Approximate counting in bounded arithmetic

Emil Jeřábek

jerabek@math.cas.cz

University of Toronto

New Directions in Proof Complexity, Cambridge, April 2006

Bounded arithmetic and complexity

There is a correspondence between theories and complexity classes:

- first-order theories (S^i_2, T^i_2): levels of polynomial hierarchy
- second-order theories: AC^0, TC^0, NC^1, L, $\dots$

Meaning of the correspondence:

- witnessing theorems, provably total computable functions
- reasoning about computation in the theories
- translation of open problems: inclusion of classes vs. conservativity of theories

Randomized classes

Theories of BA typically correspond to deterministic classes. What about probabilistic algorithms?

Examples: ZPP, BPP, AM

Connections to weak pigeonhole principle:

- [Wilkie] Σ^b_1-consequences of $S^1_2 + dWPHP(PV)$ are witnessed by $TFRP$-algorithms
- [J.] we can reason about FRP in $S^1_2 + dWPHP(PV)$

Goal of this talk: generalize to other classes of randomized algorithms

Approximate counting

We need to reason about probabilities, but we do not need exact results:

$$\Pr_{y<2^n}(A(x,y)\text{ accepts}) \geq \frac{3}{4} \text{ or } \Pr_{y<2^n}(A(x,y)\text{ accepts}) \leq \frac{1}{4}$$

Estimate of the probability within a small error suffices.

Equivalently: approximate counting of definable bounded sets

- given $X \subseteq [0, 2^n)$ defined by a poly-size circuit and $\varepsilon > 1/poly(n)$, approximate $|X|$ with accuracy $\varepsilon 2^n$

How to express it in bounded arithmetic?

Reminder

First-order bounded arithmetic [Buss 1986]:

- language: $\langle 0, S, +, \cdot, \leq, \#, |x|, \lfloor \frac{x}{2} \rfloor \rangle$
- Σ_i^b and Π_i^b formulas: count alternations of bounded quantifiers, ignore sharply bounded quantifiers
- $S_2^i = BASIC + \Sigma_i^b$-$PIND$

$$\varphi(0) \wedge \forall x \leq a \, (\varphi(\lfloor \tfrac{x}{2} \rfloor) \to \varphi(x)) \to \varphi(a)$$

Equational theory PV [Cook 1975]:

- function symbols for all poly-time algorithms
- derivation rule simulating open $PIND$

Theory PV_1 [KPT 1991]: first-order variant of PV

Dual weak pigeonhole principle

- $PHP^a_b(f)$: if we put a pigeons in $b < a$ holes, some hole must accommodate two pigeons
- $dPHP^a_b(f)$: if we put a pigeons in $b > a$ holes, some hole remains vacant

$$\exists y < b \, \forall x < a \, f(x) \neq y$$

- Weak $PHP/dPHP$: a and b differ by (much) more than 1

For our purposes: $dWPHP(f)$ means

$$\forall e \, \forall a > 0 \, dPHP^{a\,|e|}_{a(|e|+1)}(f)$$

Over S^1_2, $dWPHP(PV)$ is equivalent to $\forall a > 1 \, dPHP^a_{a^2}(PV)$, but we want PV_1 as a base theory

Counting functions

Consider $X, Y \subseteq 2^n$. We have: $|X| \geq |Y|$ iff there exists a function f which maps X onto Y

$$f : X \twoheadrightarrow Y$$

We could use it as a definition of counting, but a modification is needed to ensure

- f is computable by a poly-size circuit, if X and Y are,
- $PV_1 + dWPHP(PV)$ proves the existence of such counting functions

Counting functions (cont'd)

Definition. Let $X, Y \subseteq 2^n$ and $\varepsilon \in [0,1]$. We say that the size of Y is approximately less than the size of X with error ε, written as $Y \preceq_\varepsilon X$, if there exist

- a number $v > 0$, and
- a circuit C which maps v copies of the disjoint union of X and $[0, \varepsilon 2^n)$ onto v copies of Y

$$C \colon v \times (X \mathbin{\dot\cup} \varepsilon 2^n) \twoheadrightarrow v \times Y$$

$X \approx_\varepsilon Y$ means $X \preceq_\varepsilon Y \wedge Y \preceq_\varepsilon X$.

Counting is a special case of comparison:

$$X \approx_\varepsilon s \;:\Leftrightarrow\; X \approx_\varepsilon [0, s)$$

Nisan-Wigderson generator

The pseudorandom generator $NW_f \colon 2^\ell \to 2^n$

- seed length $\ell = O(\log n)$
- computable in time $poly(n)$
- “fools” circuits $C \colon 2^n \to 2$ of size $poly(n)$
- needs a table of a hard Boolean function f in $\Theta(\log n)$ variables

[NW 1994] $P = BPP$, if there exists $\varepsilon > 0$ and a uniform family of Boolean functions $f_k \colon 2^k \to 2$ which cannot be approximated by circuits of size $2^{\varepsilon k}$ with advantage $2^{-\varepsilon k}$.

Nisan-Wigderson generator (cont'd)

We use the NW generator to construct counting functions.

- We don't need uniformity. Nonuniformly, Boolean functions with exponential hardness exist, and $PV_1 + dWPHP(PV)$ proves it.
- The behaviour of the generator can be analyzed constructively: the conclusion

$$\left|\Pr_{x<2^n}(C(x)=1) - \Pr_{u<2^\ell}(C(NW_f(u))=1)\right| \leq 1/poly(n)$$

is witnessed by counting functions computable by small circuits, which can be extracted from the proof.

Existence of counting functions

Theorem. The following is provable in $PV_1 + dWPHP(PV)$.

Let X be a subset of 2^n definable by a Boolean circuit C, and $0 < \varepsilon < 1$ s.t. $2^{1/\varepsilon}$ exists. Then there exists $s \leq 2^n$ s.t.

$$X \approx_\varepsilon s.$$

More precisely, there exists $v \leq poly(n\varepsilon^{-1}|C|)$ and circuits G_0, H_0, G_1, H_1 of size $poly(n\varepsilon^{-1}|C|)$ such that

$$G_0\colon v(s + \varepsilon 2^n) \twoheadrightarrow v \times X \qquad G_1\colon v \times (X \,\dot\cup\, \varepsilon 2^n) \twoheadrightarrow vs$$

$$H_0\colon v \times X \hookrightarrow v(s + \varepsilon 2^n) \qquad H_1\colon vs \hookrightarrow v \times (X \,\dot\cup\, \varepsilon 2^n)$$

$$G_0(H_0(x)) = x \qquad G_1(H_1(y)) = y$$

for every $x \in v \times X$ and $y < vs$.

Applications

The rest is (mostly) easy—we can do in $PV_1 + dWPHP(PV)$:

- counting trivia: inclusion-exclusion principle, Chernoff bound, ...
- formalize randomized complexity classes: BPP, $prBPP$, APP, MA, $prMA$
  - basic definitions
  - amplify success probability
  - simulate randomness by nonuniformity
  - place it on the correct level of PH

Everything relativizes. We can do AM and $prAM$ in $T^1_2 + dWPHP(FP^{\Sigma^b_1})$.

Definability questions

Are all problems from the above mentioned classes "provably total" in $PV_1 + dWPHP(PV)$?

- syntactic classes ($prBPP$, $prMA$): trivial/meaningless
- APP: yes, it also turns out to be a syntactic class
- semantic classes (FRP, BPP, MA):
  - if true (for whatever theory), relativizing techniques cannot show it [Thapen]
  - can be reduced to provability of $\forall\Sigma_1^b$-sentences

Problems

We cannot count “sparse” sets, which arise in

- combinatorial arguments: Ramsey theorem, tournament principle, ...
- interactive protocols: graph nonisomorphism, $IP[O(1)] = AM$
- ...

Q: Does Sipser-style counting via hash functions work in bounded arithmetic?

That's the end.

Thank you for attention!