The Promise and Limitations of Al

Doug Lenat

doug@cyc.com

CYCORP

goto, chicago

🔰 follow us @gotochgo

Conclusions

□ There are 2 different ways to "power" an AI

- Statistics (induction, machine learning)
- Logic (deduction, abduction, causal models)

They have different strengths/weaknessesThey can be harnessed to work together

+ demos, fireworks, elephants, and more!

eventes:*essComp:socies=Esla eeom:*; ed block, bot sever sect:*; block soundes=DESAULT_v media.* socies: on the pet stream

dutification
additional end
add



I am the left brain. I am a scientist. A mathematician. Hove the familiar. I categorize. I am accurate. Linear. Analytical. Strategic. I am practical. Always in control. A master of words and language. Realistic. I calculate equations and play with numbers. I am order. I am logic. I know exactly who I am.

import sh.geom.*; import sh.net.*; import sh.media.*

import sh.utils.Time:

5 5100111) begin (?) // decimal adjust // results in nyb

6-8000100, 6 8001100, 6 807 6-3181100, 6 8101010, 6 811 6-30830101, 6 8011101, 6 810

> rend (a proceeding \$211). (7, per propage/dent in the set of the Automatic of the properties (b) (1, per propage) (and an all a b) (1, per properties (b) (1, per propage) (1, per properties (b) (1, per pro

R

LOGIC



PATTERNS

I am the right brain. Iam creativity. A free spirit. I am passion. Yearning, Sensuality. I am the sound of roaring laughter. I am tast. The feeling of sand beneath bare feet. I am movement. Vivid colors. I am the urge to paint on an empty canvas. I am boundless imagination. Art. Poetry. I sense. I feel. I am everything I wanted to be.

The Brittleness of Right-Brain Al's

Because you liked:



You will probably also like:





About 1,200,000 results (1.01 seconds)

May was the longest serving **home secretaries** in British history. Theresa May is Britain's first female Prime Minister since **Margaret Thatcher**. She is considered one of the most powerful Conservative women and was the longest serving **home secretaries** in British history. Mar 29, 2017

Theresa May - Prime Minister, Government Official - Biography.com https://www.biography.com/people/theresa-may-071116

Ø About this result

Feedback

People also ask	
Is Theresa May married?	\sim
Who is the new Prime Minister of UK?	~
Who is the prime minister of England right now?	~
What is the salary of the British prime minister?	~

Feedback

Theresa May - Wikipedia

https://en.wikipedia.org/wiki/Theresa_May -

Appointment. On 13 July 2016, two days after becoming Leader of the Conservative Party, May was appointed Prime Minister by Queen Elizabeth II, becoming only the second female British Prime Minister after **Margaret Thatcher**, and the first female British Prime Minister of the 21st century. Early career · Home Secretary · Minister for Women and ... · Prime Minister

The Brittleness of Left-Brain Al's



6

Skin Disease Diagnosis System:

- "Are there spots on the body?"
- "What color are the spots?"

...

• "Are there more on the trunk than elsewhere?"

Conclusion (92.71%): *"The patient has measles."*

<u>Veneer</u> of intelligence, not intelligence

ELIZA (DOCTOR)

Joe Weizenbaum, MIT

- Carl Rogers advocated *reflection Patient: "I swear a lot."*"How do you feel about the fact that you swear a lot?"
- Counted a lot on anthropomorphism
- Count on patient-to-doctor respect/awe/... *Patient: "My dog's mother died recently."*"Tell me more about your mother."

ELIZA (DOCTOR)

"blah blah blah mother blah blah blah" *Patient: "My dog's mother died recently."*"Tell me more about your mother."

ALICE (today!!)

- JUDGE11[9:26:0]I WILL ASK YOU SOME SIMPLE QUESTIONS AND GIVE SIMPLE COMMANDS
- **PROGRAM**[9:26:0]**Do you think your plan will succeed**?
- JUDGE11[9:26:16]WHAT IS THE COLOR OF A BLUE CAR
- **PROGRAM**[9:26:16]WHAT COLOR IS a blue car
- JUDGE11[9:26:35]WHERE IS SUE'S NOSE WHEN SUE IS IN HER HOUSE
- **PROGRAM**[9:26:35]Where it belongs.

Try searching the internet.

<u>Veneer</u> of intelligence, not intelligence

Each of the 2 types of AI suffers from a "brittleness bottleneck"

But each also has great power!

10

<u>Veneer</u> of intelligence, not intelligence

The power of statistics: More is More The Wisdom of the Crowd





"Diversity and independence are important because the best collective decisions are the product of disagreement and contest, not consensus"

– Francis Galton, 1906

The power of statistics

- More is More
- Low start-up cost

The power of logic demo

- Transparency: step-by-step reasoning
- 10,000-step-long proofs are still valid

"Diversity and independence are important because the best collective decisions are the product of disagreement and contest, not consensus"

– Francis Galton, 1906

The Promise and Limitations of AI: What Works

There are 2 different ways to "power" an Al

- Statistics (induction, machine learning)
- Logic (deduction, abduction, causal models)

□ They have <u>different</u> strengths/weaknesses

□ They can be harnessed to work together

ML suffers if there is little/no data (special case: very rare events)







(an application of Cyc[®])



ML can benefit by partnering with Symbolic Al if...

- there is little/no data
- showing a step-by-step justification is important
- we want the top *n* answers, and pro/con arguments for each

What factors argue <for/against> the conclusion that <ETA> <performed> <the Madrid train station attacks>?

What factors argue <for/against> the conclusion that <Al Qaida > <performed> <the Madrid train station attacks>?

What factors argue <for/against> the conclusion that <Israel> <performed> <the Madrid train station attacks>?

What factors argue <for/against> the conclusion that <Spain> <performed> <the Madrid train station attacks>?

Pro/con arguments for each of the top *n* **answers**

What factors argue <for/against> the conclusion that <ETA> <performed> <the Madrid train station attacks>?

For:

- ETA often executes attacks near national election
- ETA has performed multi-target coordinated attacks
- Over the past 30 years, ETA performed 75% of all terrorist attacks in Spain
- Over the past 30 years, 98% of all terrorist attacks in Spain were performed by Spain-based groups, and ETA is a Spain-based group.

Against:

- -ETA warns (a few minutes ahead of time) of attacks that would result in a high number civilian casualties, to prevent them. There was no such warning prior to this attack.
- -ETA generally takes responsibility for its attacks, and it did not do so this time.
- -ETA has never been known to falsely deny responsibility for an attack, and it did deny responsibility for this attack.

ML can benefit by partnering with Symbolic Al if...

- there is very little (or no!) training data
- showing a step-by-step justification is important
- we want the top *n* answers, and pro/con arguments for each
- multiple experts would disagree
- the best experts are *much* better than the average practitioner
- the user *thinks* they know what they want to ask, but often when they see it answered they revise their query (*n* times)



Super-brief "Story of Cyc"

There are two AI "power sources" (statistical ML, logic/rules models)

> Both have been fairly well understood, and used, for the last 50 years

Statistics/neural net machine learning systems were awaiting...

> CPU speed (Moore's Law; GPU's), storage cost, torrents of training data, convolution

Logic/Symbolic/Expert knowledge-based models were awaiting...

- solutions to context, inconsistency, inefficiency,... [+147 other obstacles]
- "priming the pump" so new apps aren't brittle
- CYC: Manhattan-Project-like effort to do just that
 - ➢ Now, 35 years and 2000 person-years and \$200MM later...





Specific data, facts, terms, and observations

Super-brief "Story of Cyc"

There are two AI "power sources" (statistical ML, logic/rules models)

> Both have been fairly well understood, and used, for the last 50 years

Statistics/neural net machine learning systems were awaiting...

> CPU speed (Moore's Law; GPU's), storage cost, torrents of training data, convolution

Logic/Symbolic/Expert knowledge-based models were awaiting...

- solutions to context, inconsistency, inefficiency,... [+147 other obstacles]
- "priming the pump" so new apps aren't brittle
- CYC: Manhattan-Project-like effort to do just that
 - Now, 35 years and 2000 person-years and \$200MM later...



t soment with an
a cortic valve replacement 2015."
acE natients received an ability 3 weeks in 2014 and -
"Which CCF parter d had myocarditis within 2
allograft and num - 5

Creating a CYC app is akin to training a new employee

The number of new knowledge and rules we had to add to Cyc, in order to do this entire application, was about 120,000

I.e., Cyc already knew 99.5% of what it eventually needed to know in order to carry out this clinical trial cohort-selection application for Cleveland Clinic Empirically, 95% of the rules and knowledge used in a typical question-answering "case", in this app, were <u>not</u> these new 120,000 rules but ones which Cyc had already known for years

fragments :1879 assertionsmedian date :2003/08/16mean date :2003/03/25pre Jan 07:1790 = 95.3%

combining :1020 assertionsmedian date :2003/11/14mean date :2003/06/05pre Jan 07:973 = 95.4%

ML can benefit by partnering with Symbolic Al if...

- there is very little (or no!) training data
- showing a step-by-step justification is important
- we want the top *n* answers, and pro/con arguments for each
- multiple experts would disagree
- the best experts are *much* better than the average practitioner
- the user *thinks* they know what they want to ask, but often when they see it answered they revise their query (*n* times)
- the task involves real-world objects/actions/common sense



"I'm not saying I want to forget deep learning... But we need to be able to extend it to do things like reasoning, learning causality, and exploring the world ." - Yoshua Bengio, not unlike what I have been saying since 2012 in The New Yorker. technologyreview.com/s/612434/one-o... ♡ 268 11:26 AM - Nov 21, 2018

ML can benefit by partnering with Symbolic Al if...

- there is very little (or no!) training data
- showing a step-by-step justification is important
- we want the top *n* answers, and pro/con arguments for each
- multiple experts would disagree
- the best experts are *much* better than the average practitioner
- the user *thinks* they know what they want to ask, but often when they see it answered they revise their query (*n* times)
- the task involves real-world objects/actions/common sense

It wouldn't fit in the suitcase because <u>it</u> was too big. It wouldn't fit in the suitcase because <u>it</u> was too small.

Fred was mad at Tom because **he** stole **his** lunch. Fred was mad at Tom so **he** stole **his** lunch.



When you become happy, you smile.
When someone you love accomplishes a milestone, that makes you happy.
Taking one's first step is a milestone.
Parents love their children.

the task involves real-world objects/actions/common sense

• Caption: "A man helping his daughter take her first step"



The Promise and Limitations of AI: What Works

There are 2 different ways to "power" an AI

- Statistics (induction, machine learning)
- Logic (deduction, abduction, causal models)
- They have different strengths/weaknesses
 ...and so do the various logics

□ They can be harnessed to work together



Romeo & Juliet

• Human reads it: He/she can answer an astronomical number of straightforward questions about it.

- Al reads it: Um...
 - Text processing
 - Limited logics (up to and including full first order predicate calculus)
 - Higher order logic

Romeo & Juliet: QA using text processing

- What are all the characters' names and place names, and all the references to each of those? (←using named entity recognition)
- Who knew Friar Lawrence? (←using link analysis)
- How did Romeo feel initially? (←using sentiment analysis)
- Who wore a mask in Act II? (←using text searching)
- What are the various different meanings that the word "Capulet" has? (←using latent semantic analysis)

Romeo & Juliet: QA using limited logics

Propositional Logic

Knowledge Graphs, Triple-Stores, Quad-Stores, RDF+OWL ontologies/OWL Lite, Property Graphs, Description Logics (OWL DL), Bayesian Networks, Frames (FRL), RLL, Prolog (Horn clauses), SAT inequalities, Partitioned semantic networks, CycL *circa* 1984, etc.

Full First Order Logic (Common Logic, OWL Full, KIF, CycL *circa* 1988)

Romeo & Juliet: QA using limited logics

Given: ΔRST is isosceles with vertex angle R.

Prove: $\angle S \cong \angle T$



Statements	Reasons
 ΔRST is isosceles with vertex angle R. 	1. Given
2. $\angle S$ is not congruent to $\angle T$.	2. Indirect proof assumption
3. $\overline{RS} \cong \overline{RT}$	3. Legs of an isosceles triangle are ≅.
4. $\angle S \cong \angle T$	 Angles opposite ≅ sides of a triangle are ≅.
 Contradiction between statement 2 and 4. 	 A pair of angles can't be congruent and non-congruent.
6. $\angle S \cong \angle T$	6. The negation is false.

Romeo & Juliet: QA using limited logics



What *separates* all those limited logics

Variables and quantifiers (and nesting thereof) $\forall x \exists y isa(x, American) \Rightarrow mother(x, y)$ $\exists y \forall x isa(x, American) \Rightarrow President(x, y)$ Function symbols whose values can be used as terms LessThan(YearsFn(83), AgeFn(MotherFn(Beyoncé))) Arity 3 relations: between(LA, SFO, SanDiego) Variable-arity relations: plus(-1, 34.1, 5, AgeFn(Beyoncé)) How to deal with negation (e.g., closed-world assumed) Defaults / exceptions / contexts Elaboration tolerance / non-monotonic reasoning

Different compromises on this tradeoff curve:



> Most of the time, even full First Order Logic is not expressive enough



But human-level QA requires higher order logic

- What does it mean for Romeo to say "This sentence is false."?
 Why?
- What are the relationships between Romeo and Balthasar?
 Why?
- Which statements by Juliet are lies?
 Why?
- What argument might have dissuaded Romeo from killing himself?
 Why?
- Does Juliet, at the time she drinks the feign-death potion, believe that Romeo will believe that she is alive from that time through to when she awakens? Why?
- Does Romeo believe that Juliet is dead after Balthasar tells him she is? Why?
- Which of Shakespeare's plays has Romeo heard of?
 Why?
- Does Romeo desire death after being told Juliet is dead?
 Why?

cm politics

investigation and leadership shift

45

CONGRESS

Rod Rosenstein's resignation

SUPREME COURT

stay away from talking read Charlottesville repor

2020 ELECTION

FACTS FIRST

Washington (CNN) — Deputy Attorney General Rod Rosenstein, who appointed special counsel Robert Mueller to investigate Russian interference in the 2016 presidential election, submitted his resignation letter to the White House Monday.

It is effective May 11.

"We enforce the law without fear or favor because credible evidence is not partisan, and truth is not determined by opinion polls," Rosenstein wrote. "We ignore fleeting distractions and focus our attention on the things that matter, because a republic that endures is not governed by the news cycle."



Related Article: READ: Deputy AG Rod Rosenstein's resignation letter

justice.

Rosenstein often found himself the target President Donald Trump, especially after hiring Mueller in 2017. Shortly after being confirmed as deputy attorney general, Rosenstein wrote a memo criticizing former FBI Director James Comey amid Comey's investigation into Russia's 2016 actions, which Trump used as part of his reasoning in firing Comey.

After Comey was fired, Rosenstein brought Mueller in as special counsel to continue the investigation, and Mueller also looked at the question of obstruction of Now Playing Who is Rod Rosenstein?

Now Playing Who is Rod Rosenstein? MORE FROM CNN





23andMe's most comprehensive DNA kit is finally on sale Pantone's Color of the Year is a vibrant start to 2019



Stefanie Sherk, Canadian model and actress, dies at 37



Unanswered questions as New York Times conducts review after anti-Semitic editorial cartoon

f 🌶 🖸 🔍

> Most of the time, even full First Order Logic is not expressive enough



> But that's okay, there are ways to drastically speed up HOL reasoning!







Specific data, facts, terms, and observations

Having expressivity and efficiency Meta-meta-rules .82 1% of the time: Choose/reevaluate your overall strategy .65 Meta-rules 9% of the time: Choose/reevaluate what tactic to take .97 Rules 90% of the time: Work on executing the current tactic

Having expressivity and efficiency

Micro-theories (contexts)

Contexts are first-class objects in our language

The Promise and Limitations of AI: What Works

There are 2 different ways to "power" an Al

- Statistics (induction, machine learning)
- Logic (deduction, abduction, causal models)
- They have different strengths/weaknesses
 - ...and so do the various logics

They can be harnessed to work together

Ways to connect ML + KBS to exploit their latent synergy

- Mutually-agreed-upon custom API (depictsA Image30492 Cat) → send CatRecognizer17: Image30492
 CatRecognizer17(Image30492) → send Cyc: (depictsA Image30492 Cat)
- KBS Generates vast amount of training data for ML system
 1. assertion (husband MelaniaTrump DonaldTrump) → "Melania Trump's husband is Donald Trump" → produce combinatorially many variant NLGs
 → hand the ML system the triple and that huge set of ways of saying it. ("assertion" can be a fact from, e.g., Wikidata, or hand-asserted rule in Cyc)
 2. infer trillions of simple logical entailments, by combining two or more assertions of type 1, and repeat step 1 for each



Ways to connect ML + KBS to exploit their latent synergy

- Mutually-agreed-upon custom API
 - (depictsA Image30492 Cat) → send CatRecognizer17: Image30492 CatRecognizer17(Image30492) → send Cyc: (depictsA Image30492 Cat)
- KBS Generates vast amount of training data for ML system
 - assertion (husband MelaniaTrump DonaldTrump) → "Melania Trump's husband is Donald Trump" → produce combinatorially many variant NLGs
 hand the ML system the triple and that huge set of ways of saying it. ("assertion" can be a fact from, e.g., Wikidata, or hand-asserted rule in Cyc)
 - 2. infer trillions of simple logical entailments, by combining two or more assertions of type 1, and repeat step 1 for each
- ML generates hypotheses
 KBS checks them for ill-formedness, then contradiction (with existing knowledge) reject the impossible ones

Even better: search for one or more causal explanations of that hypothesis

Example of Cyc $\leftarrow \rightarrow$ ML Synergy





Example of Cyc $\leftarrow \rightarrow$ ML Synergy





Ways to connect ML + KBS to exploit their latent synergy

- Mutually-agreed-upon custom API
 - (depictsA Image30492 Cat) → send CatRecognizer17: Image30492 CatRecognizer17(Image30492) → send Cyc: (depictsA Image30492 Cat)
- KBS Generates vast amount of training data for ML system
 - assertion (husband MelaniaTrump DonaldTrump) → "Melania Trump's husband is Donald Trump" → produce combinatorially many variant NLGs
 hand the ML system the triple and that huge set of ways of saying it. ("assertion" can be a fact from, e.g., Wikidata, or hand-asserted rule in Cyc)
 - 2. infer trillions of simple logical entailments, by combining two or more assertions of type 1, and repeat step 1 for each
- ML generates hypotheses
 KBS checks them for ill-formedness, then contradiction (with existing knowledge) reject the impossible ones

Even better: search for one or more **causal** explanations of that hypothesis

Still better: causal explanations that make independently testable predictions

Example of Cyc $\leftarrow \rightarrow$ ML Synergy





CAUSATION

UNDERSTAND the statistical result by finding plausible scenarios to account for it (or reject it by failing to find any rationalizations)

CORRELATION

UNCOVER the statistical result by integrating multiple data streams and detecting anomalous patterns



The Promise and Limitations of AI: What Works

There are 2 different ways to "power" an Al

• Statistics (induction, machine learning)

Logic (deduction, abduction, causal models)
 They have different strengths/weaknesses
 They can be harnessed to work together

Closing observation: they are more like
 2 developmental stages, not 2 species



ATAGCCATCACCAAAAAACTGCAACCTGCACTCGCCCATGTATTCCTTCATCTGCTGTCTG GCCCTGTCTGACCTGATGGTGAGTATAAGCTTGGTGCTGGAGACGGCTATCATCCTGCTG CTGGAGGCCAGGGGCCCTGGTGACCCGGGCCGCTTTGGTGCCAACAGCTGGACAATGTCATT GACGTGCTCATCTGTGGCTCCATGGTGTCCCAGTCTTTGGTGCCAGGCGTGTCCATGCCATA GACCGCTACATCTCCATCTTCTATGCATTACGTTATCAGCATCTTCTTCAGCACCCTCTTTATC GCACGACGGGCCATCGTGGGCATCTGGGTGGCCAGCATCTTCTTCTAGCCACGCTGCCCGG ACCTACTACAACCACAGCCGTCCTAATCTGCCTTGTCACCTTTCTTCTAGCCATGCTG ACCTACTACAACCACACAGCCGTCCTAATCTGCCTTGTCACCTTTCTTCTAGCCATGCTG ACCTACTACAACCACACAGCCGTCCTCATCTGCCTTGTCACCTTTCTTCTAGCCATGCTG ACCTACTACAACCACACAGCCGTCCTCATCTGCCTTGTCACCTTTCTTCTAGCCATGCTG ACCTACTACAACCACACAGCCGTCCTCATCTGCCTTGTCACCTTCTTCTAGCCATGCCATGCTG ATTGCCCAGCTCCAGAAGAGGCAGGGCTCCACCCGCCAAGGCTTCTGCCTTAAGGGTGCT GCCCTCATGGCAATTCTGTATGTCCACATGCTCACCCGAGCATACCAGCATGCTCAGGGG CCGGATGGCCTCTTCCTCAGCCTGGGGCTGGTGAGTCTGGTGGAGAATGTGCTGGTCGTG AACCTCTACCTCGTTCTCATCATCTTCAGCTCCATCGTCGACCCCCCTCATCTATGCTTT CTCACACTCATCGTCCTCTGCCCTCAGCACCCCACCTGCAGCTGCATCTTTAAGAACTTC GCCACCCTTACTATCATTCTGGGAATTTTCTTCCTGTGCTGGGGCCCCCTTCTTCCTGCAT



© Mayo Foundation for Medical Education and Research. All rights reserved.

Jesus Before the Caïf (Giotto, 1305).



The ceiling rafters show *some* convergent perspective, but it's imperfect: the ceiling has an inconsistent vanishing point and the dais is in parallel perspective, with no vanishing point.

http://www.webexhibits.org/sciartperspective/perspective1.html





http://www.renegadetribune.com/filippo-brunelleschi-rediscovery-perspective/





"There is no bilaterally symmetrical, nor excentricallyperiodic curve used in any branch of astronomy today which could not be smoothly plotted as the resultant motion of a point turning within a constellation of epicycles, finite in number, revolving upon a fixed deferent."



"The Mathematical Power of Epicyclical Astronomy", Norwood R. Hanson, Isis 51:2, U. Chicago Press, June, 1960.



Most people see this as blue & black.

Some people see this as white & gold.



Most people who see well in dim light see this as blue & black.

Most people who don't see well in dim light see this as white & gold.



If you rely a lot on your retinal cones: they do "subtractive mixing" and you perceive blue and black.

If you rely a lot on your retinal rods: they do "additive mixing" and you perceive white and gold.



CONCLUSIONS

- □ There are 2 different ways to "power" an AI: Statistics & Logic
- They have different strengths/weaknesses (and so do the logics)
- **They can be harnessed to work together**



Source States S

et us know

Remember to rate this session Thank you!

– Doug Lenat