Propositional Proof Compression

computational complexity theoretical consequences

Edward Hermann Haeusler (PUC-Rio)
joint work with Lew Gordeev (Tuebingen Univ),
Robinson Callou (PUC-Rio) and José Flávio C. Barros
(PUC-Rio - UFC -Quixadá

Summer School in Mathematics - UnB

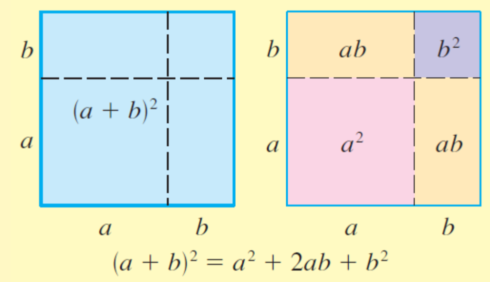
February 7, 2024

Arguments and Proofs

A proof plays two roles [Geuvers]

- A proof convinces the reader that the statement is correct.
- A proof explains why the statement is correct.

An informal proof in math



TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

A more formal proof in math

There are infinite prime numbers:

- Let P_i , $i = 1, n$, be prime numbers;
- Observe that $M = (P_1 \times P_2 \times \ldots \times P_n) + 1$ has remainder 1 when divided by any P_i, for $(\Pi_{j \neq i} P_j) P_i + 1 = M$;
- Either M is prime or there is a prime P different from any P_i, $i = 1, \ldots, n$;
- The set of primes is infinite.

The naive question

Are formal proofs for people or for machines?

Has not been adequately answered in any way.

Mechanical Theorem Proving for FOL

A Machine-Oriented Logic Based on the Resolution Principle.
J.A.Robinson (JACM, 1965)
Resolution + Unification ⇒ AI tasks base on Rule Systems

Variant Usage

SAT Solvers, Davis-Putnam (1960), DPLL or Davis-Putnam-Logemann-Loveland (1961).
Pressburger Arithmetic Solvers, Arithmetic with '+' only (1929).

Mathematical arguments are proofs ?

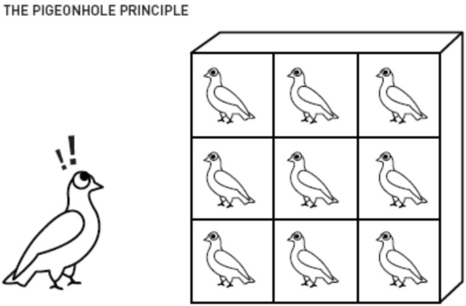
Proofs and Computers

ATP: proving of mathematical theorems by a computer program.

ITP: developing formal proofs by man-machine collaboration.

Different activities with different problems and specialized research groups

The pigeonhole principle: Logical or Mathematical ?



Theorem: If one assigns n pigeons to m pigeonholes and $n > m$ then there is at least one hole with more than one pigeon

Computational problems: It is easier to check that something is a solution than to find one?

$$PHP^n_m = \bigwedge_{i=1}^{n} \bigvee_{j=1}^{m} p_{ij} \supset \bigvee_{1 \leq j \leq k \leq n} \bigvee_{j=1}^{m} (p_ij \wedge p_{kj})$$

n	2^{2n}	Time to read the proof
5	1024	insignificant
10	1048576	1.048 sec
15	1 billion	17 min
20	1.099×10^{12}	30 hours
25	1.1252×10^{15}	35 years

Obs:Computer reads 10^6 characters per sec.

[Haken1985] Any refutation of $\neg PHP^n_m$ has size at least $2^{(n^2/2m)}$ clauses in resolution. There are at least 2^{2n} clauses in any refutation of $\neg PHP^{2n}_n$.

TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Is mathematics more objective than natural sciences?

Thierry Coquand, 2008

The history of mathematics has stories about false results that went undetected for long periods of time. However, it is generally believed that if a published mathematical argument is not valid, it will be eventually detected as such. While the process of finding a proof may require creative insight, the activity of checking a given mathematical argument is an objective activity; mathematical correctness should not be decided by a social process.

Dealing with huge proofs

Compression and efficient proof verification

Part of the computational complexity of theoremn proving and SAT is in the (Classical) Propositional Logic

Propositional proofs (I)

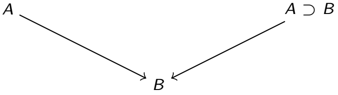
Natural Deduction

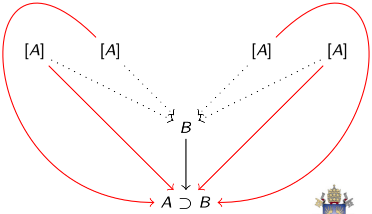
$$1\;\dfrac{\dfrac{\dfrac{[A]^1 \qquad A \supset B}{B} \qquad B \supset C}{C}}{A \supset C}$$

A convenient representation of $M_{\supset}$ in graphs

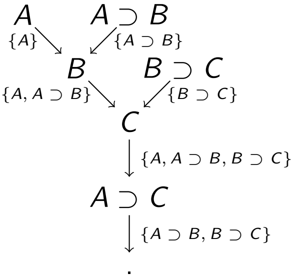
$$\frac{A \quad A \supset B}{B} \supset\text{-e}$$

$$\begin{array}{c} [A] \\ \vdots \\ \dfrac{B}{A \supset B} \end{array} \supset\text{-i}$$





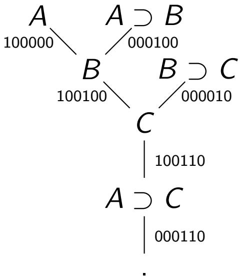
Using dependency sets to eliminate the need[1] for red (discharge) edges



[1] This approach has limitations

Using bitstrings to eliminate[2] the red (discharge) edges

Considering a total order on formulas (any)
$A \prec B \prec C \prec A \supset B \prec B \supset C \prec A \supset C$



Given the total order and the labeled tree, verifying that the conclusion is a $M_{\supset}$ tautology is polytime on the number of nodes in the tree.

[2] It has the same limitations of using dependency sets

Compressing proofs for easy proof-checking

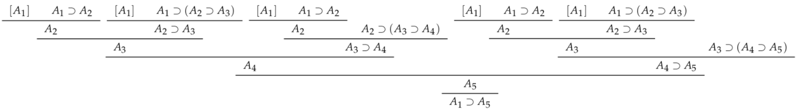


Figure 1: Deriving $A_1 \supset A_5$ from $A_1 \supset A_2$, $A_1 \supset (A_2 \supset A_3)$, $A_2 \supset (A_3 \supset A_4)$ and $A_3 \supset (A_4 \supset A_5)$

Compressing proofs for easy proof-checking

$A_1 \prec A_2 \prec A_3 \prec A_4 \prec A_5 \prec A_1 \supset A_2 \prec A_2 \supset A_3 \prec A_4 \supset A_5 \prec A_1 \supset A_5 \prec A_1 \supset (A_2 \supset A_3) \prec A_2 \supset (A_3 \supset A_4) \prec A_3 \supset (A_4 \supset A_5)$

With the sake of a better understanding, we remember that to any subset of formulas in the derivation there is a unique bitstring, see again definition 5. For example, the set $\{A_1, A_1 \supset A_2, A_1 \supset (A_2 \supset A_3)\}$ is represented by the bitstring 100001000100.

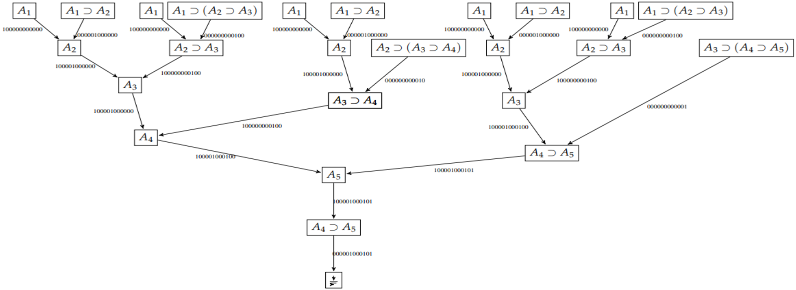


Figure 2: Labelled graph representation of derivation 1

Compressing a proof with easy proof-checking

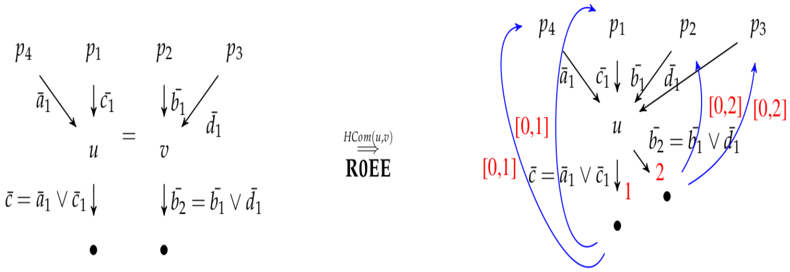


Figure 3: (a)u and v collapse (b) After collapse $HCom(u,v)$

Compressing proofs for easy proof-checking

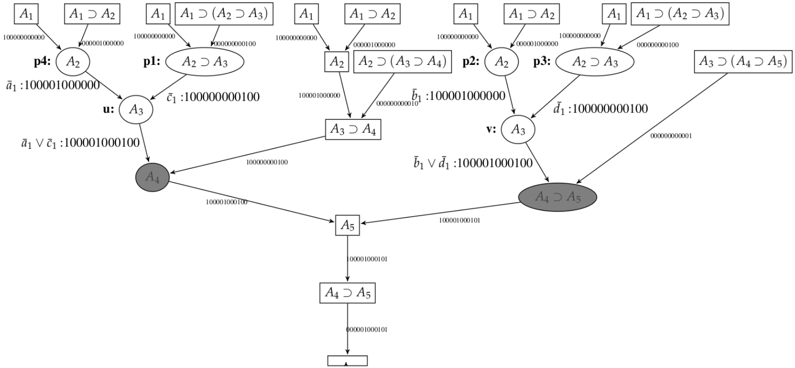


Figure 4: Matching of the HC-compression rule **R0EE** with the derivation-tree in figure 2

TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Compressing proofs for easy proof-checking

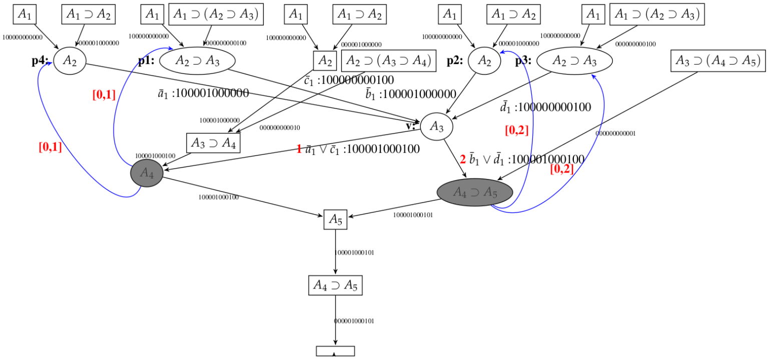


Figure 5: Result of the HC-compression rule **R0EE** application according the matching shown in figure 4

Compressing proofs for easy proof-checking

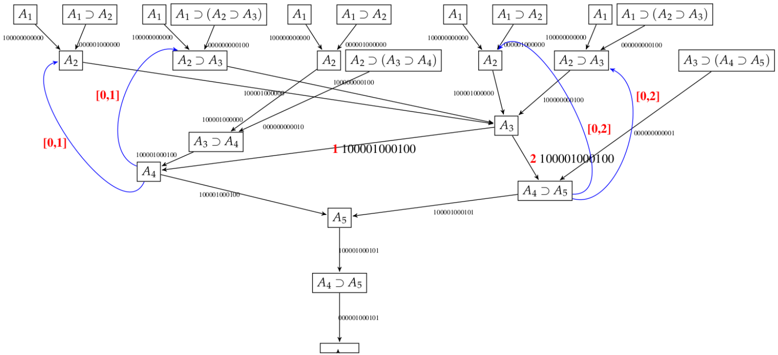


Figure 6: Defocused result of the HC-compression rule **R0EE** application appearing focused in figure 5

TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Compressing proofs for easy proof-checking

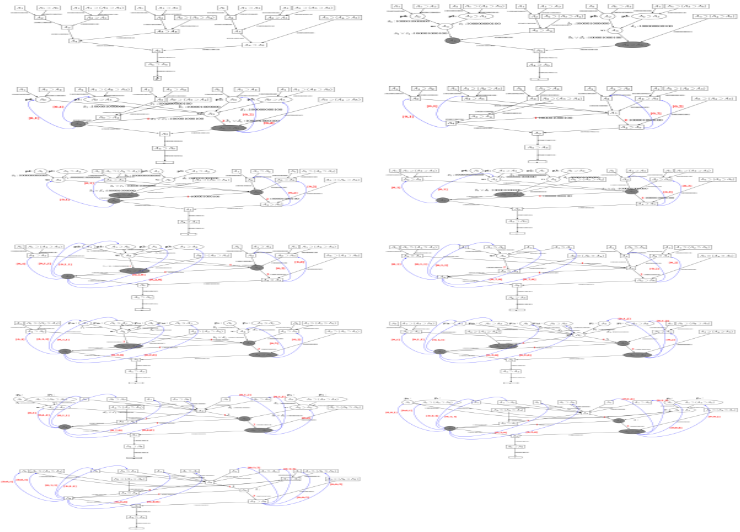


Figure 28: Summary of MUE applications to the initial derivation in [1]

Compressing a proof with easy proof-checking: MUE-rules

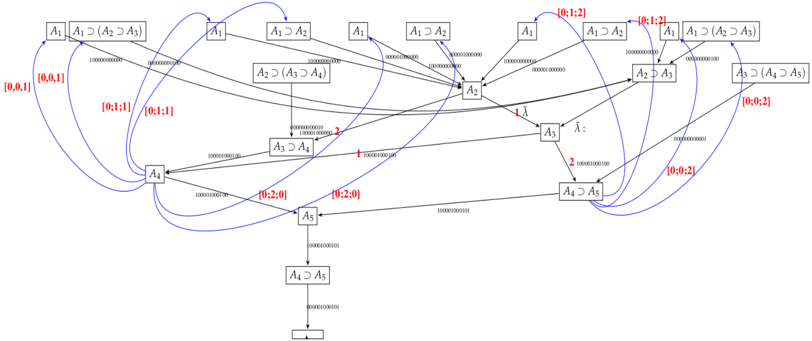


Figure 18: Defocused dag-like derivation after application of rule $\mathbf{R_e 2EE}$ as in figure 17

Compressing a proof with easy proof-checking: MDE-rules

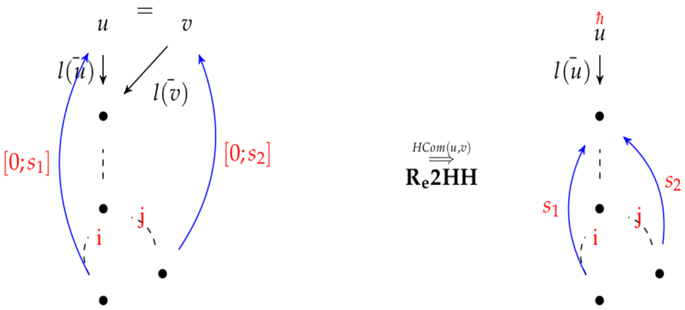


Figure 19: (a)u and v collapse (b) After collapse $HCom(u,v)$ 19

Compressing a proof with easy proof-checking: MDE-Rules

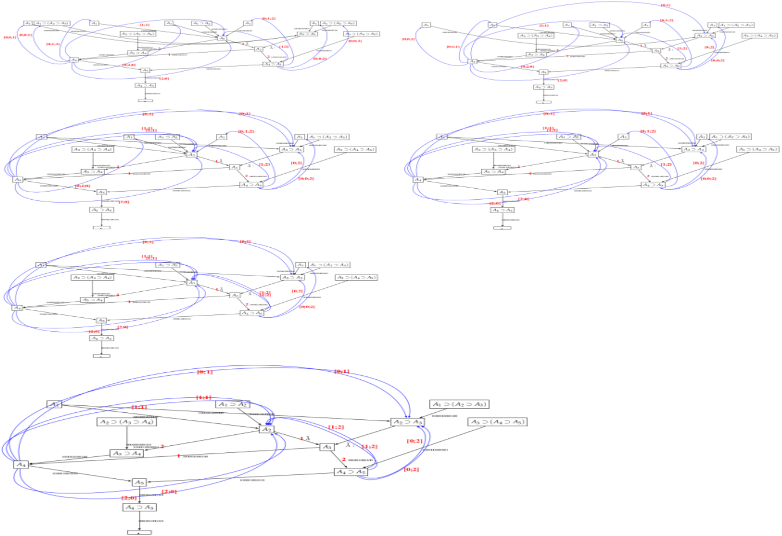


Figure 29: Summary of **MDE** applications to the initial derivation in [1]

An upper-bound for the size of the compressed proof

If h is the height of the initial derivation and m the number of formulas in it, then

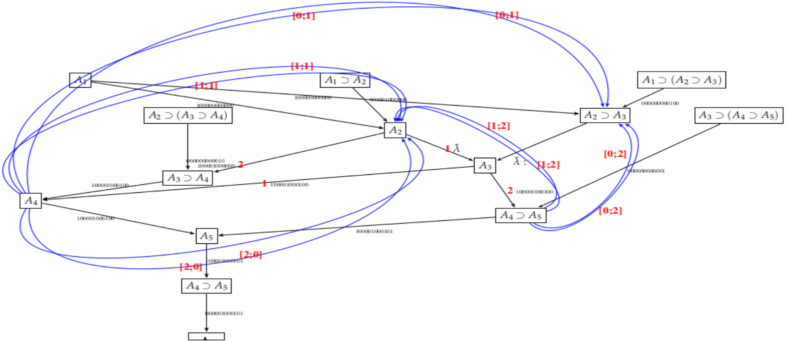
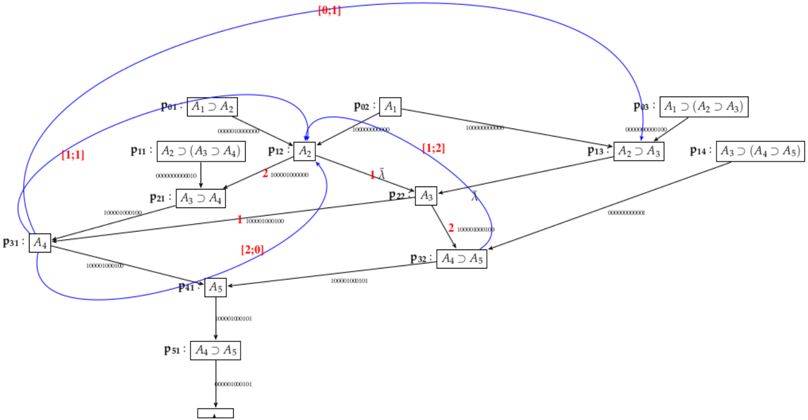


Figure 27: Defocused dag-like derivation after application of $\mathbf{R_e2HH}$ and threetimes $\mathbf{R_e2XH}$ to collapse all occurrences of A_1 in the dag-like in figure 26

The final size of the Dag-proof is $\mathcal{O}(h^2 \times m^4)$ upper-bounded.

Collapsing equal ancestor edges

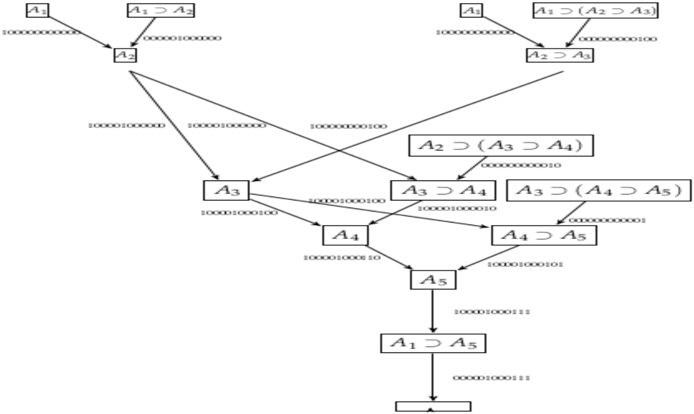


TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

The Patch Natural Deduction derivation



A glimpse into the technical details

Definition (Dag-like derivability structures DLDS)

Let Γ be a set of $M_{\supset}$ formulas and $\mathcal{O}_\Gamma$ an arbitrary linear ordering on Γ and $\mathcal{O}^0_\Gamma = \mathcal{O}_\Gamma \cup \{0, \lambda\}$[3]. A dag-like derivability structure, **DLDS for short**, is a tuple $\langle V, (E^i_D)_{i \in \mathcal{O}^i_\Gamma}, E_A, r, l, L, P\rangle$, where:

1. V is a non-empty set of nodes;
2. For each $i \in \mathcal{O}^0_\Gamma$, $E^i_D \subseteq V \times V$ is the family of sets of edges of deduction;
3. $E_A \subseteq V \times V$ is the set of edges of ancestrality;
4. $r \in V$ is the root of the **DLDS**;
5. $l : V \rightarrow \Gamma$ is a function, such that, for every $v \in V$, $l(v)$ is the (formula) label of v;
6. $L : \bigcup_{i \in \mathcal{O}^0_\Gamma} E^i_D \rightarrow \mathcal{B}(\mathcal{O}_S)$ is a function, such that, for every $\langle u, v\rangle \in E^i_D$, $L(\langle u, v\rangle)$ is a bitstring.
7. $P : E_A \rightarrow \{1, \ldots, || \Gamma ||\}^\star$, such that, for every $e \in E_A$, $P(e)$ is a string of the form $o_1; \ldots; o_n$, where each 0_i, $i = 1, n$ is an ordinal in $\mathcal{O}_\Gamma$;

TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

[3] $0 < n$, for every $n \in \mathcal{O}_\Gamma$

A glimpse into the technical details (cont)

Definition 23. *Given a structure $\mathcal{D} = \langle V, (E^i_D)_{i\in\mathcal{O}^i_\Gamma}, E_A, r, l, L, P\rangle$, we say that it is a valid* **DLDS**, *iff, the following conditions hold on it:*

- ***Color-Acyclicity*** *For each $i \in \mathcal{O}^i_\Gamma$, E^i_D does not have cycles;*
- ***Leveled-Colored*** *The rooted sub-dag $\langle V, (E^i_D)_{i\in\mathcal{O}^i_\Gamma}, r\rangle$ is leveled;*
- ***Ancestor-Edges*** *For each $\langle v_1, v_2\rangle \in E_A$, the level of v_1 is smaller than the level of v_2;*
- ***Ancestor-Backway-Information*** *For each $\langle v_1, v_2\rangle \in E_A$, $P(\langle v_1, v_2\rangle)$ is the relative address of v_1 from v_2;*
- ***Simplicity*** *The rooted sub-dag $\langle V, (E^i_D)_{i\in\mathcal{O}^i_\Gamma}, r\rangle$ is a simple graph, i.e, for each pair of nodes v_1 and v_2, there is at most an $i \in \mathcal{O}^i_\Gamma$, such that $\langle v_1, v_2\rangle \in E^i_D$;*
- ***Ancestor-Simplicity*** *The sub-dag $\langle V, E_A\rangle$ is a simple graph;*
- ***Non-Nested-Ancestor-Edges*** *For each $\langle v_1, v_2\rangle \in E_A$, there is no w in the path from v_2 to v_1, determined by $P(\langle u, v\rangle \in E_A)$, such that $\langle w, z\rangle \in E_A$, for some $z \in E_A$.*
- ***CorrectRuleApp*** *For each $w \in V$, $Flow(\mathcal{D}, w)(v)$ is well-defined for each $v \in Pre(w)$. Moreover, for each w and v, $Flow(\mathcal{D}, w)(v)$, with $v \in Pre(w)$, we have:*
  - *If $Flow(\mathcal{D}, w)(v) = \{(\vec{b}, p)\}$ then $OUT(v) = \{\langle v, v'\rangle\}$ and the color of $\langle v, v'\rangle$ is $head(p)$, i.e., $\langle v, v'\rangle \in E^{head(p)}_D$, and $\vec{b} = L(\langle v, v'\rangle)$, and;*
  - *If $Flow(\mathcal{D}, w)(v) \neq \emptyset$ and it is not a singleton either then for each $\Phi_i = \{(\vec{b}, p) \in Flow(\mathcal{D}, w)(v) : head(p) = i\}$:*
    1. *If $\Phi_i \neq \emptyset$ then there is only one v' $\langle v, v'\rangle \in E^i_D$ and if $\Phi_i = \{(\vec{b}, p)\}$ then $L(\langle v, v'\rangle) = \vec{b}$ else $L(\langle v, v'\rangle) = \lambda$, and;*
    2. *If $\Phi_i = \emptyset$ then there is no $v' \in V$, such that, $\langle v, v'\rangle \in E^i_D$*

TecMF

PUC RIO

Eberhard Karls Universität Tübingen

Main results

Theorem I: The 28 HC rules preserve the validity of the derivation. (Proved by R. Callou Filho using $L\exists\forall N$)

Theorem II: The 28 rules cover all possible cases of Horizontal Compression. (Proved by R. Callou Filho using $L\exists\forall N$)

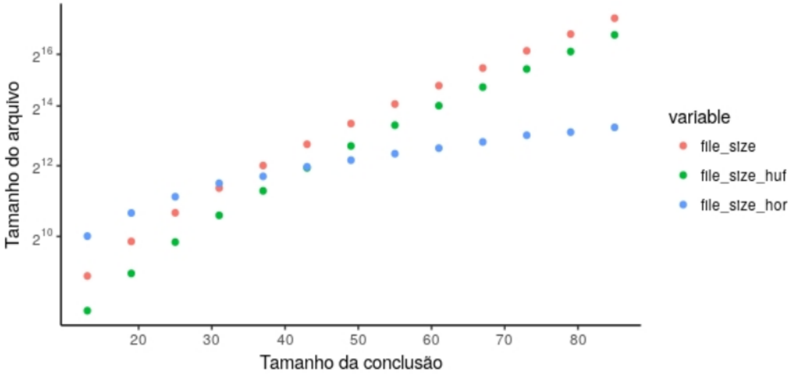
Proposition I: The 28 HC rules stops returning a fully compressed Dag-like derivation when applied on a tree-like valid derivation. (A mathematical consequence of Theorem II above and finite induction).

TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Empirical evaluation I



Empirical evaluation II

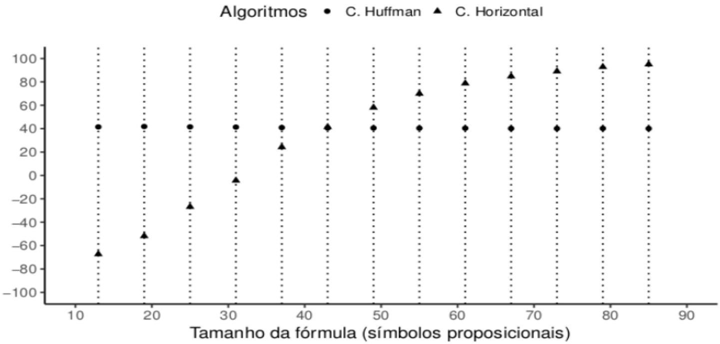
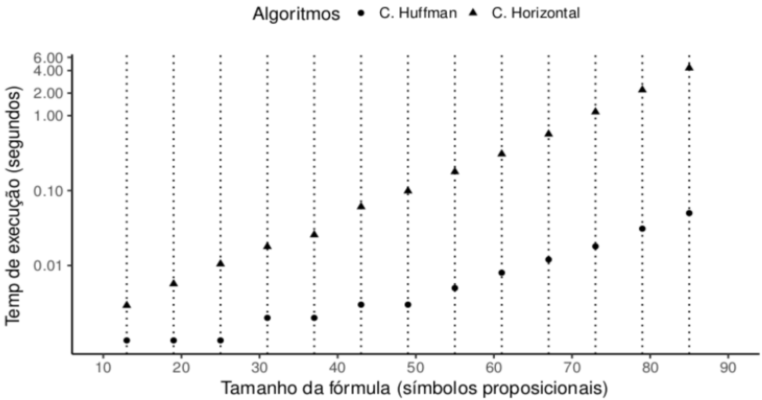


Figure 60: Compression rate comparison between Huffman and HC compression for big tautologies

Empirical evaluation III



Main Theoretical Result

The HC compression proves that $CoNP = NP$ (this presentation)

The HC compression proves that $PSPACE = NP$ (this presentation also, but need additional details)

Proof-theory can be used to prove that $CoNP = NP$

Theorem If Π is a normal proof of α then HC outputs a compressed $DLDS$ of size $\mathcal{O}(h^2.m^4)$.

Fact For any graph G with v nodes there is a formula α_G, of size $\mathcal{O}(v^3)$, that is SAT, iff, G is hamiltonian.

Proposition For any non-Hamiltonian graph G with v nodes there is a normal proof of $\neg\alpha_G$[4] with height $\mathcal{O}(v^2)$.

Theorem $CoNP \subset NP$, so $CoNP = NP$.

TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

[4] This the propositional formula that states that G is not Hamiltonian

A useful notation:

$$\begin{pmatrix} \alpha_n \\ \alpha_{n-1} \\ \vdots \\ \alpha_0 \end{pmatrix} \supset A \triangleq \alpha_n \supset (\alpha_{n-1} \ldots \supset (\alpha_0 \supset A) \ldots)$$

Moreover:

$$\begin{pmatrix} \alpha_n \\ \alpha_{n-1} \\ \vdots \\ \alpha_0 \end{pmatrix} \supset A$$

is the same of

$$\begin{pmatrix} \alpha_n \\ \alpha_{n-1} \\ \vdots \\ \alpha_1 \end{pmatrix} \supset (\alpha_0 \supset A)$$

Natural Deduction Proofs and Derivations: Usual Terminology [Prawitz1965]

Derivations and proofs are represented as labeled trees, the root is the conclusion and the leaves are assumptions , either closed or open.

Normal Proofs and Derivations in $M_{\supset}$ ND

A detour, or maximal formula, in a derivation Π is a formula occurrence μ that is, at the same time, conclusion of a $\supset$-I rule and major premiss of a $\supset$-E rule.

$$\dfrac{\dfrac{\pi_2}{A} \qquad 1\,\dfrac{\dfrac{\dfrac{[A]^1}{\pi_1}}{B}}{A \supset B}}{B}$$

$A \supset B$ is a maximal formula in the derivation above.

Maximal Formula or Detour

A branch in a derivation Π is any sequence $\alpha_0, \ldots, \alpha_i, \ldots, \alpha_k$ of formula occurrences in Π that starts in a top-formula α_0 ,ends in the conclusion of Π or some major premise of a ⊃-E rules application. Moreover α_i is a premise of the rule application that has α_{i+1} as conclusion, or vice-versa, for $i = 0, \ldots, k-1$.

Any branch has a formula μ that is the conclusion of an elimination rule, or it is an assumption, and premiss of an ⊃-introduction rule, or the last rule in the branch. μ is called *Minimal Formula*.

Normal Natural Deduction Proofs and Derivations: Usual Terminology [Prawitz1965]

A derivation π is normal, iff, it does not have any maximal formula.

Theorem Any derivation of α from $\Delta = \{\delta_0, \ldots, \delta_k\}$ gives rise to a normal derivation of α from $\Delta' \subseteq \Delta$

Apply the reduction below repeatedly.

$$\dfrac{\dfrac{\pi_2}{A} \qquad \dfrac{\begin{array}{c}[A]^1\\ \pi_1\\ B\end{array}}{A \supset B}}{B} \qquad \rhd \qquad \begin{array}{c}\pi_2\\ A\\ \pi_1\\ B\end{array}$$

Normal Atomically Expanded Proofs (NAEP)

A normal Natural Deduction derivation is Atomically Expanded, iff, all minimal formulas are atomic.

Theorem. If $\Delta \vdash \alpha$ then there is an atomically expanded derivation π of α from $\Delta' \subseteq \Delta$.

Proof. Apply the expansion below to each non-atomic minimal formula $A \supset B$, repeatedly until every minimal formula is atomic.

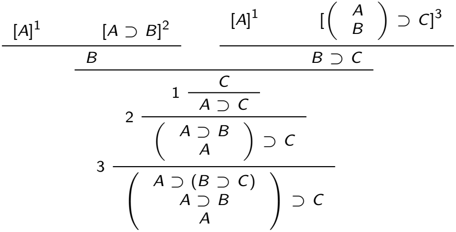
$$\dfrac{\overline{A \supset B}}{} \quad \rhd \quad a\,\dfrac{\dfrac{[A]^a \qquad \overline{A \supset B}}{B}}{A \supset B}$$

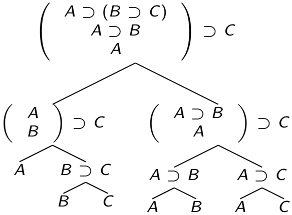
TecMF

PUC RIO

Eberhard Karls Universität Tübingen

$$
3\;\dfrac{2\;\dfrac{1\;\dfrac{\dfrac{\dfrac{[A]^1 \qquad [A \supset B]^2}{B} \qquad \dfrac{[A]^1 \qquad [A \supset (B \supset C)]^3}{B \supset C}}{C}}{A \supset C}}{(A \supset B) \supset (A \supset C)}}{(A \supset (B \supset C)) \supset (A \supset B) \supset (A \supset C)}
$$


[A]1
[A ⊃ B]2
[A]1
[(A B) ⊃ C]3
B
B ⊃ C
C
1
A ⊃ C
2
(A ⊃ B A) ⊃ C
3
(A ⊃ (B ⊃ C) A ⊃ B A) ⊃ C


(A ⊃ (B ⊃ C) A ⊃ B A) ⊃ C
(A B) ⊃ C
(A ⊃ B A) ⊃ C
A
B ⊃ C
A ⊃ B
A ⊃ C
B
C
A
B
A
C

Proposition. Let T_α be the **ST** of α and π an AENP of α. For each branch P in π:

- There is a maximal path σ in T_α *starting* in a leaf A and *finishing* at a r-child or the root of T_α, and, the reserve of σ (σ^R) is the l-part of P
- The E-part of P consists of the path from some even r-child of some formula in σ^R to its corresponding leftmost descendency, i.e. leaf.

Corollary. Given a NAEP π of α, a branch P in π, of height h and, any branch in the sub-derivation determined by P; its minimal formula is some leaf in the r-child descendency of P in T_α. Consequently there are at most $h.size(T_\alpha)$ sub-derivations determined by P.

Redundancy in Huge NAEPs

Fact. The number of leaf nodes in a binary tree is one more than the number of nodes with 2 children

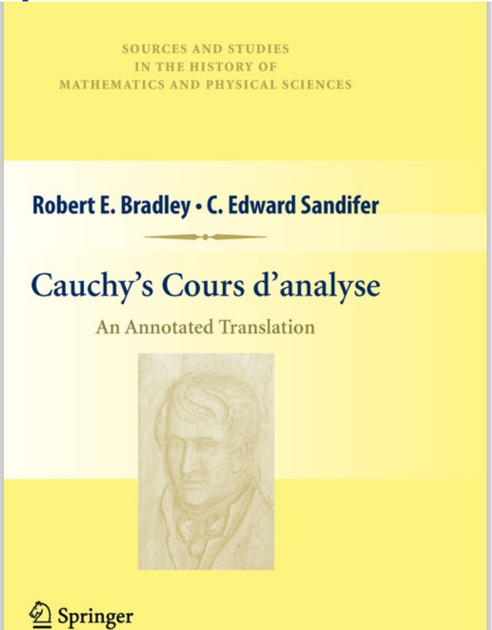
Proposition. If a NAEP π of α is such that $size(\pi) > a^{size(\alpha)}$ then there is a sub-derivation π' of π that repeats at least $a^{size(\alpha)}$ in π.

The HC compression method

Why the compression method is effective ? The most is the size of an exponetial proof the easiest is to compress it to a sub-exponential size, polynomial indeed.

Thank You

A very influential book on mathematics



An example of a mathematical argument, CoursD'Analyse, Cauchy, Section 6.2, 1821

Suppose the terms of series (1) involve some variable x. If the series is convergent and its various terms are continuous functions of x in a neighborhood of some particular value of this variable, then

$$s_n,\ r_n \text{ and } s$$

are also three functions of the variable x, the first of which is obviously continuous with repect to x in a neighborhood of the particular value in question. Given this, let us consider the increments in these three functions when we increase x by an infinitely small quantity α. For all possible values of n, the increment in s_n is an infinitely small quantity. The increment of r_n, as well as r_n itself, becomes infinitely small for very large values of n. Consequently, the increment in the function s must be infinitely small.[6] From this remark, we immediately deduce the following proposition:

Theorem I. — *When the various terms of series* (1) *are functions of the same variable x, continuous with respect to this variable in the neighborhood of a particular value for which the series converges, the sum s of the series is also a continuous function of x in the neighborhood of this particular value.*[7]

TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Cauchy's proof in modern math language

Continuous Function: f is continuous at x, iff,
$\forall\epsilon\exists\delta(\mid b \mid < \delta \supset \mid f(x+b) - f(x) \mid < \epsilon)$

($\star$) s_n is continuous at x :
$\exists\delta\forall b(\mid b \mid < \delta \supset \mid s_n(x+b) - s_n(x) \mid < \epsilon)$

($\diamond$) The series converges at x:
$\exists K \forall k > K(\mid r_n(x) \mid < \epsilon)$

($\Box$) The series converge at $x+b$:
$\exists K \forall k > K(\mid r_n(x+b) \mid < \epsilon)$

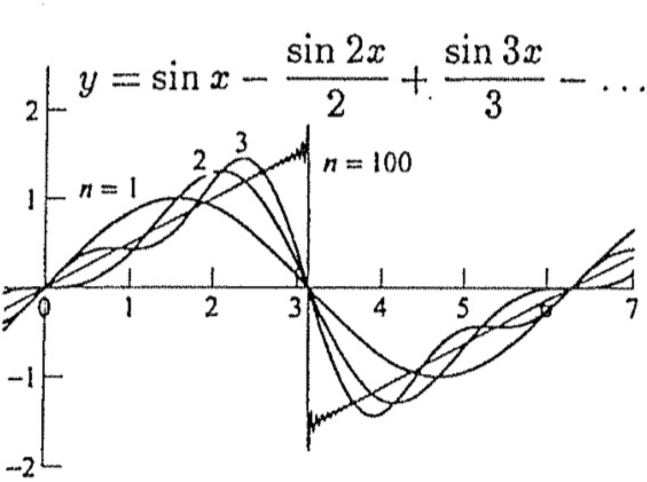
$\mid s(x+b) - s(x) \mid = \mid s_n(x+b) + r_n(x+b) - s_n(x) - r_n(x) \mid \leq \mid$
$s_n(x+b) - s_n(x) \mid + \mid r_n(x+b) \mid + \mid r_n(x) \mid \leq 3\epsilon$

TecMF

PUC RIO

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

An example of a mathematical argument



Maurice Lecat initiative in 1935

Harrison et ali. 2007

"Maurice Lecat published in 1935 a book with 130 pages of errors (500 approx) made by major mathematicians up to 1900". Nowadays would this initiative be possible ? The profusion of theorems is very higher than up to 1900.

Some remarkable achievements in ITP

- Proof of the prime number theorem (J.Avigad et. al., 2005) using Isabelle formalizes Selberg's proof. 30000 lines, 43 files.

$$\lim_{n\to\infty} \pi(n) = \frac{n}{ln(n)}$$

- Proof of the four colors theorem (G.Gonthier et al., 2008) using Coq. 60000 lines, 132 files.
- Proof of the Jordan curve theorem (Tom Hales, 2005) using HOL light. 75000 lines,15 files. Proved using Mizar later. The first correct proof is due to Veblen, 1904.

Main goals of an ITP

Andrea Asperti, 2010

The machine must be aware of the mathematical content (**the logic**) of expressions (passing from a machine readable to a machine understandable representation of mathematics).

Remarks on de Bruijn factor [see Freek Wiedijk & J. Harrison]

ITP tools or assistant proofs

- Automath [Eindhoven] (De Brujin)
- the HOL family [Cambridge] - deriving from LCF (R.Milner)
  - HOL4, HOL88 (M.Gordon), HOL90 (K.Slind)
  - HOL lite (J.Harrison)
  - Proof Power (ICL Ltd)
- Isabelle/Isar (L.Paulson,T.Nipkow) [Cambridge,Munich]
- NuPRL (Constable), MetaPearl [Cornelle]
- The COQ family
  - Coq (Huet,Coquand,Paulin-Mohring) [INRIA-France]
  - Agda (Coquand) [Chalmers]
  - Lego (Pollack) [Edinburgh]
  - Matita (Asperti,Sacerdoti Coen) [Bologna]
- PVS (N.Shankar) [Stanford]
- IMPS (W.Farmer) [McMaster]
- Mizar (A.Trybulec) [Bialystok]
- Lean (Leonardo Moura) [deMoura]

Famous uses of ITP in theorems relevant to CS

1. Needham-Schroeder authentication public key protocol breaking (1995 Lowe) and fixing (1996) correctness using CSP.
2. After Ariane V catastrophe (1996), Harrison proved (2006) that Ariane V catatrophe was caused by a programmer's diregarding the default exception-handling of IEEE 754 specs. He also proved correctness of IEEE 754 specs.
3. A proof attempt using Temporal Logic of the *ARPANET TCP three-way hand-shake protocol* revealed a very unliked but severe bug, afterwards corrected in Internet TCP/IP (1982).

There is nothing wrong with Jordan's proof o Jordan curve theorem

Studies in Logic, Grammar and Rhetoric, Tom Hales, 2007

My initial purpose in reading Jordan was to locate the error. I had completed a formal proof of the Jordan curve theorem in January 2005 and wanted to mention Jordan's error in the introduction to that paper [3]. In view of the heavy criticism of Jordan's proof, I was surprised when I sat down to read his proof to find nothing objectionable about it. Since then, I have contacted a number of the authors who have criticized Jordan, and each case the author has admitted to having no direct knowledge of an error in Jordan's proof. It seems that there is no one still alive with a direct knowledge of the error.