AI4REASON: Artificial Intelligence for Large-Scale Computer-Assisted Reasoning

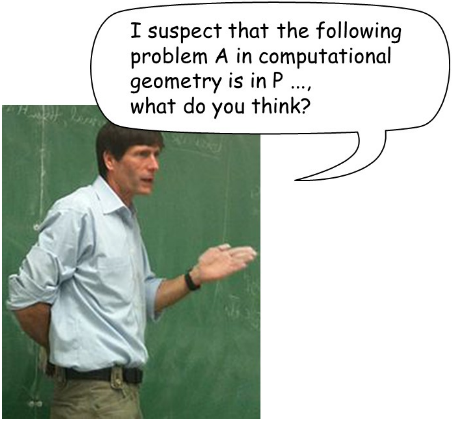
Josef Urban

Czech Technical University in Prague
Czech Institute for Informatics, Robotics and Cybernetics

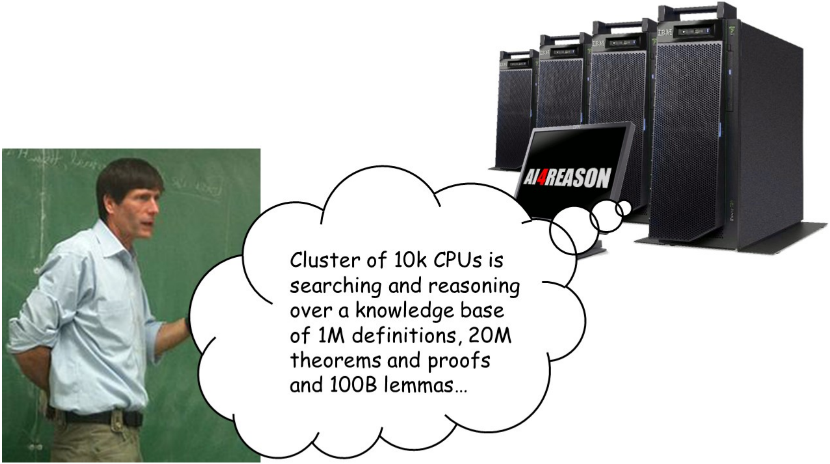
ERC Consolidator grant project No. 649043
09/2015 – 08/2020

European Research Council
Established by the European Commission

Not So Distant Future



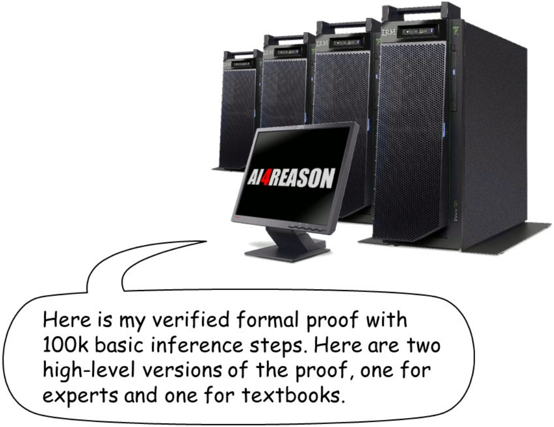
Not So Distant Future



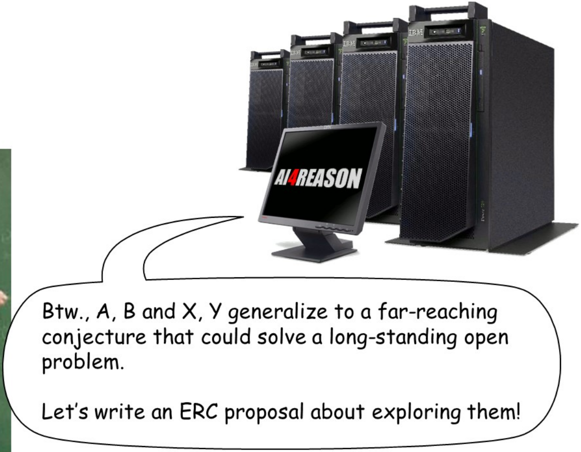
Not So Distant Future

Indeed, it is similar to a less known problem B number 13501 in my knowledge base. We can use a similar polynomial reduction to planar graphs as in B, and for the resulting constraint-solving problem we use a modified version Y of the $O(n^9)$ algorithm X published last year in Proc. of Indian Conf. on Graph Theory.

Not So Distant Future



Not So Distant Future



How Distant?

- 15 - 50 years, depending on our efforts
- Today's numbers about 100x smaller:
  - 10k-30k computer-understandable definitions
  - 200k-300k (small) theorems and proofs
  - 1B-10B primitive lemmas
- Covers roughly the Bc level in Math/CS, PhD level still far
- The main bottleneck:

WEAK AUTOMATION OF REASONING
OVER LARGE COMPUTER-UNDERSTANDABLE CORPORA

- This is where a breakthrough is necessary

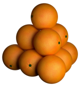
AI4REASON Goals

- Breakthrough in a hard problem in AI and reasoning: automatically proving theorems in complex theories
- Produce AI systems that combine learning and reasoning
- Thus help with automating verification of:
  - advanced mathematics and big proofs (Kepler conjecture)
  - software and hardware designs (seL4 OS microkernel)
  - advanced systems and designs (finance, industry, science)
- The idealized/perfect World of Math (Plato/Gödel): Interesting AI area – narrow or general AI?

Example: The Kepler conjecture

- J. Kepler (1611, Prague): The most compact way of stacking balls of the same size in space is a pyramid.

$$V = \frac{\pi}{\sqrt{18}} \approx 74\%$$



- Big proof: 300 pages + computations (Hales, Fergusson, 1998)
- Formal proof finished in 2014, 20000 theorems & proofs
- All of it computer-understandable and verified in HOL Light:
- `polyhedron s /\ c face_of s ==> polyhedron c`
- However, this took 20 – 30 person-years!
- Our AI methods can fully automate 40% of the proofs (2014)
- Similar verification efforts for bug-free compilers, OS, etc.

Sample of Formal Math: Irrationality of $\sqrt{2}$

```
theorem sqrt2_not_rational:
  "sqrt (real 2) ∉ ℚ"
proof
  assume "sqrt (real 2) ∈ ℚ"
  then obtain m n :: nat where
    n_nonzero: "n ≠ 0" and sqrt_rat: "¦sqrt (real 2)¦ = real m / real n"
    and lowest_terms: "gcd m n = 1" ..
  from n_nonzero and sqrt_rat have "real m = ¦sqrt (real 2)¦ * real n" by simp
  then have "real (m²) = (sqrt (real 2))² * real (n²)"
    by (auto simp add: power2_eq_square)
  also have "(sqrt (real 2))² = real 2" by simp
  also have "... * real (m²) = real (2 * n²)" by simp
  finally have eq: "m² = 2 * n²" ..
  hence "2 dvd m²" ..
  with two_is_prime have dvd_m: "2 dvd m" by (rule prime_dvd_power_two)
  then obtain k where "m = 2 * k" ..
  with eq have "2 * n² = 2² * k²" by (auto simp add: power2_eq_square mult_ac)
  hence "n² = 2 * k²" by simp
  hence "2 dvd n²" ..
  with two_is_prime have "2 dvd n" by (rule prime_dvd_power_two)
  with dvd_m have "2 dvd gcd m n" by (rule gcd_greatest)
  with lowest_terms have "2 dvd 1" by simp
  thus False by arith
qed
```

```
let SQRT_2_IRRATIONAL = prove
 (`~rational(sqrt(&2))`,
  SIMP_TAC[rational; real_abs; SQRT_POS_LE; REAL_POS] THEN
  REWRITE_TAC[NOT_EXISTS_THM] THEN REPEAT GEN_TAC THEN
  DISCH_THEN(CONJUNCTS_THEN2 ASSUME_TAC MP_TAC) THEN
  SUBGOAL_THEN `~((&p / &q) pow 2 = sqrt(&2) pow 2)`
    (fun th -> MESON_TAC[th]) THEN
  SIMP_TAC[SQRT_POW_2; REAL_POS; REAL_POW_DIV] THEN
  ASM_SIMP_TAC[REAL_EQ_LDIV_EQ; REAL_OF_NUM_LT; REAL_POW_LT;
               ARITH_RULE `0 < q <=> ~(q = 0)`] THEN
  ASM_MESON_TAC[NSQRT_2; REAL_OF_NUM_POW;
                REAL_OF_NUM_MUL; REAL_OF_NUM_EQ]);;
```

The AI4REASON Plan of Attack

WP1 AI for finding relevant knowledge in large formal corpora:

- How to capture similarity and analogy of ideas?
- How to learn from proofs, counter-examples and theories?

WP2 AI-based guiding methods for reasoning tools:

- How to efficiently apply the learned guidance?
- How to automatically learn the best reasoning strategies?

WP3 AI for suggesting plausible conjectures and concepts:

- What makes a good conjecture for a given problem?
- What concepts are good for a given problem?

WP4 Self-improving AI interleaving learning and deduction:

- How to explore easier problems to learn for harder ones?
- How to develop theories and gain most useful knowledge?

WP5 Deployment and Cross-Corpora Reuse:

- Deploy the methods as strong online services
- Translate informal math to formal

Combining Learning and Theorem Proving

- **high-level**: select relevant lemmas from a large library
- **high-level**: select good high-level strategies for a problem
- **low-level**: guide all inference steps of theorem provers
- **mid-level**: guide application of tactics to goals
- **mid-level**: invent suitable strategies for problem classes
- **mid-level**: invent suitable conjectures for a problem
- **mid-level**: invent suitable concepts/models for problems
- **proof sketches**: explore related theories to get proof ideas
- **theory exploration**: develop new theories by conjecturing
- **feedback loops**: (dis)prove, learn from it, (dis)prove more
- ...

Some Highlights So Far

- Won two divisions of the 2018 proving competition (CASC)
- 2017/18: Improved the best open prover by ML guidance
- 2018: 40% improvement of the leanCoP prover by reinforcement learning
- 2017-18: TacticToe – first ML-guided tactical system
- 2015-18: Blind Strategymaker - invent proving strategies
- First deep-learning based provers (with Google Research)
- 2015-18: Inf2formal – Translating informal math to formal, using grammar-based/semantic and neural systems
- Invited talks – Fields Inst., TYPES'18, Hales'60, AGI'18
- 2016 Google Research Award for JU
- AITP conference series started: `aitp-conference.org`
- AI/TP group at Google Research (2016), OpenAI - 2018?

What Can We Automatically Prove?

Nontrivial human-written proof that face of a polyhedron is polyhedron:

```
let FACE_OF_POLYHEDRON_POLYHEDRON = prove
 (`!s:real^N->bool c. polyhedron s /\ c face_of s ==> polyhedron c`,
  REPEAT STRIP_TAC THEN FIRST_ASSUM
   (MP_TAC o GEN_REWRITE_RULE I [POLYHEDRON_INTER_AFFINE_MINIMAL]) THEN
  REWRITE_TAC[RIGHT_IMP_EXISTS_THM; SKOLEM_THM] THEN
  SIMP_TAC[LEFT_IMP_EXISTS_THM; RIGHT_AND_EXISTS_THM; LEFT_AND_EXISTS_THM] THEN
  MAP_EVERY X_GEN_TAC
   [`f:(real^N->bool)->bool`; `a:(real^N->bool)->real^N`;
    `b:(real^N->bool)->real`] THEN
  STRIP_TAC THEN
  MP_TAC(ISPECL [`s:real^N->bool`; `f:(real^N->bool)->bool`;
                 `a:(real^N->bool)->real^N`; `b:(real^N->bool)->real`]
         FACE_OF_POLYHEDRON_EXPLICIT) THEN
  ANTS_TAC THENL [ASM_REWRITE_TAC[] THEN ASM_MESON_TAC[]; ALL_TAC] THEN
  DISCH_THEN(MP_TAC o SPEC `c:real^N->bool`) THEN ASM_REWRITE_TAC[] THEN
  ASM_CASES_TAC `c:real^N->bool = {}` THEN
  ASM_REWRITE_TAC[POLYHEDRON_EMPTY] THEN
  ASM_CASES_TAC `c:real^N->bool = s` THEN ASM_REWRITE_TAC[] THEN
  DISCH_THEN SUBST1_TAC THEN MATCH_MP_TAC POLYHEDRON_INTERS THEN
  REWRITE_TAC[FORALL_IN_GSPEC] THEN
  ONCE_REWRITE_TAC[SIMPLE_IMAGE_GEN] THEN
  ASM_SIMP_TAC[FINITE_IMAGE; FINITE_RESTRICT] THEN
  REPEAT STRIP_TAC THEN REWRITE_TAC[IMAGE_ID] THEN
  MATCH_MP_TAC POLYHEDRON_INTER THEN
  ASM_REWRITE_TAC[POLYHEDRON_HYPERPLANE]);;
```

We find an alternative shorter proof based on learning from the large library.

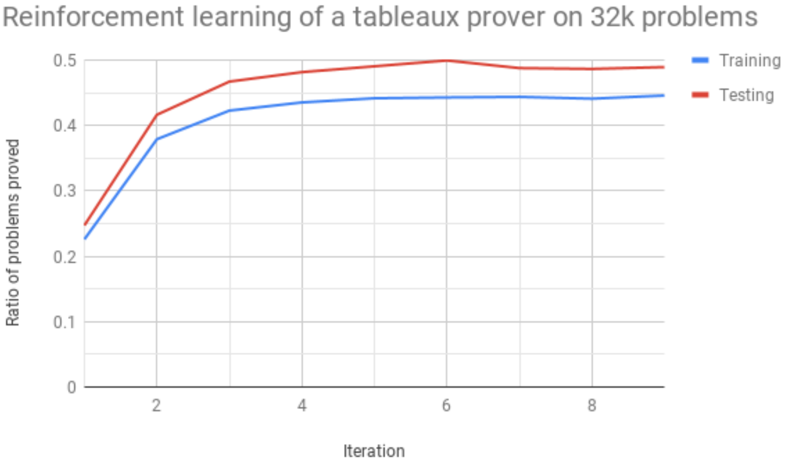
Statistical/Symbolic Guidance by Related Proofs (ProofWatch)

```
theorem Th36: :: YELLOW_5:36
for L being non empty Boolean RelStr for a, b being Element of L
holds ( 'not' (a "\/" b) = ('not' a) "/\" ('not' b)
         & 'not' (a "/\" b) = ('not' a) "\/" ('not' b) )
```

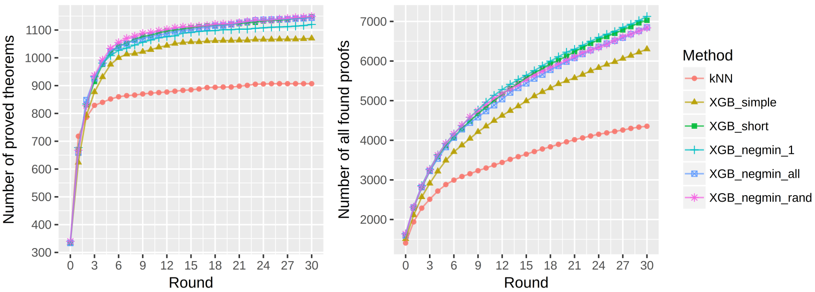
- Nontrivial proof of De Morgan laws for Boolean lattices
- Guided by continuous matching against 32 related proofs
- Most helped by a proof of a related statement for lower-bounded Heyting algebras

```
theorem :: WAYBEL_1:85
for H being non empty lower-bounded RelStr st H is Heyting holds
for a, b being Element of H holds
  'not' (a "/\" b) >= ('not' a) "\/" ('not' b)
```

Reinforcement Learning of a Tableaux Prover



Feedback Loop Generating Alternative Proofs



Informal2formal: Statistical/Semantic Parsing of Math

Strong-semantics probabilistic parser for HOL Light

Input the formula to parse. Separate symbols with spaces:

sin 0 = cos pi / 2 | Submit

debug: cache→ decode → 18 bigram&trigram features → 1024 nearest neighbours → 16 nyk parses → 12 distinct terms

Conjecture as HOL Light term:	Type info:	Automatically Provable?		Time
sin (&0) = cos pi / &2		disproved		(6.74s)
sin (&0) = cos (pi / &2)		yes	REWRITE_TAC [SIN_0; COS_PI2]	(0.87s)
csin (Cx (&0)) = Cx (cos (pi / &2))		yes	REWRITE_TAC [CSIN_0; COS_PI2]	(0.74s)
csin (Cx (&0)) = ccos (Cx (pi / &2))		yes	MESON_TAC [NUMERAL; CX_COS; CSIN_0; COS_PI2]	(0.76s)
Cx (sin (&0)) = ccos (Cx (pi / &2))		yes	MESON_TAC [SIN_0; NUMERAL; CX_COS; COS_PI2]	(0.70s)
Cx (sin (&0)) = Cx (cos (pi / &2))		yes	REWRITE_TAC [SIN_0; COS_PI2]	(0.80s)

exp (ii * x) = ii * (sin x) + (cos x)

Conjecture as HOL Light term:	Type info:	Automatically Provable?	
cexp (ii * A0) = ii * (csin A0 + ccos A0)	A0:real^2	no advice	
cexp (ii * A0) = ii * csin A0 + ccos A0	A0:real^2	yes	REWRITE_TAC [COMPLEX_ADD_SYM; CEXP_EULER]
cexp (ii * Cx A0) = ii * Cx (sin A0 + cos A0)	A0:real	disproved	
cexp (ii * Cx A0) = ii * (csin (Cx A0) + ccos (Cx A0))	A0:real	no advice	
cexp (ii * Cx A0) = ii * csin (Cx A0) + ccos (Cx A0)	A0:real	yes	REWRITE_TAC [COMPLEX_ADD_SYM; CEXP_EULER]

Neural Informal2formal: Performance after Training

Rendered LaTeX

Suppose s_8 is convergent and s_7 is convergent . Then $\lim(s_8+s_7) = \lim s_8 + \lim s_7$

Input LaTeX

```
Suppose $ { s _ { 8 } } $ is convergent and $ { s _ { 7 } }
$ is convergent . Then $ \mathop { \rm lim } ( { s _ { 8 }
} { + } { s _ { 7 } } ) \mathrel { = } \mathop { \rm lim }
{ s _ { 8 } } { + } \mathop { \rm lim } { s _ { 7 } } $ .
```

Correct

```
seq1 is convergent & seq2 is convergent implies lim ( seq1
+ seq2 ) = ( lim seq1 ) + ( lim seq2 ) ;
```

Snapshot-1000

```
x in dom f implies ( x * y ) * ( f | ( x | ( y | ( y | y )
) ) ) = ( x | ( y | ( y | ( y | y ) ) ) ) ) ;
```

Snapshot-2000

```
seq is summable implies seq is summable ;
```

Snapshot-3000

```
seq is convergent & lim seq = 0c implies seq = seq ;
```

Snapshot-4000

```
seq is convergent & lim seq = lim seq implies seq1 + seq2
is convergent ;
```

Snapshot-5000

```
seq1 is convergent & lim seq2 = lim seq2 implies lim_inf
seq1 = lim_inf seq2 ;
```

Snapshot-6000

```
seq is convergent & lim seq = lim seq implies seq1 + seq2
is convergent ;
```

Snapshot-7000

```
seq is convergent & seq9 is convergent implies
lim ( seq + seq9 ) = ( lim seq ) + ( lim seq9 ) ;
```

Snapshot-12000

```
seq1 is convergent & seq2 is convergent implies
  lim ( seq1 + seq2 ) = ( lim seq1 ) + ( lim seq2 ) ;
```

Team and Collaborations

- Chad Brown, Jan Jakubův, Martin Suda, Thibault Gauthier, Bartosz Piotrowski, Zarathustra Goertzel, Shawn Wang
- External scientific advisors
  - Prof. Stephan Schulz (Autom. reasoning, DHBW Stuttgart)
  - Prof. Robert Veroff (Autom. reasoning, U. of New Mexico)
  - Prof. Tom Heskes (AI, Radboud U. Nijmegen)
- Further Collaborations
  - Dr. Cezary Kaliszyk, U. of Innsbruck (ERC in 2016)
  - Dr. Jasmin Blanchette, VU Amsterdam (ERC in 2016)
  - Prof. Larry Paulson, U. of Cambridge (ERC in 2017)
  - Prof. Geoff Sutcliffe, U. of Miami
  - Dr. Christian Szegedy, Google Research
  - Prof. Herman Geuvers, Radboud U. Nijmegen
- over 20 research visits so far
- large related national grant awarded to JU in 2017

Future Potential - Science

- Use strong AI/reasoning and formal verification for:
- Science
  - Routinely verify complex math, software, hardware?
  - Make all of math/science computer-understandable?
  - Strong AI assistants for math/science?
- Examples
  - Automatically understand/verify/explain all arXiv papers?
  - Can we train a superhuman system like AlphaGo/Zero for math/physics? What will it take?
  - Can we prove that the Amazon Cloud cannot be hacked?
  - The same for critical government/private IT systems?

Future Potential - Society

- Use strong AI/reasoning and formal verification for:
- Society
  - Leibniz's dream: Let us Calculate! (solve any dispute)
  - J. McCarthy: Mathem. Objectivity and the Power of Initiative
  - AI/reasoning assistants for law/regulations
  - Verification of financial, transport/traffic systems, ...
  - Explainable and very securely verified systems
- Examples
  - Prove that two Paris metro trains will never crash?
  - Prove that a trading system doesn't violate regulations?
  - Prove that a new law is inconsistent with an old one?
  - Automatically debunk fallacies in political campaigns?

Possible Pitfalls and Avoiding Them

Keep informed, don't fall for the hype

- AI is much more than just (deep) learning/neural nets
- E.g., SAT/SMT/model-checking may be one of the biggest recent AI successes – Amazon, Facebook, Microsoft, etc.
- Don't expect miracles/singularity due to the current hype
- We can train image recognition & language models, but ...
- .. don't know what it takes to solve hard science problems
- However, some breakthroughs can happen quickly
- Researchers/society/lawmakers need to talk more/faster
- AI infrastructure for EU (CLAIRE) could serve this purpose

Possible Pitfalls and Avoiding Them

Don't let US, China, ...

- ... take away the best EU science minds
- In reasoning and formal methods EU is the leader!
- Make a deal with big AI companies to seriously support open university-based research
- Example: PRAIRIE institute in Paris,
- ... CLAIRE centers modelled after that?
- Infrastructure like CLAIRE very needed in countries like CR
- Larger brain-drain and local incompetence aggravating it
- Use such infrastructure to impose EU values on AI

Links and Impacts on Other AI Areas

- Main areas: Machine Learning, Automated Reasoning
- Needs advances in Representation Learning
- AI needs intuition, but also reasoning and explanations
- Impact on Formal Verification (SW, HW, etc.)
- Potentially on any (hard) science/thinking/arguing
- Alan Turing, 1950, AI:

“We may hope that machines will eventually compete with men in all purely intellectual fields.”

Outlook - Bets from 2014

- In 20 years, 80% of Flyspeck and Mizar toplevel theorems will be provable automatically (about 40% in 2014)
- The same in 30 years - I'll give you 2:1, In 10 years: 60% (getting there)
- In 25 years, 50% of the toplevel statements in LaTeX-written Msc-level math curriculum textbooks will be parsed automatically and with correct formal semantics

Outlook – Scientific Revolution

- (from a talk about Kepler and Hales)
- What did Kepler, Galileo & Co start to do in 1600s?
- What are we trying to do today?
- Kepler's Conjecture in Strena in 1611 (with many others)
- Kepler's laws, Newton, ..., age of science, math, machines
- ..., Hilbert, ..., Turing, ... age of computing machines?
- 1998 machine helps to find a proof of Kepler's Conjecture
- 2014 machine verifies a proof of Kepler's Conjecture
- ... 2050? machine finds a proof of Kepler's Conjecture?
- (no betting ;-)

Thanks and Advertisements

- Thanks for your attention!
- More examples of our systems at
  `http://ai4reason.org/demos.html`
- AITP – Artificial Intelligence and Theorem Proving
- April 7–12, 2019, Obergurgl, Austria,
  `aitp-conference.org`
- ATP/ITP/Math vs AI/Machine-Learning people, Computational linguists
- Discussion-oriented and experimental
- Grown to 60 people in 2018
- 2019: Hales, Goertzel, Gonthier, Marques Silva, Mikolov, Szegedy, Sutskever (?), ...