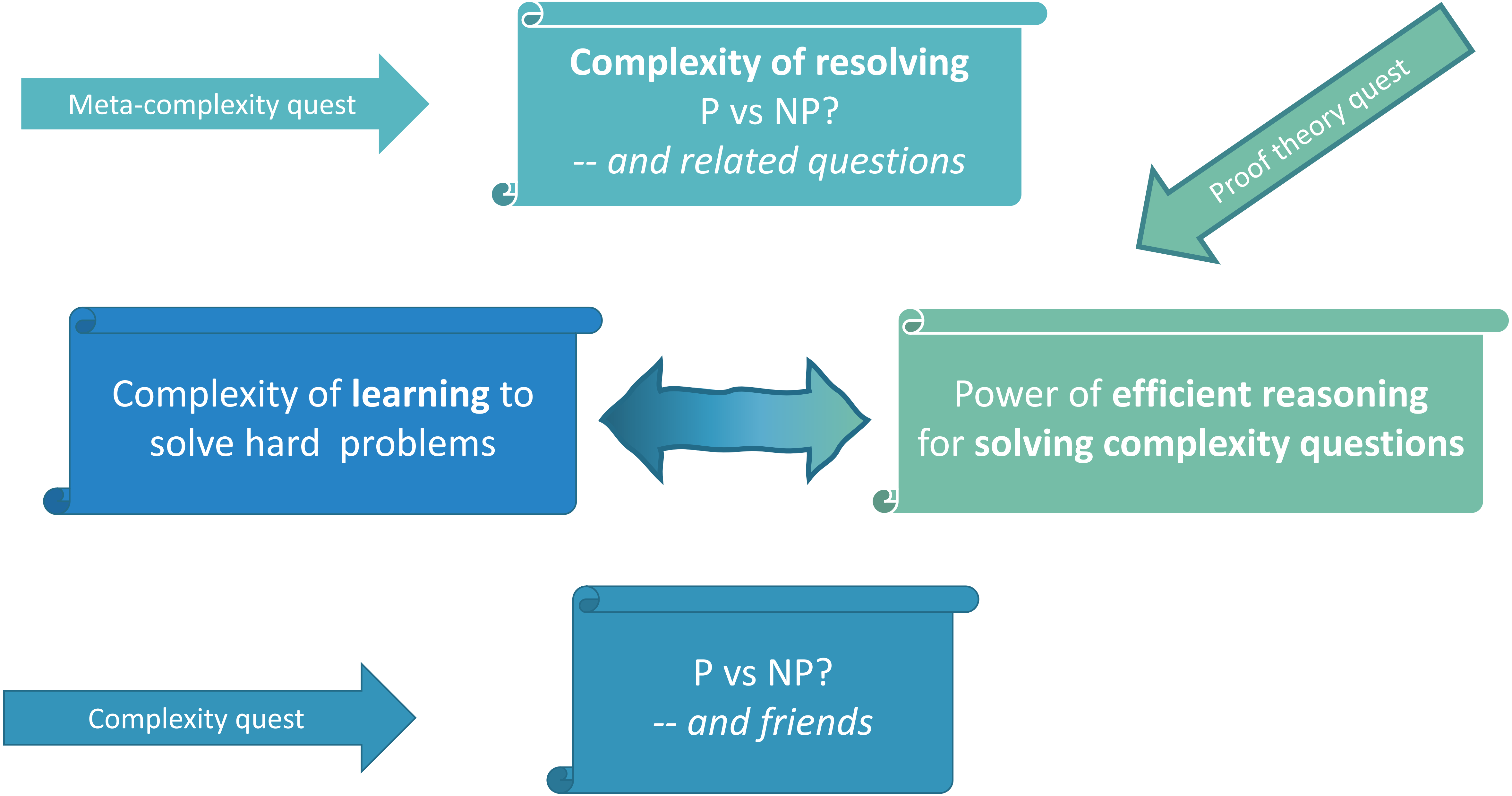
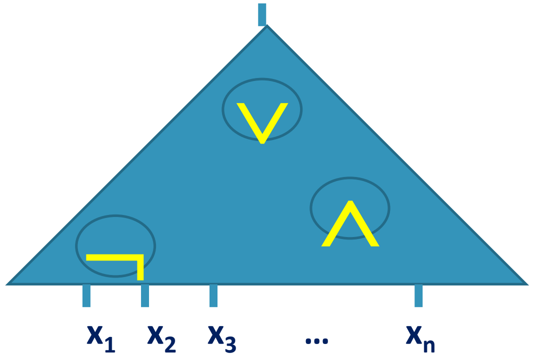
Learning from Bounded Arithmetic

Antonina Kolokolova (Memorial University of Newfoundland)
joint with Marco Carmosino, Valentine Kabanets and Igor C. Oliveira

Proof Theory Virtual Seminar, June 30, 2021


Meta-complexity quest
Complexity of resolving
P vs NP?
-- and related questions
Proof theory quest
Complexity of learning to solve hard problems
Power of efficient reasoning for solving complexity questions
Complexity quest
P vs NP?
-- and friends

Non-uniformity



- Proving P ≠ NP seems hard… So let's try to solve an even harder problem!

- Non-uniform computation:
  - A family of Boolean circuits, each solving the problem on inputs of specific length
  - Can solve an undecidable problem with a family of constant-size circuits!

- Efficient non-uniform computation = circuits are small
  - Small: polynomial size (number of gates)
  - Most Boolean functions require exponential size circuits [Shannon'49, Lupanov'58]
  - Best we know for a problem in NP: slightly more than 3n gates.

- Can all problems in NP be solved by polynomial-size circuits (is NP ⊂ P/poly)?
  - Can Boolean satisfiability (SAT) be solved by a family of circuits of size $O(n^k)$ for some constant k?

- Can all problems in P be solved by circuits of size $O(n^2)$, that is, is $P \subset SIZE[O(n^2)]$?
  - Or $SIZE[O\left(n^k\right)]$ for some other constant k? Even linear?

- What power of reasoning (weak system of arithmetic) do we need to prove these statements, if they are true?

An upper bound statement

"A language L is computable by a family of circuits of size n^k"

$$\forall n \;\; \exists C \;\; size(C) \le n^k \wedge \;\forall Z \;\; (|Z| = n \rightarrow C(Z) = L(Z)\;)$$

Input length

Circuit

Input string

- We want to study provability of such statements in weak theories of arithmetic.

*A **proof** of the existence of an object often provides more information about the object than just its existence.*

Herbrand's theorem

- Let T be a universal theory, and $\varphi(x,y)$ a quantifier-free formula.

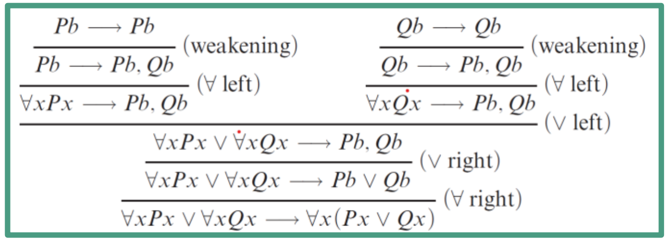
- Let $T \vdash \forall x\, \exists y\ \ \varphi(x,y)$.

Then there is a constant k and terms $t_1, \dots, t_k$ in the language of T such that

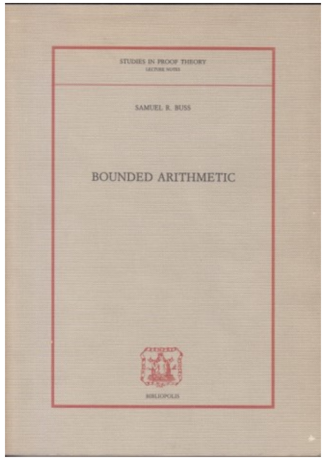
- $T \vdash \forall x\ (\varphi(x, t_1(x)) \vee \varphi(x, t_2(x)) \vee \cdots \vee \varphi(x, t_k(x)))$

Provability gives us a way to find (witness) existentially quantified objects

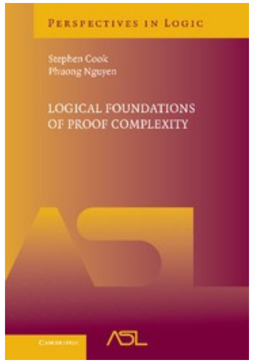
Bounded arithmetic



- Weak theories of arithmetic: all quantifiers bounded by terms
  - Cook's PV, Buss' S_2^1, T_2^1, etc, Jerabek's APC^1…
- Power of reasoning: definability + witnessing theorems
  - PV: polynomial-time reasoning
  - APC^1: probabilistic polynomial time reasoning
- We use two-sorted theories, following Cook/Nguyen
  - Universally axiomatized theories for many classes, including within P.
  - Have analogues of all first-order theories above:
    - "large numbers" become strings, "small numbers" are just numbers (indices)



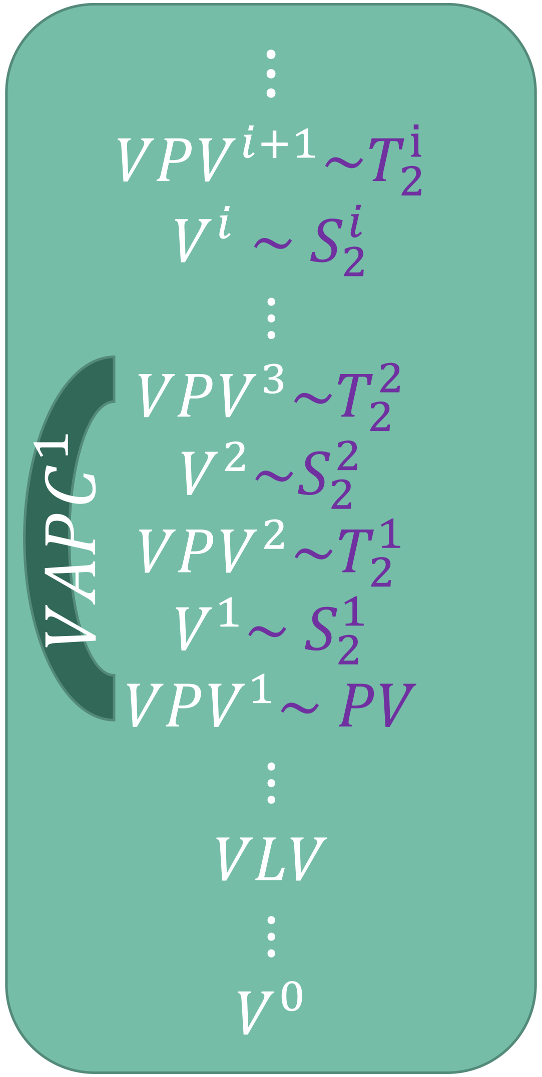
Two-sorted theories of bounded arithmetic



- Language: 2-sorted arithmetic (numbers + strings)
  - $0, 1, +, *, \leq$ for the number sort, length $|X|$ for strings, = for both sorts, $n \in X$.
- Axioms:
  - For numbers: standard ($x + 1 \neq 0$, etc)
  - For strings: defining length and string equality
    - $X(y) \rightarrow y < |X|$, $y + 1 = |X| \rightarrow X(y)$,..
  - **Comprehension**: for a class of formulas Φ (different Φ give different theories.)
    - $\exists X \leq n\, \forall z < n\, \big(X(z) \leftrightarrow \varphi(z)\big)$ for $\varphi \in \Phi$
    - Can also add induction on string length (provable in all our theories):
      - $X(0) \wedge \forall y < n\, \big(X(y) \rightarrow X(y+1)\big) \rightarrow X(n)$

Two-sorted theories of bounded arithmetic

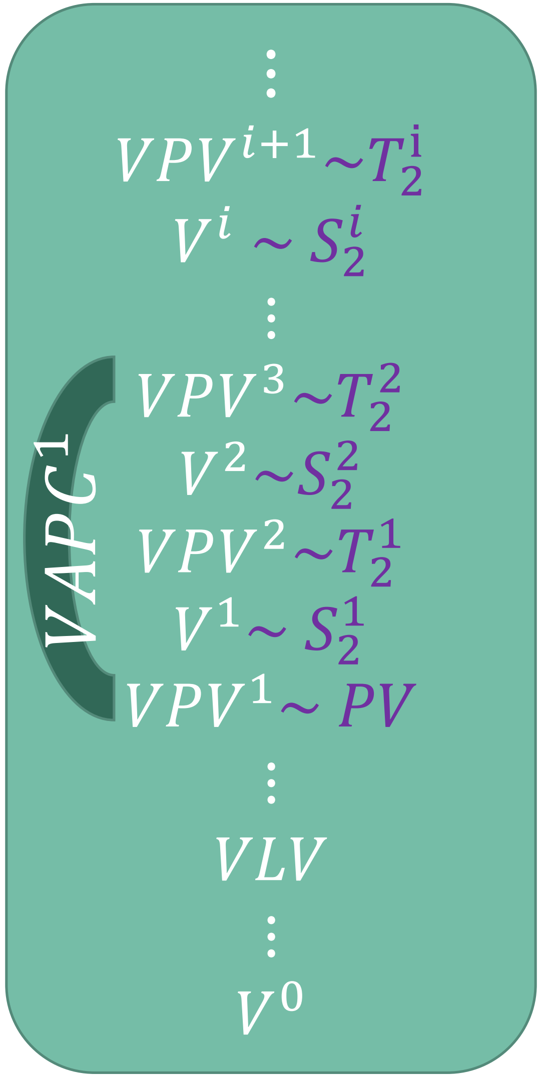
- Theory V^0: comprehension over formulas with no second-sort quantifiers
  - We actually need its universally axiomatized conservative extension, $\overline{V^0}$.
- V^1: $\Phi = \Sigma_1^B$, formulas with one (bounded) existential string quantifier
  - V^1 is equivalent to Buss' S_2^1 via RSUV isomorphism
  - Similarly, V^i with $\Phi = \Sigma_i^B$
- $V\mathrm{PV}^1$: universal theory for polynomial time
  - $\overline{V^0}$ + function symbols for all polytime functions with their defining axioms.
  - Similarly, theories for complexity classes other than P (eg VLV for logspace)
  - VPV^2, with functions from P^{NP}, is conservative over (two-sorted variant of) T_2^1
- $VAPC^1$: VPV+dual weak pigeonhole principle for all polytime functions.
  - dWPHP: $\forall n \, \forall S \, \exists Y \, (|Y| = n+1, \text{"} Y < 2^n + 2^n/n \text{"} \;\; \forall X \, (|X| = n) \;\; F(S,X) \neq Y$



Power of reasoning in bounded arithmetic

- VPV^1:
  - Captures polytime computation
  - Proves Cook-Levin, PCP theorem [Pich'15],…
  - Does not prove $P \subseteq SIZE[n^k]$ [Krajicek/Oliveira'17, BOK'20]
    - Previous conditional collapses [Cook/Krajicek'07]

- $VAPC^1$ [Jerabek'05,'07]
  - Captures probabilistic polytime
  - Formalizes much of known complexity theory
  - Proves $Parity \notin$ constant-depth Boolean circuits, etc



Buss's witnessing theorem

Let $\varphi(X, Y)$ be a formula with no second-sort quantifiers (Σ_1^B also OK)
If $V^1 \vdash \forall X \, \exists Y \, \varphi(X, Y)$, then there exists a polytime function F such that
$V^1(F) \vdash \forall X \, \varphi(X, F(X))$

- So provability of a $\forall\exists$ formula in V^1 gives a polynomial-time algorithm to witness the existential quantifier.
  - Works for multiple variables of both sorts for both quantifiers.
  - Scales up the polynomial-time hierarchy.
- Corollary: If $V^1 \vdash$ "$Primality\ \ is\ in\ P$“, get polytime algorithm for factoring

What if our formula is $\forall\exists\forall$, such as that upper bound statement?

KPT witnessing theorem [Krajicek/Pudlak/Takeuti]

Let $\varphi(X,Y,Z)$ be a formula with no string quantifiers, T a universal theory.

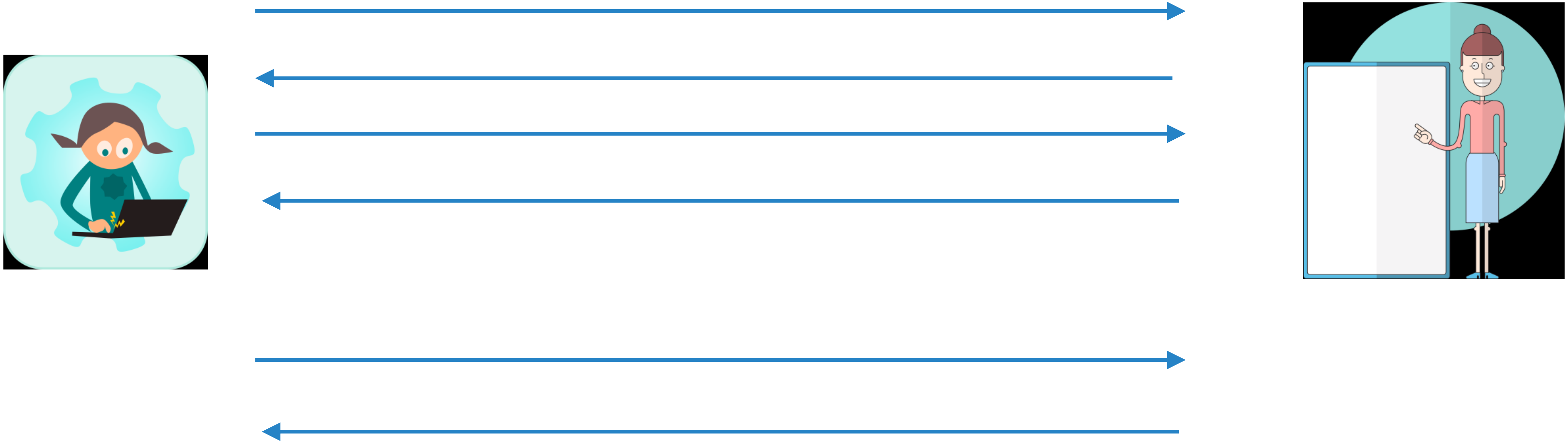
If $T \vdash \forall X\, \exists Y\, \forall Z\;\; \varphi(X,Y,Z)$, then there exists a finite sequence of terms $F_1, \dots, F_k$ in the language of T such that

$$\mathrm{T}\;\vdash \forall X\, \forall Z_1 \dots \forall Z_k\;\; \varphi(X, F_1(X), Z_1) \vee \varphi(X, F_2(X, Z_1), Z_2) \vee \cdots \vee \\ \vee\, \varphi(X, F_k(X, Z_1, \dots, Z_{k-1}), Z_k)$$

- Proved using Herbrand's theorem

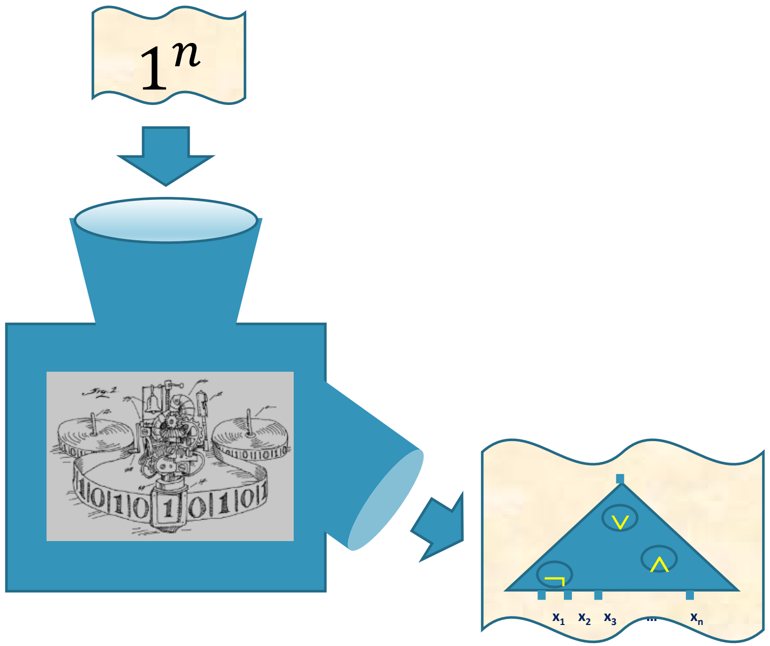
What kind of an algorithm for computing Y does KPT theorem give?

Student-teacher game / counterexample computation



In computational learning theory, this is known as "learning with equivalence queries" [Angluin'87].

Uniformity

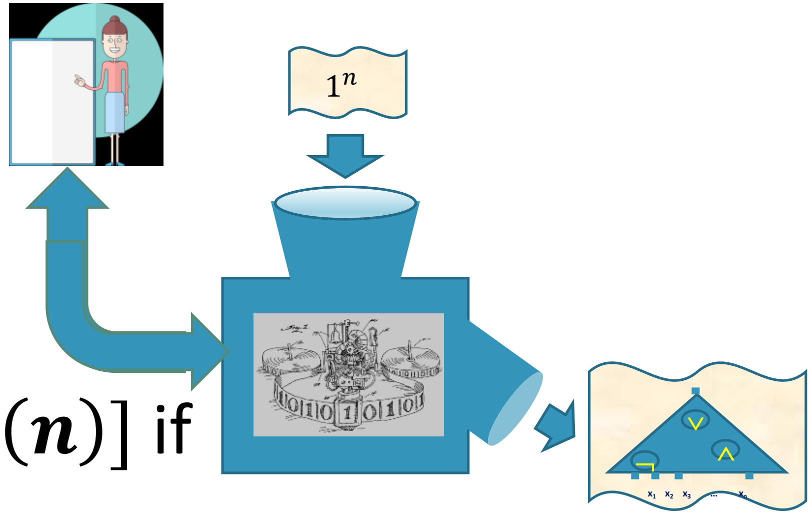


- A problem is in **CLASS-uniform $\boldsymbol{SIZE}[\boldsymbol{t}(\boldsymbol{n})]$** if there is an algorithm from CLASS that, given input length n, outputs a description of a circuit of size $\leq t(n)$ solving that problem on all inputs of length n.
  - Usually, n is in unary.

Algorithm ... P-uniform circuit family ... Non-uniform = $FZPP^{NP}$

- [Santhanam/Williams'14]:
  - For every $k \geq 1$ there is a problem $L \in P$ such that $L \notin P$-uniform $SIZE[n^k]$

$LEARN^{EQ}$-uniformity



- A problem L is in $\boldsymbol{LEARN^{EQ[r]}}$**-uniform** $\boldsymbol{SIZE[t(n)]}$ if there is a student-teacher game with
  - a polytime student algorithm
  - which, on input 1^n
  - after at most r equivalence queries to the teacher
  - outputs a circuit of size $\leq t(n)$ solving L on inputs of size n

Algorithm … P-uniform … $\boldsymbol{LEARN^{EQ}}$**-uniform** Non-uniform = $FZPP^{NP}$

Extending [SW'14] to $LEARN^{EQ}$-uniformity

For every $k \geq 1$,

1. $P \not\subseteq LEARN^{EQ[\boldsymbol{O(1)}]}$-uniform $SIZE[n^k]$

2. $NP \not\subseteq LEARN^{EQ[\boldsymbol{n^{o(1)}}]}$ -uniform SIZE$[n^k]$

3. $NP \not\subseteq LEARN^{EQ[\boldsymbol{n^{O(1)}}]}$ -uniform SIZE$[n^k]$

   or SearchSAT $\notin$ $LEARN^{SearchSAT-EQ[\boldsymbol{n^{O(1)}}]}$ -uniform SIZE$[n^k]$

- Also get lower bounds for randomized uniformity (eg. ZPP, FZPP…) and randomized $LEARN^{EQ}$

$LEARN^{EQ}$-uniformity bounds to unprovability

- Let T be a universal theory (eg VPV^1) and suppose that

$$T \vdash \forall n \;\; \exists C \;\; size(C) \leq n^k \;\wedge\; \forall Z \;\; (|Z| = n \;\rightarrow\; C(Z) = L(Z))$$

- By KPT witnessing theorem, get a $LEARN^{EQ[O(1)]}$-uniform family of circuits for L.
    - T cannot prove a truly non-uniform upper bound, only $LEARN^{EQ[O(1)]}$-uniform.

- If $L \notin LEARN^{EQ[O(1)]}$-uniform $SIZE[n^k]$, then T cannot prove a non-uniform upper bound for L.

Unprovability results

For all $k \geq 1$,

1. $VPV^1 \nvdash P \subseteq SIZE[n^k]$
2. $V^1 \nvdash NP \subseteq SIZE[n^k]$
   - Moreover, $V^1 \nvdash NP \subseteq ioSIZE[poly] \cap ioSIZE^{SAT}[n^k]$
3. $VPV^2 \nvdash P^{NP} \subseteq SIZE^{SAT}[n^k]$

Also in [BKO'20]

4. VLV does not prove that logspace has branching programs size $O(n^k)$

5. $VAPC^1 \nvdash SAT \in ioSIZE[poly]$, $or\ ZPP^{NP[O(1)]} \nsubseteq ioSIZE^{SAT}[n^k]$

Limits of provability

- $VPV^1 \nvdash (NP \nsubseteq P) \wedge (NP \subseteq ioSIZE[poly])$

- $VAPC^1 \nvdash (NP \nsubseteq BPP) \wedge (NP \subseteq ioSIZE[poly])$

Feasible reasoning cannot simultaneously prove non-uniform upper bounds and uniform lower bounds!

- Proof idea: KPT witnessing for the upper bound statement pitched against Buss' witnessing for the uniform lower bound.

$VPV^1 \nvdash (NP \nsubseteq P) \wedge (NP \subseteq ioSIZE[poly])$

- Upper bound statement: $NP \subseteq ioSIZE[poly]$
  - equivalent to $\mathrm{SAT} \in ioSIZE[poly]$

  $\forall m\, \exists n > m\ \exists C\ \ size(C) \leq n^k\ \wedge\ \forall(\varphi, w) \leq n\ (\varphi(w) \rightarrow \varphi(C(\varphi)))$

  - If provable, can apply KPT theorem to get a learning algorithm for SearchSAT

- Lower bound statement: $NP \nsubseteq$ P
  - Equivalent to $SAT \notin P$. For every function G that tries to solve SAT,

  $\forall m\, \exists n \geq m\, \exists(\varphi, w) \leq n\ (\varphi(w) \wedge \neg\varphi\big(G(\varphi)\big))$

  - If provable, can use Buss' witnessing to eliminate equivalence queries.

Open problems

- Show that $V^1 \nvdash P \subseteq SIZE[n^k]$, not just $NP \subseteq SIZE[n^k]$
- Show that $VAPC^1 \nvdash ZPP \subseteq SIZE[n^k]$
  - May need to get better randomized-uniformity lower bounds first
- Show independence of some natural circuit complexity statement from at least VPV^1.

A **proof** of the existence of an object often provides more information about the object than just its existence.

$\vdots$

$VPV^{i+1} \sim T_2^i$

$V^i \sim S_2^i$

$\vdots$

$VPV^3 \sim T_2^2$

$V^2 \sim S_2^2$

$VPV^2 \sim T_2^1$

$V^1 \sim S_2^1$

$VPV^1 \sim PV$

$VAPC^1$

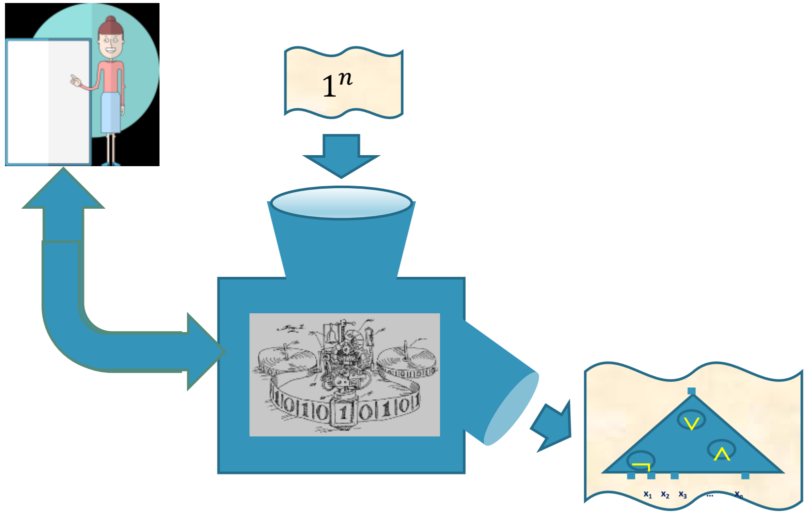
$\vdots$

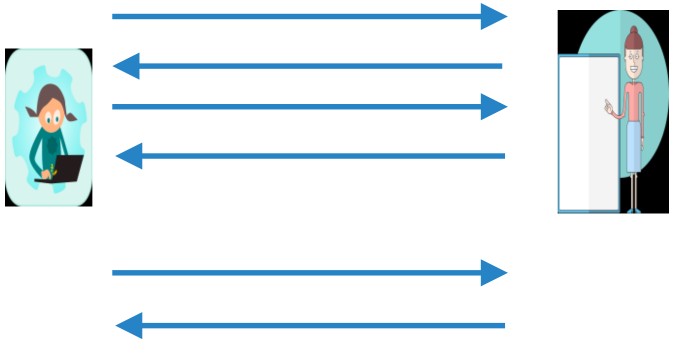
VLV

$\vdots$

V^0

Thank you!





Algorithm … P-uniform … $\boldsymbol{LEARN^{EQ}}$**-uniform** Non-uniform = $FZPP^{NP}$