On the computational content of intuitionistic propositional proofs

Samuel R. Buss[1,3] Pavel Pudlák[2,3]

1 Introduction

The intuitionistic calculus was introduced to capture reasoning in constructive mathematics. As such it has much more constructive character than classical logic. This property of the intuitionistic calculus has been extensively studied, but mostly from the point of view of computability and little has been proved about computational complexity. The aim of this paper is to show that the constructive character of intuitionistic logic manifests itself not only on the level of computability but, in case of the propositional fragment, also on the level of *polynomial time* computability.

Recent progress in proof complexity of propositional logic, which concerns various proof systems, suggest that the study of the complexity of intuitionistic propositional proofs may be a fruitful area. In particular for several classical calculi a so-called feasible interpolation theorem was proved [5, 7, 9]. Such theorems enable one to extract a boolean circuit from a proof; the size of the circuit is polynomial in the size of the proof. Indeed, feasible interpolation theorem was proved for the intuitionistic sequent calculus in [8]. The proof was based on the result of Buss and Mints [3] which shows that the well-known disjunction property can be witnessed by polynomial algorithms in case of the propositional fragment of the intuitionistic calculus.

In this paper we further generalize the two results on the intuitionistic propositional calculus. The ultimate aim is to obtain a realizability theorem for intuitionistic propositional proofs based on polynomial time computations. We prove a result in this direction (Theorem 3), but we suspect that it is not the best possible result of this type. On the other hand, we show that boolean circuits cannot be replaced by a more restricted type of computation (section 5).

Our proof technique is extracted from [3]. In this paper we make it more explicit (Theorem 1) and use the sequent calculus instead of the natural deduction system used in [3]. Goerdt [4] has also proved some related extensions of

[1]Supported in part by NSF grant DMS-9803515.
[2]Supported in part by a grant A1019901 of the Academy of Sciences of the Czech Republic.
[3]Supported in part by a cooperative research grant INT-9600919/ME-103 from the NSF (USA) and the MŠMT (Czech Republic)

the results in [3], but his main result is weaker than our Theorem 1.

In section 6 we prove some corresponding results for first order intuitionistic logic.

2 Cut elimination

We are working exclusively with propositional logic for now. The sequent calculus for intuitionistic logic is formulated in the usual way. Each sequent has at most one formula in the succedent. We shall adopt the convention that the antecedents of sequents are multisets of formulas, and freely use notions like "ancestor" and "direct ancestor" which can readily be defined similarly to the definitions in [2]. Our propositional language contains the logical symbols $\wedge$, $\vee$, $\supset$ and $\bot$. A negation $\neg A$ is treated as being an abbreviation of $A \supset \bot$. Initial sequents are $A \longrightarrow A$ with A required to be a propositional variable, and $\bot \longrightarrow$.

Definition Let P be a proof. The *closure, $cl(P)$, of P* is the smallest set of sequents which contains the sequents of P and is closed under the cut rule and weakenings.

Note that for intuitionistic proofs P, the sequents of P are Horn clauses. Therefore SLD resolution algorithms may be used to solve the following problem in polynomial time (see [10]): (a) Given P and Γ, list the set of formulas A in P such that $\Gamma \longrightarrow A$ is in $Cl(P)$. Hence also: (b) Given P, Γ and A, is $\Gamma \longrightarrow A$ in $Cl(P)$, and is $\Gamma \longrightarrow$ in $Cl(P)$?

Theorem 1 *Let P be a propositional intuitionistic proof of $\Gamma \longrightarrow A$. Then there is a cut-free proof P' of $\Gamma \longrightarrow A$ such that $cl(P') \subseteq cl(P)$.*

Proof We shall prove the theorem by showing that it is possible to eliminate the cuts in P one at a time, without adding any new sequents to the closure of the proof. Unlike the usual proof of the cut-elimination theorem where the proof is transformed by a series of "local transformations" of the proof using induction on right rank and left rank, we shall use a series of global transformations similar to the approach used in [2] (although the method used there will not work for the present proof).

First, consider the case where a cut on an implication $A \supset B$ is to be removed from the proof. Consider the subproof of P which ends with the cut inference

$$\dfrac{\begin{matrix} \ddots\vdots\cdot^{Q} \\ \Gamma_1 \longrightarrow A \supset B \end{matrix} \qquad \begin{matrix} \ddots\vdots\cdot^{R} \\ A \supset B, \Gamma_2 \longrightarrow C \end{matrix}}{\Gamma_1, \Gamma_2 \longrightarrow C}$$

We call the sequents of Q that contain direct ancestors of $A \supset B$ in their succedent, the *lower part of Q*. At the upper boundary of the lower part of Q, there are k many inferences which have a direct ancestor of the cut formula as principal formula:

$$\dfrac{\dfrac{\vdots\, Q^i}{\Pi_i, A \longrightarrow B}}{\Pi_i \longrightarrow A \supset B}$$

for $i = 1, \ldots k$. There may also be direct ancestors of the cut formula introduced by weakening inferences, but no direct ancestor appears in an initial sequent because of our convention that initial sequents have atomic formulas.

Similarly, in the proof R, consider the subproofs R_j which have a direct ancestor of the cut formula as principal formula:

$$\dfrac{\dfrac{\vdots}{\Delta_j \longrightarrow A} \qquad \dfrac{\vdots}{B, \Delta'_j \longrightarrow D_j}}{A \supset B, \Delta_j, \Delta'_j \longrightarrow D_j}$$

for $j = 1, \ldots m$. There may also be direct ancestors of the cut formula introduced by weakening inferences.

For each $i = 1, \ldots k$, form the proof R^i by modifying R as follows: First replace the last inference of each R_j with two cuts:

$$\dfrac{\dfrac{\dfrac{\vdots}{\Delta_j \longrightarrow A} \qquad \dfrac{\vdots\, Q^i}{\Pi_i, A \longrightarrow B}}{\Pi_i, \Delta_j \longrightarrow B} \qquad \dfrac{\vdots}{B, \Delta'_j \longrightarrow D_j}}{\Pi_i, \Delta_j, \Delta'_j \longrightarrow D_j}$$

Then modify the rest of the lower part of R by replacing direct ancestors of the cut formula with the cedent Π_i. (Direct ancestors of the cut formula which are introduced by weakening inferences may be so replaced by using a series of weakening inferences to introduce the formulas in Π_i.) This changes sequents of the form $A \supset B, \Delta \longrightarrow D$ to sequents of the form $\Pi_i, \Delta \longrightarrow D$. In this way, we obtain a proof R^i with end-sequent $\Pi_i, \Gamma_2 \longrightarrow C$. It is easy to check that $cl(R^i) \subseteq cl(P)$.

To finish the elimination of the cut on $A \supset B$ from P, we replace each subproof Q^i of P with the proof R^i and we replace each sequent $\Pi \longrightarrow A \supset B$ in the lower part of Q with $\Pi, \Gamma_2 \longrightarrow C$. The result is a proof of $\Gamma_1, \Gamma_2 \longrightarrow C$ in which the cut on $A \supset B$ has been eliminated. It is easy to check that the closure of the new proof is a subset of $cl(P)$.

Now consider the case of removing a cut with principal formula a disjunction. Let some subproof of P end with a cut

$$\dfrac{\dfrac{\vdots\, Q}{\Gamma_1 \longrightarrow A \vee B} \qquad \dfrac{\vdots\, R}{A \vee B, \Gamma_2 \longrightarrow C}}{\Gamma_1, \Gamma_2 \longrightarrow C}$$

Define the *lower part of* Q similarly to the previous case. At the upper boundary of the lower part of Q, there are k many inferences which have a direct ancestor of the cut formula as principal formula:

$$\dfrac{\begin{array}{c}\vdots\, Q^i \\ \Pi_i \longrightarrow X_i\end{array}}{\Pi_i \longrightarrow A \vee B}$$

for $i = 1, \ldots k$ where each X_i is either A or B. There may also be direct ancestors of the cut formula introduced by weakening inferences, but no direct ancestor appears in an initial sequent.

Similarly, in the proof R, consider the subproofs R_j which have a direct ancestor of the cut formula as principal formula:

$$\dfrac{\begin{array}{c}\vdots \\ A, \Delta_j \longrightarrow D_j\end{array} \qquad \begin{array}{c}\vdots \\ B, \Delta'_j \longrightarrow D_j\end{array}}{A \vee B, \Delta_j, \Delta'_j \longrightarrow D_j}$$

for $j = 1, \ldots m$. There may also be direct ancestors of the cut formula introduced by weakening inferences.

For each $i = 1, \ldots k$, form the proof R^i by modifying R as follows. We assume X_i is A: the case where X_i is B is completely similar. First, replace the last inference of each R_j by a cut and weakenings:

$$\dfrac{\dfrac{\begin{array}{c}\vdots\, Q^i \\ \Pi_i \longrightarrow A\end{array} \qquad \begin{array}{c}\vdots \\ A, \Delta_j \longrightarrow D_j\end{array}}{\Pi_i, \Delta_j \longrightarrow D_j}}{\Pi_i, \Delta_j, \Delta'_j \longrightarrow D_j}$$

Then modify the rest of the lower part of R by replacing direct ancestors of the cut formula with the cedent Π_i. This changes sequents of the form $A \vee B, \Delta \longrightarrow D$ to sequents of the form $\Pi_i, \Delta \longrightarrow D$. In this way, we obtain a proof R^i with end-sequent $\Pi_i, \Gamma_2 \longrightarrow C$. It is easy to check that $cl(R^i) \subseteq cl(P)$.

To finish the elimination of the cut on $A \vee B$ from P, we replace each subproof Q^i of P with the proof R^i and we replace each sequent $\Pi \longrightarrow A \supset B$ in the lower part of Q with $\Pi, \Gamma_2 \longrightarrow C$. The result is a proof of $\Gamma_1, \Gamma_2 \longrightarrow C$ in which the cut on $A \vee B$ has been eliminated. It is easy to check that the closure of the new proof is a subset of $cl(P)$.

Now consider the case where the cut formula is a conjunction. Let some subproof of P end with a cut

$$\dfrac{\begin{array}{c}\vdots\, Q \\ \Gamma_1 \longrightarrow A \wedge B\end{array} \qquad \begin{array}{c}\vdots\, R \\ A \wedge B, \Gamma_2 \longrightarrow C\end{array}}{\Gamma_1, \Gamma_2 \longrightarrow C}$$

At the upper boundary of the lower part of Q, there are k many inferences which have a direct ancestor of the cut formula as principal formula:

$$\dfrac{\overset{\vdots\; Q^i_A}{\Pi_i \longrightarrow A} \qquad \overset{\vdots\; Q^i_B}{\Pi'_i \longrightarrow B}}{\Pi_i, \Pi'_i \longrightarrow A \wedge B}$$

for $i = 1, \ldots k$.

Similarly, in the proof R, consider the subproofs R_j which have a direct ancestor of the cut formula as principal formula:

$$\dfrac{\overset{\vdots}{X_j, \Delta_j \longrightarrow D_j}}{A \wedge B, \Delta_j \longrightarrow D_j}$$

for $j = 1, \ldots m$, where X_j is either A or B.

For each $i = 1, \ldots k$, form the proof R^i by modifying R as follows. We assume X_j is A: the case where X_j is B is completely similar. First, replace the last inference of each R_j by a cut and weakenings:

$$\dfrac{\dfrac{\overset{\vdots\; Q^i}{\Pi_i \longrightarrow A} \qquad \overset{\vdots}{A, \Delta_j \longrightarrow D_j}}{\Pi_i, \Delta_j \longrightarrow D_j}}{\Pi_i, \Pi'_i, \Delta_j \longrightarrow D_j}$$

Then modify the rest of the lower part of R by replacing direct ancestors of the cut formula with the cedent Π_i, Π'_i. This changes sequents of the form $A \wedge B, \Delta \longrightarrow D$ to sequents of the form $\Pi_i, \Pi'_i, \Delta \longrightarrow D$. In this way, we obtain a proof R^i with end-sequent $\Pi_i, \Pi'_i, \Gamma_2 \longrightarrow C$. It is easy to check that $cl(R^i) \subseteq cl(P)$.

To finish the elimination of the cut on $A \wedge B$ from P, we replace each subproof Q^i of P with the proof R^i and we replace each sequent $\Pi \longrightarrow A \wedge B$ in the lower part of Q with $\Pi, \Gamma_2 \longrightarrow C$. The result is a proof of $\Gamma_1, \Gamma_2 \longrightarrow C$ in which the cut on $A \wedge B$ has been eliminated. It is easy to check that the closure of the new proof is a subset of $cl(P)$.

Finally consider the case where a cut on an atomic formula A is to be removed from the proof. Let the subproof P end with the cut inference

$$\dfrac{\overset{\vdots\; Q}{\Gamma_1 \longrightarrow A} \qquad \overset{\vdots\; R}{A, \Gamma_2 \longrightarrow C}}{\Gamma_1, \Gamma_2 \longrightarrow C}$$

This cut can be eliminated as follows: everywhere where a sequent $\Pi \longrightarrow A$ appears in Q with the succedent A a direct ancestor of the cut formula, replace this sequent with $\Pi, \Gamma_2 \longrightarrow C$. Where the direct ancestor was in an initial sequent, the initial sequent is also replaced by a copy of the subproof R. Where the direct ancestor was introduced by weakening, the formulas in Γ_2, C are introduced by weakening inferences. It is easy to verify that this process eliminates the cut and does not add new sequents to the closure of the proof.

Any case where the cut formula is introduced on either the left or the right only by weakening inferences is entirely trivial. This includes the elimination of cuts where the cut formula is $\perp$, since the formula $\perp$ can be introduced on the right only by a weakening inference.

3 A realizability theorem

We shall call a formula a *disjunction,* a *conjunction,* or an *implication,* if its outermost connective is the corresponding connective.

Lemma 2 *(a) Let P be a cut-free proof of a sequent $\Gamma \longrightarrow B_0 \vee B_1$ such that Γ contains only atomic formulas and implications. Then*

1. *For some $i = 0, 1$, $\Gamma \longrightarrow B_i$ is in $Cl(P)$, or,*
2. *For some implication $C \supset D$ in Γ, we have $\Gamma \longrightarrow C$ in $Cl(P)$.*

(b) The same assuming that P is a general proof (not necessarily cut-free) and $\Gamma \longrightarrow B_0 \vee B_1$ is only an element of $Cl(P)$.

Proof (a) Consider the last inference in the proof. W.l.o.g., the last inference is not a weakening:left. Therefore we may assume the last inference is an application of one of the following rules: weakening:right, $\vee$:right, or $\supset$:left. If it is weakening:right or $\vee$:right, then we get case 1 of the lemma. If it is $\supset$:left, then we get case 2 of the lemma. Indeed, the last two lines of the proof have the form

$$\frac{\Sigma \longrightarrow C \quad D, \Pi \longrightarrow B_0 \vee B_1}{C \supset D, \Sigma, \Pi \longrightarrow B_0 \vee B_1}$$

which gives case 2.

(b) From Theorem 1 it follows that we can prove the same without the assumption that the proof P is cut-free. If we only have $\Gamma \longrightarrow B_0 \vee B_1 \in Cl(P)$, then we first construct a proof P' of $\Gamma \longrightarrow B_0 \vee B_1$ using only sequents from $Cl(P)$ and then apply part (a).

Now we can prove our main theorem. To state the theorem we need to introduce a special type of interactive computations. Let a proof P of a sequent $\Gamma \longrightarrow B_0 \vee B_1$ be given. An interactive computation is carried out by an oracle Turing machine, with the aim of constructing a proof of a sequent $\Delta \longrightarrow B_i$ for some $i = 0$ or 1 and a set of propositions Δ which are subformulas of Γ and are implied by Γ. We will call Δ the set of *established* formulas. Initially Δ

will equal Γ, and then the computation of the Turing machine will iteratively update Δ in cooperation with an oracle. The following are the rules.

1. The Turing machine starts with a proof P of a sequent $\Gamma \longrightarrow B_0 \vee B_1$. We denote the current set of *established* formulas by Δ, and initially Δ equals Γ.

2. The machine may replace a conjunction $C \wedge D$ in Δ by the two formulas C and D.

3. If Δ contains an implication $C \supset D$ and if the machine succeeds in finding an intuitionistic proof of C from the established formulas Δ, then the machine may replace the formula $C \supset D$ with D.

4. If Δ contains a disjunction $C \vee D$, the machine can pose the query $C \vee D$ to the oracle. The oracle is obliged to pick one of the disjuncts. Then the machine replaces $C \vee D$ by the disjunct picked by the oracle.

5. The computation ends when the machine succeeds in finding a proof of $\Delta \longrightarrow B_i$, for some $i = 0, 1$.

Theorem 3 *There exists an oracle machine obeying the rules above which for every proof P of a sequent $\Gamma \longrightarrow B_0 \vee B_1$ finishes the computation in polynomial time in the size of P.*

In order to prove the theorem, we extend the notion of the closure of a proof:

Definition The *extended closure* of a proof P, denoted by $Cl^+(P)$, is defined to equal the closure under weakening and the cut rule of the set S of sequents defined as follows: S contains all the sequents of P, plus the sequents $A \wedge B \longrightarrow A$, $A \wedge B \longrightarrow B$, $C \longrightarrow C \vee D$, $D \longrightarrow C \vee D$ and $E \longrightarrow F \supset E$ for all formulas $A \wedge B$, $C \vee D$ and $E \supset F$ which occur (possibly as subformulas) in P.

Since the number of sequents in S is certainly bounded by twice the number of formulas occurring in the proof, one can test the presence of a sequent in the extended closure in polynomial time, in the same way as for the ordinary closure. Also it is clear that Theorem 1 and Lemma 2 still hold with the ordinary closure replaced by the extended closure.

Proof (of Theorem 3). First observe that when the machine modifies the sequent according to the rules, the sequent remains in $Cl^+(P)$ and it is always reduced to a simpler one. So we need only to show that one of the rules can always be applied and each round can be done in polynomial time. That follows from Lemma 2, part 2 and the fact that we can test the presence of a sequent in $Cl^+(P)$ in polynomial time, and, in case it is there, we can construct its proof in polynomial time.

Theorem 3 can be further generalized to arbitrary formulas in the succedent as follows. Given a sequent of the form $\Gamma \longrightarrow A$ with a general A, the computation is defined by the following clauses:

1. if A is a disjunction, then compute as above to get one of the terms of the disjunction; then replace A by this term;

2. if A is a conjunction, then the computation splits into two branches corresponding to the terms of A;

3. if A is an implication, then move the antecedent of the implication to the antecedent of the sequent;

4. the computation stops, if A is a propositional variable or $\perp$.

It follows from Theorem 3 that such a computation always stops after a polynomial number of steps in the size of a proof of the sequent $\Gamma \longrightarrow A$. Notice that the parallelism inherent in the treatment of conjunctions can be removed if clause 2. is replaced by a clause saying that the oracle chooses one of the terms.

4 The disjunction property with Harrop hypotheses and the feasible interpolation theorem

Definition The *disjunction problem for propositional intuitionistic logic* is the following problem: Given an intuitionistic proof of a disjunction $A \vee B$, determine one of A and B to be intuitionistically valid.

It is well-known that a disjunction $A \vee B$ is intuitionistically valid iff A or B is intuitionistically valid. A classical result of Harrop generalizes this result to sequents $\Gamma \longrightarrow A \vee B$ where Γ is a set of so called *Harrop formulas.* These are defined by:

1. every atomic formula is Harrop, $\perp$ is Harrop;

2. if A and B are Harrop, then $A \wedge B$ is Harrop;

3. if A is arbitrary and B is Harrop, then $A \supset B$ is Harrop;

4. no other formulas are Harrop.

Some variations on the disjunction property include (a) the *disjoint variable disjunction property* where A and B are required to have no variables in common, or (b) the *strong disjunction property* where an intuitionistic proof of either A or B must be produced.

From Buss-Mints [3], we know the strong disjunction property is in *PTIME.* In this section we shall prove such a result for sequents whose antecedents consist of disjunctions of Harrop formulas. It is a corollary of Theorem 3. In Section 5 we shall prove converse results, in particular a lower bound on the complexity of the disjunction property.

Theorem 4 *There is a polynomial time algorithm which for a given intuitionistic proof P of a sequent*

$$A_1, \dots, A_n \longrightarrow B_1 \vee \dots \vee B_m, \tag{1}$$

where A_k are Harrop formulas, constructs a proof P' of

$$A_1, \dots, A_n \longrightarrow B_i,$$

for some $1 \le i \le m$.

Proof The algorithm from Theorem 3 can obviously be generalized so that when it is applied to a sequent $\Gamma \longrightarrow B_1 \vee \dots \vee B_m$, it eventually produces a proof of $\Delta \longrightarrow B_i$, for some i and some established formulas Δ. Since Γ consists of only Harrop formulas, it is easy to see that only Harrop formulas can be obtained as established formulas. In particular, it never happens in the course of computation that we get a disjunction in the antecedent, and thus the machine never queries the oracle about a disjunction. In addition, all antecedents, in particular the last one, are Cl^+-derivable from Γ. Hence $A_1, \dots, A_n \longrightarrow B_i$ is in $Cl^+(P)$ for some i, which fact can be tested in polynomial time.

Corollary 5 *Let P be an intuitionistic proof of*

$$A_{1,1} \vee \dots \vee A_{1,l_1}, \dots, A_{n,1} \vee \dots \vee A_{n,l_n} \longrightarrow B_1 \vee, \dots \vee B_m, \tag{2}$$

with all $A_{k,j}$ Harrop. Then for every $j_1, \dots, j_n$, where $1 \le j_k \le l_k$, there exists an i, $1 \le i \le m$, such that

$$A_{1,j_1}, \dots, A_{n,j_n} \longrightarrow B_i \tag{3}$$

is intuitionistically valid. Moreover such an assignment $j_1, \dots, j_n \to i$ can be computed in polynomial time in the size of P and also proofs of the corresponding sequents (3) can be computed in polynomial time.

Proof Given a proof P of (2) and $j_1, \dots, j_n$ we can construct easily a proof of

$$A_{1,j_1}, \dots, A_{n,j_n} \longrightarrow B_1 \vee \dots \vee B_m$$

by adding the proofs of the sequents

$$A_{k,j_k} \longrightarrow A_{k,1} \vee \dots \vee A_{k.l_k}$$

Now the corollary follows from Theorem 4.

Corollary 6 (Feasible Interpolation Theorem) *Suppose an intuitionistic proof P of*

$$x_1 \vee \neg x_1, \dots, x_n \vee \neg x_n \longrightarrow B_0 \vee B_1 \tag{4}$$

is given. Then it is possible to construct a circuit $\mathcal{C}(\vec{x})$ whose size is polynomial in the size of P such that for every input $\vec{a} \in \{0,1\}^n$, if $\mathcal{C}(\vec{a}) = i$, then $B_i(\vec{x}/\vec{a})$ (i.e., B_i where we substitute for variables x_j $\bot$, if $x_j = 0$ and $\top$, if $x_j = 1$) is a tautology.

We do not require that the variables x_i are the only common variables of B_0 and B_1, but we do not know of any application of the case when the formulas share more variables.

The interpretation of the statement, when $\vec{x}$ are the only common variables, is as follows. Suppose $B_0(\vec{x},\vec{y}) \vee B_1(\vec{x},\vec{z})$ is a classical tautology. Then for any substitution of truth values for the common variables one of the two subformulas must be a tautology. In the intuitionistic calculus such a disjunction cannot be a tautology, unless, trivially, one of the subformulas is. But it is possible that (4) is an intuitionistically valid sequent, since the excluded middle laws for variables x_i express that we "know the truth values of these variables". The statement demonstrates the constructive character of the intuitionistic calculus: having a proof of (4) and "knowing" the variables x_i we should be able to tell which of the subformulas is true.

Corollary 7 *If $NP \cap coNP \not\subseteq P/poly$ (more generally, if there exists a pair of disjoint NP sets which cannot be separated by a $P/poly$ set), then the lengths of shortest proofs in the intuitionistic propositional calculus cannot be bounded by a polynomial of the size of the proved formula.*

Proof This is proved from Corollary 6 using ideas of Mundici [6]. Suppose that Q is a predicate in $NP \cap coNP$. Then there are families of formulas $B_{0,i}(\vec{x},\vec{y})$ and $B_{1,i}(\vec{x},\vec{z})$, with the i specifying the number of $\vec{x}$ variables, such that $\exists\vec{y}B_{0,i}(x,\vec{y})$ is equivalent to $Q(\vec{x})$ and $\exists\vec{z}B_{1,i}(x,\vec{z})$ is equivalent to $\neg Q(\vec{x})$. The formulas $B_{0,i} \vee B_{1,i}$ are tautologies and hence intuitionistically provable. If there is a polynomial bound on the size of the intuitionistic proofs of these tautologies, then, by Corollary 6, Q is in $P/poly$. □

It is generally accepted as plausible conjectures that factoring and discrete logarithm cannot be computed in polynomial time. Both conjectures imply $NP \cap coNP \not\subseteq P/poly$. Note that the well-known $PSpace$-completeness of the propositional intuitionistic calculus implies that there is no polynomial bound on the proofs assuming $PSpace \not\subseteq NP$. These two complexity theoretical assumptions do not seem to be comparable.

Corollary 8 *Assuming that factoring is not computable in polynomial time, there is more than polynomial speed-up between classical and intuitionistic propositional calculus, i.e., there are intuitionistic tautologies that have polynomial size proofs in the classical sequent calculus, but no polynomial size proofs in the intuitionistic sequent calculus.*

Proof Bonnet, Pitassi and Raz [1] constructed tautologies which have polynomial size proofs in the classical sequent calculus and which cannot have such proofs in any system admitting feasible interpolation, provided that factoring is hard.

Let us note that such a speed-up follows also from the assumption that $PSpace \not\subseteq NP$, but the last corollary gives more concrete examples on which this speed-up is achieved.

5 The P-hardness of the disjunction property

The following is, in some sense, a converse to Corollary 2.

Theorem 9 *Let $\mathcal{C}(x_1, \ldots, x_n)$ a boolean circuit be given. Then it is possible to construct in logarithmic space formulas B_0, B_1 and an* intuitionistic *proof P of (4) such that for all $\vec{a} \in \{0,1\}^n$, we have $\mathcal{C}(\vec{a}) = i$ if and only if $B_i(\vec{a}, \vec{u})$ is an intuitionistic tautology. Further, when $\mathcal{C}(\vec{a}) = i$, the intuitionistic proof of $B_i(\vec{a}, \vec{u})$ can can be constructed in polynomial time given $\mathcal{C}$ and $\vec{a}$.*

Proof Given a circuit $\mathcal{C}$ with inputs $x_1, \ldots, x_n$, we construct the formulas B_0 and B_1 as follows. Without loss of generality the only gates in $\mathcal{C}$ are NOT gates and AND gates. With each input signal and each gate in $\mathcal{C}$, associate a distinct Boolean variable y_i $(i = 1, \ldots, m)$. With each y_i we associate two or three formulas, depending on how y_i is computed in $\mathcal{C}$:

In case y_i is an input signal x_j, the two formulas associated with y_i are

$$x_i \supset y_i \quad \text{and} \quad \neg x_i \supset \neg y_i.$$

In case y_i is the output of a NOT gate with input y_j, the two formulas associated with y_i are:

$$y_j \supset \neg y_i \quad \text{and} \quad \neg y_j \supset y_i$$

In case y_i is the output of an AND gate with inputs y_j and y_k the three formulas associated with y_i are:

$$y_j \wedge y_k \supset y_i \quad \text{and} \quad \neg y_j \supset \neg y_i \quad \text{and} \quad \neg y_k \supset \neg y_i.$$

Now define the formula $C(\vec{x}, \vec{y})$ to be the conjunction of all the formulas associated with the y_i's. Let $B_0(\vec{x}, \vec{y})$ be the formula $C(\vec{x}, \vec{y}) \supset y_m$ and let $B_1(\vec{x}, \vec{y})$ be the formula $C(\vec{x}, \vec{y}) \supset \neg y_m$, where y_m is the output signal. For conciseness, let $EM(\vec{x})$ be the conjunction of the formulas $x_i \vee \neg x_i$.

We shall show that a proof of the sequent

$$EM(\vec{x}) \longrightarrow B_0(\vec{x}, \vec{y}) \vee B_1(\vec{x}, \vec{y}).$$

can be constructed in logarithmic space. Then the rest of the theorem follows easily from the construction of the sequent.

To construct the proof, proceed inductively on i giving intuitionistic proofs of the sequents

$$EM(\vec{x}) \longrightarrow (C(\vec{x}, \vec{y}) \supset y_i) \vee (C(\vec{x}, \vec{y}) \supset \neg y_i).$$

Both the base step and the inductive steps are easy, we shall consider only the inductive steps. Let y_i be the output of a NOT gate with input y_j, then the sequents

$$C(\vec{x}, \vec{y}) \supset y_j \longrightarrow C(\vec{x}, \vec{y}) \supset \neg y_i$$

and

$$C(\vec{x}, \vec{y}) \supset \neg y_j \longrightarrow C(\vec{x}, \vec{y}) \supset y_i.$$

have simple, short intuitionistic proofs. Let y_i be the output of an AND gate with inputs y_j and y_k. We first derive

$$EM(\vec{x}) \longrightarrow (C(\vec{x},\vec{y}) \supset y_j \wedge y_k) \vee (C(\vec{x},\vec{y}) \supset \neg y_j) \vee (C(\vec{x},\vec{y}) \supset \neg y_k)$$

and then we apply the clauses for the gate in a similar fashion as above.

Thus we get only

$$EM(\vec{x}) \longrightarrow B_0(\vec{x},\vec{y}) \vee B_1(\vec{x},\vec{y}),$$

ie., the formulas in the disjunction share also the variables $\vec{y}$. In order to get (4) we only need to prove inductively slightly more complicated sequents

$$EM(\vec{x}) \longrightarrow [(C(\vec{x},\vec{y}) \supset y_i) \wedge (C(\vec{x},\vec{z}) \supset z_i)] \vee [(C(\vec{x},\vec{y}) \supset \neg y_i) \wedge (C(\vec{x},\vec{z}) \supset \neg z_i)],$$

and otherwise proceed similarly. For $i = m$ such a sequent is clearly stronger than the sequent (4) that we need.

The construction can be performed in logarithmic space, since each step of the proof is explicitly and easily determined by the circuit.

Corollary 10 *The disjunction property is P-hard with respect to logarithmic space reductions. In fact, the disjoint variable disjunction property is P-hard.*

Proof The previous theorem gives actually a logarithmic space reduction of the P-complete problem circuit value to the disjunction problem.

6 Cut elimination for first-order logic

In this section we extend the definition of the closure of a proof to sequent calculus proofs in first-order logic and prove the analogue of Theorem 1 that cut elimination can be performed on intuitionistic proofs without adding new sequents to the closure of the proof.

Definition Let P be a sequent calculus proof in first-order logic. The *closure, $cl(P)$, of P* is the smallest set of sequents which contains the sequents of P and is closed under the cut rule, under weakening, and under term substitution.

By "term substitution", we mean uniformly substituting a term for a free variable in the sequent.

Unlike the situation for propositional logic, we no longer have a polynomial time algorithm for deciding membership of sequents in the closure of P.

Theorem 11 *Let P be a first-order intuitionistic proof of $\Gamma \longrightarrow A$. Then there is a cut-free proof P' of $\Gamma \longrightarrow A$ such that $cl(P') \subseteq cl(P)$.*

Proof The general idea of the proof is exactly like the proof of Theorem 1. We will consider only the new cases where the cut to be eliminated has a cut formula with outermost connective a quantifier. Without loss of generality, the proof is in free variable normal form and is converted back to free variable normal after each elimination of a cut.

The cases of eliminating cuts on formulas which have outermost connective propositional, or which are atomic are exactly as in the proof of Theorem 1, so we do not repeat them here.

Instead, first consider the case where the cut formula has outermost connective a universal quantifier. Let some subproof of P end with a cut

$$\dfrac{\begin{array}{c}\ddots\vdots\cdot^{\cdot}Q\\ \Gamma_1\longrightarrow(\forall x)A(x)\end{array}\qquad\begin{array}{c}\ddots\vdots\cdot^{\cdot}R\\ (\forall x)A(x),\Gamma_2\longrightarrow C\end{array}}{\Gamma_1,\Gamma_2\longrightarrow C}$$

At the upper boundary of the lower part of Q, there are k many inferences which have a direct ancestor of the cut formula as principal formula:

$$\dfrac{\begin{array}{c}\ddots\vdots\cdot^{\cdot}Q^i\\ \Pi_i\longrightarrow A(b_i)\end{array}}{\Pi_i\longrightarrow(\forall x)A(x)}$$

for $i = 1, \ldots k$, where the b_i's are distinct eigenvariables.

Similarly, in the proof R, consider the subproofs R_j which have a direct ancestor of the cut formula as principal formula:

$$\dfrac{\begin{array}{c}\ddots\vdots\cdot^{\cdot}\\ A(t_j),\Delta_j\longrightarrow D_j\end{array}}{(\forall x)A(x),\Delta_j\longrightarrow D_j}$$

for $j = 1, \ldots m$.

For each $i = 1, \ldots k$, form the proof R^i by modifying R as follows. First, form the proof $Q^i(t_j/b_i)$ by replacing each occurrence of b_i with t_j. Then, replace the last inference of each R_j by

$$\dfrac{\begin{array}{c}\ddots\vdots\cdot^{\cdot}Q^i(t_j/b_i)\\ \Pi_i\longrightarrow A(t_j)\end{array}\qquad\begin{array}{c}\ddots\vdots\cdot^{\cdot}\\ A(t_j),\Delta_j\longrightarrow D_j\end{array}}{\Pi_i,\Delta_j\longrightarrow D_j}$$

Then modify the rest of the lower part of R by replacing direct ancestors of the cut formula with the cedent Π_i. Weakening inferences which introduce direct ancestors are replaced by weakenings which introduce the formulas in Π_i. This changes sequents of the form $(\forall x)A(x), \Delta\longrightarrow D$ to sequents of the form $\Pi_i, \Delta\longrightarrow D$. In this way, we obtain a proof R^i with end-sequent $\Pi_i, \Gamma_2\longrightarrow C$. It is easy to check that $cl(R^i) \subseteq cl(P)$.

To finish the elimination of the cut on $\forall xA(x)$ from P, we replace each subproof Q^i of P with the proof R^i and we replace each sequent $\Pi\longrightarrow(\forall x)A(x)$

in the lower part of Q with $\Pi, \Gamma_2 \longrightarrow C$. The result is a proof of $\Gamma_1, \Gamma_2 \longrightarrow C$ in which the cut on $(\forall x)A(x)$ has been eliminated. It is easy to check that the closure of the new proof is a subset of $cl(P)$.

Now consider the case where the cut formula has outermost connective n existential quantifier. Let some subproof of P end with a cut

$$\frac{\begin{array}{c}\ddots\vdots\cdot^{\,Q}\\ \Gamma_1 \longrightarrow (\exists x)A(x)\end{array} \qquad \begin{array}{c}\ddots\vdots\cdot^{\,R}\\ (\exists x)A(x), \Gamma_2 \longrightarrow C\end{array}}{\Gamma_1, \Gamma_2 \longrightarrow C}$$

At the upper boundary of the lower part of Q, there are k many inferences which have a direct ancestor of the cut formula as principal formula:

$$\frac{\begin{array}{c}\ddots\vdots\cdot^{\,Q^i}\\ \Pi_i \longrightarrow A(t_i)\end{array}}{\Pi_i \longrightarrow (\exists x)A(x)}$$

for $i = 1, \ldots k$.

Similarly, in the proof R, consider the subproofs R_j which have a direct ancestor of the cut formula as principal formula:

$$\frac{\begin{array}{c}\ddots\vdots\cdot\\ A(b_j), \Delta_j \longrightarrow D_j\end{array}}{(\exists x)A(x), \Delta_j \longrightarrow D_j}$$

for $j = 1, \ldots m$.

For each $i = 1, \ldots k$, form the proof R^i by modifying R as follows. First, form the proof $R_j(t_i/b_j)$ from R_j by replacing each occurrence of b_j with t_i. Then, replace the last inference of each $R_j(t_i/b_j)$ by

$$\frac{\begin{array}{c}\ddots\vdots\cdot^{\,Q^i}\\ \Pi_i \longrightarrow A(t_i)\end{array} \qquad \begin{array}{c}\ddots\vdots\cdot\\ A(t_i), \Delta_j \longrightarrow D_j\end{array}}{\Pi_i, \Delta_j \longrightarrow D_j}$$

Then modify the rest of the lower part of R by replacing direct ancestors of the cut formula with the cedent Π_i. As usual, weakening inferences which introduce direct ancestors are replaced by weakenings which introduce the formulas in Π_i. This changes sequents of the form $(\exists x)A(x), \Delta \longrightarrow D$ to sequents of the form $\Pi_i, \Delta \longrightarrow D$. In this way, we obtain a proof R^i with end-sequent $\Pi_i, \Gamma_2 \longrightarrow C$. It is easy to check that $cl(R^i) \subseteq cl(P)$.

To finish the elimination of the cut on $\exists x A(x)$ from P, we replace each subproof Q^i of P with the proof R^i and we replace each sequent $\Pi \longrightarrow (\exists x)A(x)$ in the lower part of Q with $\Pi, \Gamma_2 \longrightarrow C$. The result is a proof of $\Gamma_1, \Gamma_2 \longrightarrow C$ in which the cut on $(\exists x)A(x)$ has been eliminated. It is easy to check that the closure of the new proof is a subset of $cl(P)$.

References

[1] M. L. Bonet, T. Pitassi, and R. Raz, *No feasible interpolation for TC^0-Frege proofs*, in Proceedings of the 38th Annual Symposium on Foundations of Computer Science, Piscataway, New Jersey, 1997, IEEE Computer Society, pp. 264–263.

[2] S. R. Buss, *An introduction to proof theory*, in Handbook of Proof Theory, S. R. Buss, ed., North-Holland, 1998, pp. 1–78.

[3] S. R. Buss and G. Mints, *The complexity of the disjunction and existence properties in intuitionistic logic*, Annals of Pure and Applied Logic, 99 (1999), pp. 93–104.

[4] A. Goerdt, *Efficient interpolation for the intuitionistic sequent calculus*, preprint, Technische Universität Chemnitz, CSR-00-02, January 2000.

[5] J. Krajíček, *Interpolation theorems, lower bounds for proof systems and independence results for bounded arithmetic,* JSL 62, (1997) pp.. 457-486.

[6] D. Mundici, *Tautologies with a unique Craig interpolant*, Annals of Pure and Applied Logic, 27 (1984), pp. 265–273.

[7] P. Pudlák, *Lower bounds for resolution and cutting plane proofs and monotone computations,* JSL 62, (1997) pp. 981-998.

[8] P. Pudlák, *On the complexity of propositional calculus, Sets and proofs*, in Logic Colloquium '97, Cambridge University Press, 1999, pp. 197–218.

[9] P. Pudlák and J. Sgall, *Algebraic models of computation and interpolation for algebraic systems*, DIMACS Series in Discrete Math. and Theor. Comp. Sci. 39, (1998), pp. 279-295.

[10] U. Schöning, *Logik für Informatiker*, Wissenschaftsverlag, 1989.