Proof Compression and \mathcal{NP} vs \mathcal{PSPACE}

L. Gordeev

Uni-Tübingen, Uni-Ghent, PUC-Rio

CL2016, Hamburg, September 10-11, 2016

§1. Reminder -1-

・ロン ・部 と ・ ヨ と ・ ヨ と …

æ

Classes $\mathcal{NP}\text{, } \textbf{co}\mathcal{NP}\text{ and } \mathcal{PSPACE}\text{:}$

▲圖 ▶ ▲ 臣 ▶ ▲ 臣 ▶ …

æ

Classes \mathcal{NP} , co \mathcal{NP} and \mathcal{PSPACE} :

L ⊆ {0,1}* is in NP, resp. coNP, if there exists a polynomial p and a polytime TM M such that

$$x \in L \Leftrightarrow \left(\exists u \in \{0,1\}^{p(|x|)}\right) M(x,u) = 1$$

resp.
$$x \in L \Leftrightarrow \left(\forall u \in \{0,1\}^{p(|x|)}\right) M(x,u) = 1$$

holds for every $x \in \{0,1\}^*$.

Classes \mathcal{NP} , co \mathcal{NP} and \mathcal{PSPACE} :

• $L \subseteq \{0,1\}^*$ is in \mathcal{NP} , resp. $co\mathcal{NP}$, if there exists a polynomial p and a polytime TM M such that

resp.
$$x \in L \Leftrightarrow \left(\exists u \in \{0,1\}^{p(|x|)}\right) M(x,u) = 1$$
$$x \in L \Leftrightarrow \left(\forall u \in \{0,1\}^{p(|x|)}\right) M(x,u) = 1$$

holds for every $x \in \{0,1\}^*$.

 L ⊆ {0,1}* is in PSPACE if there exists a polynomial p and a TM M such that for every input x ∈ {0,1}*, the total number of non-blank locations that occur during M's execution on x is at most p(|x|), while x ∈ L ⇔ M(x) = 1.

§1. Reminder -2-

・ロン ・部 と ・ ヨ と ・ ヨ と …

æ

æ

御 と く ヨ と く ヨ と …

• $\mathcal{NP} \subseteq \mathcal{PSPACE}$ and $co\mathcal{NP} \subseteq \mathcal{PSPACE}$.

伺 と く ヨ と く ヨ と

э

- $\mathcal{NP} \subseteq \mathcal{PSPACE}$ and $co\mathcal{NP} \subseteq \mathcal{PSPACE}$.
- NP = PSPACE implies NP = coNP. The latter conjecture seems more natural and/or plausible, as it reflects an idea of logical equivalence between model theoretical (re: NP) and proof theoretical (re: coNP) interpretations of non-deterministic polytime computability.

- $\mathcal{NP} \subseteq \mathcal{PSPACE}$ and $co\mathcal{NP} \subseteq \mathcal{PSPACE}$.
- NP = PSPACE implies NP = coNP. The latter conjecture seems more natural and/or plausible, as it reflects an idea of logical equivalence between model theoretical (re: NP) and proof theoretical (re: coNP) interpretations of non-deterministic polytime computability.
- $\mathcal{NP} = co\mathcal{NP}$ (resp. $\mathcal{NP} = \mathcal{PSPACE}$) follows from global polynomial-size provability of tautologies in classical and/or intuitionistic (resp. minimal) logic.

- $\mathcal{NP} \subseteq \mathcal{PSPACE}$ and $co\mathcal{NP} \subseteq \mathcal{PSPACE}$.
- NP = PSPACE implies NP = coNP. The latter conjecture seems more natural and/or plausible, as it reflects an idea of logical equivalence between model theoretical (re: NP) and proof theoretical (re: coNP) interpretations of non-deterministic polytime computability.
- $\mathcal{NP} = co\mathcal{NP}$ (resp. $\mathcal{NP} = \mathcal{PSPACE}$) follows from global polynomial-size provability of tautologies in classical and/or intuitionistic (resp. minimal) logic.
- Claim [L.G.+E.H.Haeusler]: $\mathcal{NP} = \mathcal{PSPACE}$ is provable by DAG-like proof-compression techniques in Prawitz's Natural Deduction for minimal logic.

伺 ト く ヨ ト く ヨ ト

æ

御 と く ヨ と く ヨ と

§2. The proof [L.G.]: Overview

- Formalize minimal propositional logic as fragment LM_{\rightarrow} of Hudelmaier's tree-like cutfree intuitionistic sequent calculus. For any LM_{\rightarrow} proof ∂ of sequent $\Rightarrow \rho$:
 - $h(\partial)$ (= the height) is polynomial (actually linear) in $|\rho|$,
 - φ(∂) (= total number of formulas) and μ(∂) (= maximal formula length) are also polynomial in |ρ|.
- Show that there exists a constructive (1)+(2) preserving embedding F of LM→ into Prawitz's tree-like natural deduction formalism NM→ for minimal logic.
- $\textcircled{\sc 0}$ Elaborate polytime verifiable DAG-like deducibility in $NM_{\rightarrow}.$
- Elaborate and apply *horizontal tree-to-DAG proof compression* in NM→. For any tree-like NM→ input ∂, the weight of DAG-like output ∂^C is bounded by h(∂) × φ(∂) × μ(∂). Hence the weight of (F(∂))^C for any given tree-like LM→ proof ∂ of ρ is polynomially bounded in |ρ|. Since minimal logic is PSPACE-complete, conclude that NP = PSPACE.

§2. More on conclusion $\mathcal{NP} = \mathcal{PSPACE}$

L. Gordeev Proof Compression and NP vs PSPACE

§2. More on conclusion $\mathcal{NP} = \mathcal{PSPACE}$

Proof.

L. Gordeev Proof Compression and NP vs PSPACE

æ

イロン 不聞 とくほとう ほどう

Proof.

Recall that the validity problem for minimal propositional logic is PSPACE-complete. It will suffice to show that it is a NP problem. So consider any purely implicational formula ρ . By Hudelmaier's result, ρ is valid in the minimal logic iff there exists a tree-like LM_{\rightarrow} proof ∂ of ρ . Hence, by the embedding theorem and soundness and completeness of DAG-like NM_{\rightarrow} , ρ is valid in the minimal logic iff we can "guess" a DAG-like NM_{\rightarrow} proof $\hat{\partial}$ of ρ , whose weight is polynomial in $|\rho|$ (witness: $(F(\partial))^{C}$). Moreover, we know that ' ∂ is an encoded DAG-like NM_{\rightarrow} proof of ρ ' is decidable in polynomial time with respect to $|\rho|$. Thus the existence of DAG-like NM_{\rightarrow} proof of ρ is verifiable in polynomial time by a non-deterministic algorithm, and hence so is the problem of ρ validity in the minimal logic, Q.E.D.

- 冬 医 🕨 🔍 医 🕨 👘

3

First question: Which proof system(s) could/should be used?

▶ ★ Ξ ▶

3

First question: Which proof system(s) could/should be used? **Basic options:**

- Łukasiewicz-Tarski-Hilbert-Bernays style modus ponens calculi a/o Gentzen-Schütte-style sequent calculi with cut.
- 2 Cutfree sequent calculi.
- S Prawitz style *natural deduction* (not necessarily normal).

First question: Which proof system(s) could/should be used? **Basic options:**

- Łukasiewicz-Tarski-Hilbert-Bernays style modus ponens calculi a/o Gentzen-Schütte-style sequent calculi with cut.
- 2 Cutfree sequent calculi.
- S Prawitz style *natural deduction* (not necessarily normal).

First answer: 2 and 3 interactive (1 is too loose).

First question: Which proof system(s) could/should be used? **Basic options:**

- Łukasiewicz-Tarski-Hilbert-Bernays style modus ponens calculi a/o Gentzen-Schütte-style sequent calculi with cut.
- 2 Cutfree sequent calculi.
- S Prawitz style *natural deduction* (not necessarily normal).

First answer: 2 and 3 interactive (1 is too loose). **Second question:** What about geometric structure of proofs involved?

First question: Which proof system(s) could/should be used? **Basic options:**

- Łukasiewicz-Tarski-Hilbert-Bernays style modus ponens calculi a/o Gentzen-Schütte-style sequent calculi with cut.
- 2 Cutfree sequent calculi.
- S Prawitz style *natural deduction* (not necessarily normal).

First answer: 2 and 3 interactive (1 is too loose).

Second question: What about geometric structure of proofs involved?

Basic options:

- Standard tree-like proofs.
- **②** DAG-like proofs (DAG = *directed acyclic graph*).

First question: Which proof system(s) could/should be used? **Basic options:**

- Łukasiewicz-Tarski-Hilbert-Bernays style modus ponens calculi a/o Gentzen-Schütte-style sequent calculi with cut.
- 2 Cutfree sequent calculi.
- S Prawitz style *natural deduction* (not necessarily normal).

First answer: 2 and 3 interactive (1 is too loose).

Second question: What about geometric structure of proofs involved?

Basic options:

- Standard tree-like proofs.
- OAG-like proofs (DAG = directed acyclic graph).

Second answer: 1 for 2 and 2 for 3.

▶ ★ 문 ▶ ★ 문 ▶

æ

 DAGs in question arise by proof compression that is obtained by merging distinct nodes labeled with identical proof objects (formulas or sequents) in tree-like inputs.

- DAGs in question arise by proof compression that is obtained by merging distinct nodes labeled with identical proof objects (formulas or sequents) in tree-like inputs.
- Proofs in (cutfree, et al) sequent calculi admit full DAG-like compression in which distnict nodes contain distint sequents, which makes the size bounded by total number of sequents.

- DAGs in question arise by proof compression that is obtained by merging distinct nodes labeled with identical proof objects (formulas or sequents) in tree-like inputs.
- Proofs in (cutfree, et al) sequent calculi admit full DAG-like compression in which distnict nodes contain distint sequents, which makes the size bounded by total number of sequents.
- This is good, but not good enough. Because polynomial bounds on the number of subformulas fail to provide polynomial bounds on the number of sequents involved. In contrast, proof objects of natural deductions are single formulas, so there is a hope to overcome this obstacle.

- DAGs in question arise by proof compression that is obtained by merging distinct nodes labeled with identical proof objects (formulas or sequents) in tree-like inputs.
- Proofs in (cutfree, et al) sequent calculi admit full DAG-like compression in which distnict nodes contain distint sequents, which makes the size bounded by total number of sequents.
- This is good, but not good enough. Because polynomial bounds on the number of subformulas fail to provide polynomial bounds on the number of sequents involved. In contrast, proof objects of natural deductions are single formulas, so there is a hope to overcome this obstacle.

However, full compression of natural deductions should be weakened to (say) *horizontal compression*, to save Prawitz's discharging rule(s). But this weakening still yields the result, provided that the height and the total number of formulas are polynomially bounded (via emebedding of sequent proofs).

L. Gordeev Proof Compression and NP vs PSPACE

→ 문 ▶ 문

Any given tree-like deduction ∂ with root formula ρ can be compressed to a DAG-like deduction ∂^c of the same conclusion ρ such that the size of ∂^c is at most h(∂) × φ(∂).

- Any given tree-like deduction ∂ with root formula ρ can be compressed to a DAG-like deduction ∂^c of the same conclusion ρ such that the size of ∂^c is at most h(∂) × φ(∂).
- The operation ∂ → ∂^c (called *horizontal compression*) runs by bottom-up recursion on h(∂) such that for any n ≤ h(∂), the nth horizontal section of ∂^c is obtained by merging all nodes with identical formulas occurring in the nth horizontal section of ∂ (this operation is called *horizontal collapsing*).

- Any given tree-like deduction ∂ with root formula ρ can be compressed to a DAG-like deduction ∂^c of the same conclusion ρ such that the size of ∂^c is at most h(∂) × φ(∂).
- The operation ∂ → ∂^C (called *horizontal compression*) runs by bottom-up recursion on h(∂) such that for any n ≤ h(∂), the nth horizontal section of ∂^C is obtained by merging all nodes with identical formulas occurring in the nth horizontal section of ∂ (this operation is called *horizontal collapsing*).
- Thus the horizontal compression is obtained by bottom-up iteration of the horizontal collapsing.

- Any given tree-like deduction ∂ with root formula ρ can be compressed to a DAG-like deduction ∂^c of the same conclusion ρ such that the size of ∂^c is at most h(∂) × φ(∂).
- The operation ∂ → ∂^c (called *horizontal compression*) runs by bottom-up recursion on h(∂) such that for any n ≤ h(∂), the nth horizontal section of ∂^c is obtained by merging all nodes with identical formulas occurring in the nth horizontal section of ∂ (this operation is called *horizontal collapsing*).
- Thus the horizontal compression is obtained by bottom-up iteration of the horizontal collapsing.
- The size and weigth estimates |∂^C| ≤ h(∂) × φ(∂) resp.
 ||∂^C|| ≤ h(∂) × φ(∂) × μ(∂) are obvious, as the size of every (compressed) nth horizontal section of ∂^C can't exceed φ(∂).

§3. Background: Crucial technical features -1-

L. Gordeev Proof Compression and NP vs PSPACE

留 と く ヨ と く ヨ と …

æ

§3. Background: Crucial technical features -1-

 The notion of DAG-like deducibility/provability is highly untrivial due to the corresponding DAG-like discharging (of chosen assumptions α). For in a given DAG-like deduction ∂ there are different maximal threads connecting α with root formula ρ. This is due to inverse-branching nodes (which don't occur in tree-like deductions).

§3. Background: Crucial technical features -1-

- The notion of DAG-like deducibility/provability is highly untrivial due to the corresponding DAG-like discharging (of chosen assumptions α). For in a given DAG-like deduction ∂ there are different maximal threads connecting α with root formula ρ. This is due to inverse-branching nodes (which don't occur in tree-like deductions).
- Thus every DAG-like deduction requires additional information on "legitimate" maximal deduction threads that determine the sets of open/closed assumptions. This is achieved by adding a suitable function ℓ^{G} that determines "legitimate" parents of inverse-branching nodes (regarded as "road signs" showing allowed ways from the leaves down to the root).

- The notion of DAG-like deducibility/provability is highly untrivial due to the corresponding DAG-like discharging (of chosen assumptions α). For in a given DAG-like deduction ∂ there are different maximal threads connecting α with root formula ρ. This is due to inverse-branching nodes (which don't occur in tree-like deductions).
- Thus every DAG-like deduction requires additional information on "legitimate" maximal deduction threads that determine the sets of open/closed assumptions. This is achieved by adding a suitable function $\ell^{\rm G}$ that determines "legitimate" parents of inverse-branching nodes (regarded as "road signs" showing allowed ways from the leaves down to the root).
- A given DAG-like deduction with root formula ρ is called a proof of ρ iff all assumptions are closed.

留 と く ヨ と く ヨ と …

æ

• Corresponding "legitimate" DAG-like provability remains sound and complete w.r.t. minimal logic (just as its canonical tree-like version).

- Corresponding "legitimate" DAG-like provability remains sound and complete w.r.t. minimal logic (just as its canonical tree-like version).
- Horizontal compression is supplied with corresponding $\ell^{\rm G}$ compression that preserves closed assumptions (and hence provability). So if ∂ is a canonical tree-like proof of ρ then $\partial^{\rm C}$ is a DAG-like proof of ρ .

- Corresponding "legitimate" DAG-like provability remains sound and complete w.r.t. minimal logic (just as its canonical tree-like version).
- Horizontal compression is supplied with corresponding $\ell^{\rm G}$ compression that preserves closed assumptions (and hence provability). So if ∂ is a canonical tree-like proof of ρ then $\partial^{\rm C}$ is a DAG-like proof of ρ .
- DAG-like provability in question is encoded by appropriate *local proof correctness* conditions that are polytime verifiable (just as in standard tree-like case).

§4. Hudelmaier's sequent calculus LM_{\rightarrow} -1-

L. Gordeev Proof Compression and \mathcal{NP} vs \mathcal{PSPACE}

▶ 《 문 ▶ 《 문 ▶ …

æ

$\S4$. Hudelmaier's sequent calculus $LM_{ ightarrow}$ -1-

Axiom and rules of implicational minimal logic:

$$(MA): \quad \Gamma, p \Longrightarrow p$$

$$(MA): \quad \Gamma, p \Longrightarrow p$$

$$(M/1 \to): \quad \frac{\Gamma, \alpha \Longrightarrow \beta}{\Gamma \Longrightarrow \alpha \to \beta} \quad [(\nexists\gamma): (\alpha \to \beta) \to \gamma \in \Gamma]$$

$$(M/2 \to): \quad \frac{\Gamma, \alpha, \beta \to \gamma \Longrightarrow \beta}{\Gamma, (\alpha \to \beta) \to \gamma \Longrightarrow \alpha \to \beta}$$

$$(ME \to P): \frac{\Gamma, p, \gamma \Longrightarrow q}{\Gamma, p, p \to \gamma \Longrightarrow q} \quad [q \in VAR(\Gamma, \gamma), p \neq q]$$

$$(ME \to \to): \frac{\Gamma, \alpha, \beta \to \gamma \Longrightarrow \beta}{\Gamma, (\alpha \to \beta) \to \gamma \Longrightarrow q} \quad [q \in VAR(\Gamma, \gamma)]$$

- * 注 * * 注 * -

э

§4. Hudelmaier's sequent calculus LM_{\rightarrow} -2-

L. Gordeev Proof Compression and NP vs PSPACE

御 と く き と く き と …

æ

§4. Hudelmaier's sequent calculus LM_{\rightarrow} -2-

Theorem (Hudelmaier + L.G.)

э

Theorem (Hudelmaier + L.G.)

 LM_{\rightarrow} is sound and complete with respect to minimal propositional logic and tree-like deducibility. So any given formula ρ is valid in the minimal logic iff sequent $\Longrightarrow \rho$ is tree-like deducible in LM_{\rightarrow} . Moreover:

Theorem (Hudelmaier + L.G.)

 LM_{\rightarrow} is sound and complete with respect to minimal propositional logic and tree-like deducibility. So any given formula ρ is valid in the minimal logic iff sequent $\Longrightarrow \rho$ is tree-like deducible in LM_{\rightarrow} . Moreover:

The height of any tree-like LM_→ deduction ∂ of sequent S is linear in |S|. In particular if S is ⇒ ρ, then h(∂) ≤ 3 |ρ|.

Theorem (Hudelmaier + L.G.)

 LM_{\rightarrow} is sound and complete with respect to minimal propositional logic and tree-like deducibility. So any given formula ρ is valid in the minimal logic iff sequent $\Longrightarrow \rho$ is tree-like deducible in LM_{\rightarrow} . Moreover:

- The height of any tree-like LM→ deduction ∂ of sequent S is linear in |S|. In particular if S is ⇒ ρ, then h(∂) ≤ 3 |ρ|.
- The foundation of any tree-like LM→ deduction ∂ of sequent S is at most quadratic in |S|. In particular if S is ⇒ ρ, then φ(∂) ≤ (|ρ| + 1)², while |α| ≤ |ρ| for any α occurring in ∂.

L. Gordeev Proof Compression and NP vs PSPACE

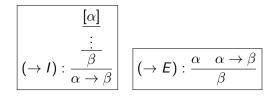
御 と く き と く き と …

æ

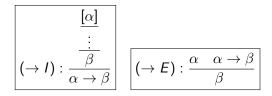
• We consider Prawitz's purely implicational proof system NM_{\rightarrow} for minimal propositional logic that contains just two rules

글 🕨 🔺 글 🕨 👘

• We consider Prawitz's purely implicational proof system NM_{\rightarrow} for minimal propositional logic that contains just two rules

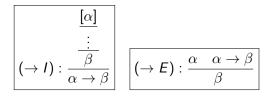


 $\bullet~$ We consider Prawitz's purely implicational proof system NM_{\rightarrow} for minimal propositional logic that contains just two rules



where $\alpha, \beta, \gamma, \cdots$ denote arbitrary formulas over propositionl variables p, q, r, \cdots and one propositional connective \rightarrow .

• We consider Prawitz's purely implicational proof system NM_{\rightarrow} for minimal propositional logic that contains just two rules



where $\alpha, \beta, \gamma, \cdots$ denote arbitrary formulas over propositionl variables p, q, r, \cdots and one propositional connective \rightarrow .



• We consider Prawitz's purely implicational proof system NM_{\rightarrow} for minimal propositional logic that contains just two rules

$$(\rightarrow I): \frac{\frac{[\alpha]}{\vdots}}{\alpha \rightarrow \beta} \quad (\rightarrow E): \frac{\alpha \quad \alpha \rightarrow \beta}{\beta}$$

where $\alpha, \beta, \gamma, \cdots$ denote arbitrary formulas over propositionl variables p, q, r, \cdots and one propositional connective \rightarrow .

Theorem (Prawitz)

 NM_{\rightarrow} is sound and complete with respect to minimal propositional logic and tree-like deducibility.

§6. Embedding theorem

L. Gordeev Proof Compression and NP vs PSPACE

æ

< E

§6. Embedding theorem

Theorem (L.G.)

L. Gordeev Proof Compression and \mathcal{NP} vs \mathcal{PSPACE}

æ

▶ 《 문 ▶ 《 문 ▶

There exists a recursive operator F that transforms any given tree-like LM_{\rightarrow} deduction ∂ of $\Gamma \implies \rho$ into a tree-like NM_{\rightarrow} deduction F (∂) with root-formula ρ and assumptions occurring in Γ . Moreover ∂ and F (∂) share the semi-subformula property, linearity of the height and polynomial upper bounds on the foundation. In particular if $\Gamma = \emptyset$, then F (∂) is a NM_{\rightarrow} proof of ρ such that:

There exists a recursive operator F that transforms any given tree-like LM_{\rightarrow} deduction ∂ of $\Gamma \Longrightarrow \rho$ into a tree-like NM_{\rightarrow} deduction F (∂) with root-formula ρ and assumptions occurring in Γ . Moreover ∂ and F (∂) share the semi-subformula property, linearity of the height and polynomial upper bounds on the foundation. In particular if $\Gamma = \emptyset$, then F (∂) is a NM_{\rightarrow} proof of ρ such that:

1 $h(F(\partial)) \le 18 |\rho|,$

There exists a recursive operator F that transforms any given tree-like LM_{\rightarrow} deduction ∂ of $\Gamma \Longrightarrow \rho$ into a tree-like NM_{\rightarrow} deduction F (∂) with root-formula ρ and assumptions occurring in Γ . Moreover ∂ and F (∂) share the semi-subformula property, linearity of the height and polynomial upper bounds on the foundation. In particular if $\Gamma = \emptyset$, then F (∂) is a NM_{\rightarrow} proof of ρ such that:

- **1** $h(F(\partial)) \le 18 |\rho|,$
- **2** $\phi(F(\partial)) < (|\rho|+1)^2(|\rho|+2),$

There exists a recursive operator F that transforms any given tree-like LM_{\rightarrow} deduction ∂ of $\Gamma \Longrightarrow \rho$ into a tree-like NM_{\rightarrow} deduction F (∂) with root-formula ρ and assumptions occurring in Γ . Moreover ∂ and F (∂) share the semi-subformula property, linearity of the height and polynomial upper bounds on the foundation. In particular if $\Gamma = \emptyset$, then F (∂) is a NM_{\rightarrow} proof of ρ such that:

- **1** $h(F(\partial)) \le 18 |\rho|,$
- **2** $\phi(F(\partial)) < (|\rho|+1)^2(|\rho|+2),$
- $(F(\partial)) \leq 2 |\rho|.$