shall be come approached and the day When little srain would be attain'd into being never fed, And who is but a chain and subjects of his death, I should not sleep.

#### Second Senator:

They are away this miseries, produced upon my soul, Breaking and strongly should be buried, when I perish The earth and thoughts of many states.



http://karpathy.github.io/2015/05/21/rnneffectiveness/

## MACHINE LEARNING: GPT-2

- Meaning is implicit in the data
- Thought is the transformation of learned representations

Troilus and Cressida, Act 4 Scene 2. Maybe.

SCENE II. The same. Court of Pandarus' house. Enter TROILUS and CRESSIDA CRESSIDA Come, come, beshrew your heart! you'll ne'er be good, Nor suffer others. PANDARUS Ah, what is thy humour? CRESSIDA All you that see me, what does your wife say? PANDARUS Does she say what? CRESSIDA Your wife says that the worst rascal of all Has not yet come to dine. PANDARUS To dine! 'tis the biggest villain Who can see another to dinner. CRESSIDA What does your wife say?

SCENE II. The same. Court of Pandarus' house. Enter TROILUS and CRESSIDA CRESSIDA Come, come, beshrew your heart! you'll ne'er be good, Nor suffer others. PANDARUS Ha! ha! Alas, poor wretch! ah, poor capocchia! hast not slept to-night? would he not, a naughty man, let it sleep? a bugbear take him! CRESSIDA Did not I tell you? Would he were knock'd i' the head!

### Knocking within

Who's that at door? good uncle, go and see. My lord, come you again into my chamber: You smile and mock me, as if I meant naughtily

## **REASONING AND PERCEPTION**

#### Querv:

What values of OLD-TYPE, SLOT, and TYPE are there such that Copy Cat Motors should TYPE instead of OLD-TYPE, since doing so will maximize a SLOT?

#### Answer:

Copy Cat Motors should manufacture electric vehicle extended range battery pack systems using United States Patent 8450974 instead of manufacturing in which an electric vehicle standard range battery pack system plays a product role whose instrumental role includes Commercially Licensable Patent #4727343344, since doing so will maximize battery capacity

#### Detailed Justification: Linear Justification: Because:

- V Copy Cat Motors should manufacture electric vehicle extended range battery pack systems using United States Patent 8450974 instead of manufacturing in which an electric vehicle standard range battery pack system plays a product role whose instrumental role includes Commercially Licensable Patent #4727343344, since doing so will maximize battery capacity.
- If two patents allow very similar products to be created, and the lower cost one is better in some important respect, then to maximize that improvement, the maker should utilize the lower cost patent
- Manufacturing in which an electric vehicle extended range battery pack system plays a product role whose instrumental role includes United States Patent 8450974 and manufacturing in which an electric vehicle standard range battery pack system plays a product role whose instrumental role includes Commercially Licensable Patent #4727343344 are the same sort of event with respect to how their outputs function with respect to providing power modulo any differences due to United States Patent 8450974 and Commercially Licensable Patent #4727343844 being different.
- Copy Cat Motors is known to manufacture electric vehicle standard range battery pack systems using Commercially Licensable Patent #4727343344.
- United States Patent 8450974 describes a manufacturing process for electric vehicle extended range battery pack systems
- Commercially Licensable Patent #4727343344 describes a manufacturing process for electric vehicle standard range battery pack systems.
- Electric vehicle extended range battery pack systems have a greater amount of battery capacity than electric vehicle standard range battery pack systems.
- Every electric vehicle extended range battery pack system provides power for some electric car.
- Every electric vehicle standard range battery pack system provides power for some electric car.
- Copy Cat Motors is known to manufacture electric cars.
- The annual fee of Commercially Licensable Patent #4727343344 is larger than United States Patent 8450974's.
- ► II
  - X is a freely usable patent
  - and Y is a patent that requires a licensing fee,
  - then the annual fees of Y are larger than X's.
- Commercially Licensable Patent #4727343344 is a patent that requires a licensing fee.
- United States Patent 8450974 is a freely usable patent.
- If a certain type of patent becomes available under "open source" license, then each of those patents is freely usable.
- Elon Musk's open sourcing of every patent that Tesla is the patent holder of, was making something available via open source license.
- Elon Musk's open sourcing of every patent that Tesia is the patent holder of, occurred on June 13, 2014.
- Every patent that Tesla is the patent holder of was made available in Elon Musk's open sourcing of every patent that Tesla is the patent holder of
- United States Patent 8450974 is a patent.

#### [Refresh] [Debug Show Cycl

#### **Describes without errors** Describes with minor errors

#### Somewhat related to the image



A person riding a

motorcycle on a dirt road.





A skateboarder does







Two hockey players are fighting A group of young people playing a game of frisbee.





over the puck.



A little girl in a pink hat is

blowing bubbles.

A red motorcycle parked on the side of the road

#### http://googleresearch.blogspot.com/2014/11/a-picture-is-worth-thousand-coherent.html

on a couch.



A herd of elephants walking across a dry grass field.

-	2003	Very large datasets: genomic sequences
	2007	Very large datasets: WAC (web as corpus)
	2011	IBM Watson wins Jeopardy
	2012	US state of Nevada permits driverless cars
	2010s	Deep learning takes over: recommendation systems, image analysis, board games, machine translation, pattern recognition
	2017	Google AlphaGo beats the world's best Go player, Ke Jie AlphaZero learns boardgames by itself and beats the best programs
I	2019	IBM Debater assembles coherent arguments out of existing text
	2020	GPT-3 "language model" performs a variety of NL tasks without retraining
	2050 (roughly)	Single largest computer exceeds all humans in computational power.





http://yann.lecun.com/exdb/publis/pdf/sermanet-ijcnn-11.pdf



James the photographer - http://flickr.com/photos/22453761@N00/592436598/







IGITA



http://mashable.com/2015/08/29/computer-photos/#HcSaLeWp5mqU

### BUT WHAT HAPPENED TO KNOWLEDGE AND REASONING?

### &

### HOW CAN WE COMBINE THEM WITH SKILLS AND LEARNING?

Current AI falls significantly short of human-problem solving, including question-answering: it does not recursively decompose problems for solution, it does not follow that decomposition to assemble answers, and it does not store and apply salient background knowledge for decomposition, partial solution, or answer composition.



### SYMBOLIC KNOWLEDGE (TEXT-LIKE)



High precision and reliable reuse, Low expressiveness Low precision and usability High expressiveness

### SYMBOLIC KNOWLEDGE (PICTURE-LIKE)



High precision and reliable reuse, Low expressiveness Low precision and usability High expressiveness

## TIMELINE FOR THE FUTURE

- Large Scale Reasoning
- Assistants
- Better than human speech
- General Conversation
- Better than human translation
- Broadly capable robotics
- Broadly capably corporate AI
- General superintelligence

## BROAD CONCERNS

- Loss of constrained sensing and manipulation jobs
- Loss of constrained intellectual jobs
- Consequential grossly unequal access to resources
- Broad job loss
- Broad loss of human control over corporations
- Concentrated rent-seeking over AI value production
- General inutility of humans (how to recognize inherent value)
- Superintelligence general loss of human autonomy

### LEGAL QUESTIONS

- Legal Status of AI Systems
- "IP" and its effects on constructed minds
- Universal enforceability and legal decision making

# BROAD AI

### now and later

Michael Witbrock, PhD University of Auckland Broad Al Lab @Witbrock Website for BroadAl Lab soon: broadai.ac.nz