Information in propositional proof and algorithmic proof search

J. Krajíček (Charles U.)

[seminar 1.III.2021]

Ref's:

- there is a preprint with the same title on my web page

- [Sec. x.y], [p...] refer to my 2019 book

- these slides are on my web page as well

Motivation:

"Is there an optimal way to search for propositional proofs?"

- formulate precisely
- relate to proof complexity
- [Secs. 1.5 & 21.5]

notation / basics:

- a Cook-Reckhow proof system (a pps):

$$P\text{-time } P : \{0,1\}^* \twoheadrightarrow TAUT$$

onto

↖ De Morgan !.

- the lengths-of-proof f.:

$$s_P(\tau) := \min\{|w| \mid P(w) = \tau\}$$

- 2 key problems :→ $NP \overset{?}{=} coNP$: $\left[\exists_{pps}\, P : s_P(\tau) \le |\tau|^{O(1)}\right]$

↘ $\exists^{?}$ optimal P : $s_P(\tau) \le s_Q(\tau)^{O(1)}$

call Q.

Definition

A <u>proof search algorithm</u> is a pair (A, P) s.t.:

- P is a pps.
- A is a deterministic alg. that stops on all inputs
- for all $\tau \in TAUT$: $P(A(\tau)) = \tau$

$\uparrow$ w.l.o.g.

$time_A(w) :=$ "the time A needs to stop on w"

Perhaps the canonical way how to compare (A, P) with (B, Q) is the time:

$$(A, P) \geq_t (B, Q) \overset{\leftarrow}{\underset{df.}{\rightarrow}} time_A(\tau) \leq time_B(\tau)^{O(1)},$$

all $\tau \in TAUT$

↑ a quasi-ordering

do not care for $u \notin TAUT$

<u>Formal Q</u>: Is there a $\geq_t$-max (A, P)?

Lemma: For every ppt P there is det. alg. A_P s.t. $(A_P, P) \geq_\epsilon (B, P)$, all alg's B.

I.e. (A_P, P) is time-optimal among all (B, P).

Pf: ... an instance of Levin's universal search. □

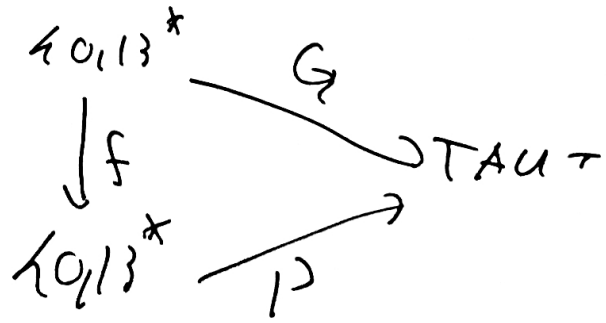
Notation: A_P - one such fixed universal alg

Theorem : Let P be any pps containing R (= reduction) and having the property that $\exists c \geq 1$ $\forall \mathcal{I}$ $\forall \mathcal{I}'$ obtained from $\mathcal{I}$ by substituting constants for some atoms it holds $S_P(\mathcal{I}') \leq S_P(\mathcal{I})^c$.

Then P is p-optimal $\Leftrightarrow$ (A_P, P) is tem-optimal (i.e. $\mathcal{I}_t$-max) among all (B, G).

□

$P \geq_p Q \Leftrightarrow \exists$ p-hom f :



p-simulation

Pf - sketch :

$\Rightarrow$ if (B, G) is a proof scheme adv. for

$$\left.\begin{array}{l} (f \circ B, P) \geq_t (B, G) \\ \text{but} \quad (A_p, P) \geq_t (f \circ B, G) \end{array}\right\} \Rightarrow (A_p, P) \geq_t (B, G).$$

$\Leftarrow$ Take any pps G and fix $\langle Ref_G \rangle_n$, $n \geq 1$.

Key property: If there is p-time $g : 1^n \xrightarrow{g}$ P-proof of $\langle Ref_G \rangle_n$

then $P \geq_p G$.

Define (B, G'): $B(\tau)$ $\longrightarrow$ if $\tau = \langle Ref_G \rangle_n$ then $:= 1^{(n)}$

$\searrow$ if not, $B(\tau)\Sigma$ is any G-proof of τ

↑

i.e. produced by an arbitrary fixed adv.

By the hypothesis $(A_p, P) \geq_t (B, G')$ so

$$\mathrm{time}_{A_p}(\langle \mathrm{Ref}_Q \rangle_n) \leq t_B(\langle \mathrm{Ref}_G \rangle_n) = n^{O(1)},$$

and (by the key property) $P \geq_p G' \geq_p Q$.

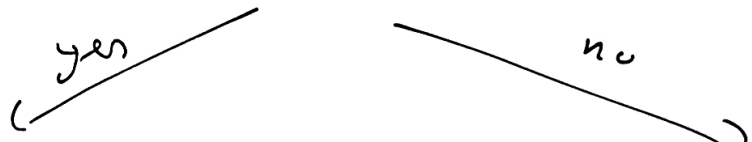
□

Remark: There is a second proof in the preprint.

Is $\geq_t$ the right quasi-order?

E. situation: (A, R): any resolution searching algo

(B, ER): - B checks if $T = PHP_n$ for some n ($\leq |T|$)



produces ER-proof of PHP_n

produces $A(T)$!

[Can we really claim that B is better than A? If only remembers one type of a simple fla - ?]

Informal ideas

- A and B should be compared on far Σ "where they actually do something non-trivial"

- We may consider a quasi-ordering that allows to take out simple flips as PHP_n : but it is unnatural
  → it may not help in breaking the $\subseteq$ direction in the theorem.

- Drawback: the decision to take PHP_n is based on the ambient uniform sequence $\langle PHP_n \rangle_{n \geq 1}$, and not on $P \# P_n$ alone

algorithmic information theory

- every $e \in \{0,1\}^*$ codes a T. machine
- a p-time universal machine $U(e,u,1^{(t)})$: simulates e on input u for time $\le t$, if it does not stop produces e.g. 0.

$$Kt_e(w \mid u) := \min\{|e| + \log t \mid U(e,u,1^{(t)}) = w\}$$

$$Kt(w) := Kt_e(w \mid 0)$$

the time-bounded Kolmogorov complexity

(def. by Levin)

Simple extensions :

$$\log(|u|) \leq K(u|u) \leq K(u) \leq |u| + (\log(|u|) + O(1))$$

- $\log(|u|)$: time need to write u down
- $K(u|u)$: e : hard-wired in u; time $O(|u|)$
- $K(u)$
- $|u| + (\log(|u|) + O(1))$: fixed machine

We would like to have :

$$Kt(u) \leq Kt(u|u) + K(u)$$

↑ machine e_2 ↑ machine e_1

compose e_1 with e_2

But the length of the code of (e_1, e_2) is not $|e_1| + |e_2|$ but rather

$$|e_1| + |e_2| + O(\log(|e_1|) + \log(|e_2|))$$

[there are codes where $|(e_1, e_2)| = |e_1| + |e_2|$ but they are $\log$-longer than trivial codes]

We get:

$$Kt(u) \leq Kt(u|u) + Kt(u) + \log\text{-terms}$$

and a bit more generally

$$Kt(u|u) \leq Kt(u|v) + Kt(v|u) + \log \text{ terms.}$$

[already in Kolmogorov 1980s]

Definition

Let P be a pps. The <u>information-efficiency function</u>

$i_P : TAUT \to \mathbb{N}$ is defined by:

$$i_P(\tau) := \min\{ K\varepsilon(u \mid \tau) \mid P(u) = \tau \}.$$

[7]

minimal amount of information any P-proof of τ has to contain, <u>knowing what τ is</u>

Lemma 3.2: $P \geq_p Q \Rightarrow i_P(\tau) \leq O(i_Q(\tau))$, all τ. ☐

Lemma 3.3: Let (A, P) be any proof scheme. Then for $\tau \in TAUT$:

$$i_P(\tau) \leq Kt(A(\tau)|\tau) \leq |A| + \log(time_A(\tau)).$$

In particular, $time_A(\tau) \geq \Omega(2^{i_P(\tau)})$. ☐

Lemma 3.4: $\forall$ pps P $\exists$ alg B, $\forall \tau \in TAUT$:

$$Kt(B(\tau)|\tau) = i_P(\tau)$$

and

$$time_B(\tau) \leq 2^{O(i_P(\tau))}.$$

$\forall$ via universal search

↖ a form of information-automatizability ☐

Def: $P \geq_i Q \iff_{\text{def.}} i_P(\tau) \leq O(i_Q(\tau))$, all τ.

↑

a quasi-ordering coarser than $\geq_P$

But nevertheless:

Theorem: (i) (A_P, P), (B_P, P) are $\mathcal{R}_E$-equivalent

(ii) a pps P is $\geq_i$-max $\Rightarrow$ P is p-optimal.

□

Prf: analogous to Re than about $\mathcal{R}_t$. □

Summary:

- The existence of optimal proof systems (AIP), either in the time-sense or info-sense

  (i) depends only on P

  (ii) is equivalent to $\exists$ p-optimal pps.

I.e. in either formalism we do not get a new problem.

But: the ip(—) measure allows us, in principle, to target time lower bounds for individual fns and not as asymptotic properties of entire sequences, similarly as we can target size l. bounds

Size vs. information

The time$_A(\tau)$, for (A,P), is bounded by $sp(\tau)$

but – in principle – $ip(\tau)$ can give <u>super-poly</u>

<u>better</u> bounds (via Ch 3.3) of:

$$ip(\tau) \notin O(\log(sp(\tau)))$$

That is, when $ip(\tau) \geq \omega(\log(sp(\tau)))$ then

<u>information</u> is better than <u>size</u>.

can we get such flas?

Thm: A pps P is non-automatizable iff

$\exists$ infinite $\mathcal{T} \subseteq TAUT$ of unbounded size s.t.

$$i_P(\tau) \geq \omega(\log s_P(\tau)), \text{ for } \tau \in \mathcal{T}.$$

W.l.o.g.: $|\tau| = n$, $n \to \infty$

$$s_P(\tau) \leq n^{O(1)} \quad \text{and} \quad i_P(\tau) \geq \omega(\log n)$$

Can we establish this for individual (specific) τ?

Proofs of non-autom. w.r.t p-reductions of bounded sets to a set of tt with short proofs s.t. $\mathcal{T}$ maps to far with long (as ∞) proofs. Is asymptotic, not for individual flas.

A calculation:

$$Kt(\pi) \le Kt(\pi \mid \tau) + Kt(\tau) + \log\text{-terms}$$

Assume: $(|\tau| = m)$

(1) $Sp(\tau) \le m^{O(1)}$ [τ has short proofs]

(2) $Kt(\tau) \le O(\log m)$ [τ is <u>simple</u>]

(3) $P(\pi) = \tau \to Kt(\pi) \ge \omega(\log m)$ [all proofs are <u>complex</u>]

Then (*) yields:

$$\omega(\log m) \le Kt(\pi \mid \tau), \text{ all } \pi.$$

For non-simple τ (condition (2)) this extension does not work but we can retrofit the good in a way that makes sense for complex τ as well:

$$\mathrm{It}(\tau : \pi) := K\mathrm{t}(\pi) - \min_{p} K\mathrm{t}(p|\tau) \le K\mathrm{t}(x) - c_2(\tau)$$

↑ information that τ conveys about π [Kolmogorov's notion]

and we want $\mathrm{It}(\tau : \pi)$ as small as possible, say $O(\log n)$.

Candidate from

$h : \{0,1\}^* \to \{0,1\}^*$ our , $h_n := h|_{\{0,1\}^n}$

$B(x)$: it hard bit

$b \in \{0,1\}^n$

- $\mu_b := [\, h_n(x) = b \longrightarrow B(x) = B(h^{-1}(b)) \,]$

Short

~~proof~~ :

π_b — prove h_n is 1-to-1,

— pick $a \in \{0,1\}^n$ s.t. $h(a) = b$,

— show $h_n(x) = b \to x = a$,

— show $B(a) = B(h^{-1}(b))$.

π_b : p-time but $K^t(\pi_b | \mu_b)$ is large for most b as

π_b contains a

Q : Can you make $ip(\mu_b)$ small in some p ?

YES NO

(candidate flas)

$$\gamma_b := [h_n(x) = b \longrightarrow \varphi_n(x)]$$

where $\varphi_n(x)$ is <u>knock</u> for P.

Same short proof: - using that h is 1-to-1
- substitute $a := h^{-1}(b)$
- evaluate $\varphi_n(a) = 1$

Q: An argument "specific" to b ?

Remark 1: random flas

Müller – Tzameret (using Feige-Kim-Ofek): random 3CNF with $\Omega(n^{1.4})$ clauses has poly-size TC^0-F proof.

showed that "short witnesses" can be found in exp-time

it is thus consistent that these formulas realize the separation of size from width

Remark 2 : proof complexity generators

$$g : \{0,1\}^n \longrightarrow \{0,1\}^m \quad , \quad m >> n$$

$$b \in \{0,1\}^m \setminus Rng(g)$$

$$\tau_b := \text{“} b \notin Rng(g) \text{”}$$

- $Kt(a)$ small for $a \in Rng(g)$, so slowly that $Kt(b)$ is large implies τ_b
- hence if τ_b ought to be hard to prove (the hardness of PV) else “$K(u) > n/2$” ought to be hard

i.e.

Remark 3 : further connections to

↳ implicit proof systems

↳ proof system w/ advice

[suggested by Igor Oliveira]

.
.
.