Electronic Colloquium on Computational Complexity, Report No. 142 (2006)

On the Deduction Theorem and Complete Disjoint NP-Pairs

Olaf Beyersdorff*

Institut für Informatik, Humboldt-Universität zu Berlin, Germany
beyersdo@informatik.hu-berlin.de

Abstract. In this paper we ask the question whether the extended Frege proof system EF satisfies a weak version of the deduction theorem. We prove that if this is the case, then complete disjoint NP-pairs exist. On the other hand, if EF is an optimal proof system, then the weak deduction theorem holds for EF. Hence the weak deduction property for EF is a natural intermediate condition between the optimality of EF and the completeness of its canonical pair. We also exhibit two conditions that imply the completeness of the canonical pair of Frege systems.

1 Introduction

Although disjoint NP-pairs were already introduced into complexity theory in the 80's by Grollmann and Selman [GS88] it was only during recent years that disjoint NP-pairs have fully come into the focus of complexity-theoretic research (cf. e.g. [Pud03,GSSZ04,GSS05,GSZ06]). This interest mainly stems from the applications of disjoint NP-pairs to such different areas as cryptography [GS88,HS92] and propositional proof complexity [Raz94,Pud03,Kra04,Bey04].

Similarly as for other promise classes it is not known whether the class of all disjoint NP-pairs contains pairs that are complete under the appropriate reductions. This question, posed by Razborov [Raz94], is one of the most prominent open problems in the field. On the positive side, it is known that the existence of optimal proof systems suffices to guarantee the existence of complete pairs [Raz94]. More towards the negative, a body of sophisticated relativization results underlines the difficulty of the problem. Glaßer et al. [GSSZ04] provided an oracle under which there exist complete disjoint NP-pairs. On the other hand, they also constructed an oracle relative to which there exist complete pairs but optimal proof systems do not exist.

Further information on the problem is provided by a number of different characterizations. Glaßer, Selman, and Sengupta [GSS05] obtained a condition in terms of uniform enumerations of machines and also proved that the question of the existence of complete pairs receives the same answer under reductions of different strength. Additionally, the problem was characterized by provability conditions in propositional proof systems and shown to be robust under an increase of the number of components from two to arbitrary constants [Bey06b].

In this paper we exhibit several sufficient conditions for the existence of complete disjoint NP-pairs which involve properties of concrete proof systems such as Frege systems and their extensions. In particular, we link the problem

* Supported by DFG grant KO 1053/5-1

ISSN 1433-8092

on complete pairs with the question whether the extended Frege proof system satisfies a weak version of the deduction property. The deduction theorem for propositional logic explains how a proof of a formula ψ from an extra hypothesis φ is transformed to a proof of $\varphi \rightarrow \psi$. This property is known to hold for Frege systems [BB93], but fails for extended or substitution Frege systems as neither the extension nor the substitution rule is sound. We therefore relax the condition by requiring the extra hypothesis φ to be tautological.

Whether this weakened version of the deduction property holds for EF appears to be a natural problem of intermediate strength between the problems on the existence of optimal proof systems and complete NP-pairs. Namely, from the existence of optimal proof systems we infer the deduction property for some extension of EF, and this in turn implies the existence of complete NP-pairs. Conversely, it is not clear whether any of the opposite implications is valid. While the deduction property for EF relates to the completeness of the canonical EF-pair, we also exhibit two conditions that imply the completeness of the canonical pair of Frege systems. In particular, we demonstrate that the existence of complete NP-pairs is tightly connected with the question whether EF is indeed more powerful than ordinary Frege systems.

The paper is organized as follows. In Sect. 2 we provide some background information on propositional proof systems and disjoint NP-pairs. In Sect. 3 we define and discuss different versions of the deduction property. Section 4, after a series of lemmas, contains the main results connecting the deduction property for strong systems with the existence of complete NP-pairs. Finally, in Sect. 5 we conclude with some open problems.

2 Preliminaries

Propositional Proof Systems. Propositional proof systems were defined in a very general way by Cook and Reckhow in [CR79] as polynomial time functions P which have as its range the set of all tautologies. A string π with $P(\pi) = \varphi$ is called a P-proof of the tautology φ. By $P \vdash_{\leq m} \varphi$ we indicate that there is a P-proof of φ of size $\leq m$.

Proof systems are compared according to their strength by simulations introduced in [CR79]. A proof system Q *p-simulates* a proof system P (denoted by $P \leq_p Q$), if there exists a function that computes in polynomial time from a P-proof a Q-proof of the same formula. A proof system is called *p-optimal* if it p-simulates all proof systems. Whether or not p-optimal proof systems exist is an open problem posed by Krajíček and Pudlák [KP89].

A prominent example of a class of proof systems is provided by *Frege systems* which are usual textbook proof systems based on axioms and rules. In the context of propositional proof complexity these systems were first studied by Cook and Reckhow [CR79] and it was proven there that all Frege systems, i.e., systems using different axiomatizations and rules, are polynomially equivalent.

Augmenting Frege systems by the possibility to abbreviate complex formulas by propositional variables we arrive at the *extended Frege proof system EF*. This extension rule might further reduce the proof size, but it is not known whether EF is really stronger than ordinary Frege systems. Both Frege and the extended

Frege system are very strong systems for which no non-trivial lower bounds to the proof size are currently known.

Another way to enhance the power of Frege systems is to allow substitutions not only for axioms but also for all formulas that have been derived in Frege proofs. Augmenting Frege systems by this substitution rule leads to the *substitution Frege system* SF. The extensions EF and SF were introduced by Cook and Reckhow [CR79]. While it was already proven there that EF is simulated by SF, the converse simulation is considerably more involved and was shown independently by Dowd [Dow85] and Krajíček and Pudlák [KP89]. For more detailed information on Frege systems and its extensions we refer to the monograph [Kra95].

Under the notion of *line based proof systems* we subsume all proof systems that have as proofs sequences of formulas, and formulas in such a sequence are derived from earlier formulas in the sequence by the rules available in the proof system. In particular, Frege systems and its extensions are line based in this sense. Line based proof systems P can be enhanced by additional axioms in two different ways. Namely, we can form a proof system $P + \Phi$ augmenting P by a polynomial time computable set Φ of tautologies as new axiom schemes. This means that formulas from Φ as well as substitution instances of these formulas can be freely introduced as new lines in $P+\Phi$ -proofs. In contrast to this we use the notation $P \cup \Phi$ for the proof system that extends P only by formulas from Φ but not by their substitution instances as new axioms. In our applications the set Φ will mostly be *printable*, meaning that it is even possible to generate the formulas from Φ in polynomial time.

Disjoint NP-Pairs. A pair (A, B) is called a *disjoint* NP*-pair* if $A, B \in \mathsf{NP}$ and $A \cap B = \emptyset$. Grollmann and Selman [GS88] defined the following reduction between disjoint NP-pairs (A, B) and (C, D): $(A, B) \leq_p (C, D)$ if there exists a polynomial time computable function f such that $f(A) \subseteq C$ and $f(B) \subseteq D$.

The link between disjoint NP-pairs and propositional proof systems was established by Razborov [Raz94], who associated a canonical disjoint NP-pair $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ with a proof system P, where the first component $\mathrm{Ref}(P) = \{(\varphi, 1^m) \,|\, P \vdash_{\leq m} \varphi\}$ contains information about proof lengths in P and $\mathrm{SAT}^* = \{(\varphi, 1^m) \mid \neg\varphi \in \mathrm{SAT}\}$ is a padded version of SAT. This canonical pair is linked to the automatizablility and the reflection property of the proof system [Pud03]. Simulations between proof systems are reflected in reductions between canonical pairs as the next easy, but useful proposition shows:

Proposition 1 (Pudlák [Pud03]). *If P and Q are proof systems with $P \leq_p Q$, then the canonical pair of P is $\leq_p$-reducible to the canonical pair of Q.*

Proof. Let f compute the simulation of P by Q. Then the reduction is given by $(\varphi, 1^m) \mapsto (\varphi, 1^{p(m)})$ where p is a polynomial bounding the running time of f. □

More information on the connection between disjoint NP-pairs and propositional proof systems can be found in [Pud03,Bey04,Bey06a,GSZ06].

3 Deduction Properties for Frege Systems

The deduction theorem of propositional logic states that in a Frege system F a formula ψ is provable from a formula φ if and only if $\varphi \to \psi$ is provable in F. Because proof complexity is focusing on the length of proofs it is interesting to analyse how the proof length is changing in the deduction theorem. An F-proof of $\varphi \to \psi$ together with the axiom φ immediately yields the formula ψ with one application of modus ponens. Therefore it is only interesting to ask for the increase in proof length when constructing a proof of $\varphi \to \psi$ from an F-proof of ψ with the extra axiom φ. This was analysed in detail in [Bon93,BB93].

The main application of the deduction property is to simplify proofs of complex formulas. Namely, to prove an implication $\varphi \to \psi$ it suffices to construct a proof of ψ from φ. In particular, φ can be any formula and is not necessarily a tautology. It is clear that such a deduction property is doomed to fail for strong systems like EF or SF that can immediately produce substitution instances from φ. For instance, by one application of the substitution rule we get $SF \cup \{p\} \vdash q$, whereas $p \to q$ is not even a tautology.

Aiming in particular at such strong proof systems we therefore restrict φ to tautologies and make the following general definition.

Definition 2. *A line based proof system P allows* efficient deduction *if there exists a polynomial p such that for all finite sets Φ of tautologies $P \cup \Phi \vdash_{\leq m} \psi$ implies $P \vdash_{\leq p(m+m')} (\bigwedge_{\varphi \in \Phi} \varphi) \to \psi$ where $m' = |\bigwedge_{\varphi \in \Phi} \varphi|$.*

This efficient deduction property is known to hold for Frege systems (cf. [BB93]):

Theorem 3 (Deduction theorem for Frege systems). *Every Frege system F allows efficient deduction.*

Proof. For every F-rule

$$R_i = \frac{\psi_1 \quad \dots \quad \psi_r}{\psi}$$

we fix an F-proof π_i of the tautology $((q \to \psi_1) \wedge \dots \wedge (q \to \psi_r)) \to (q \to \psi)$. In particular, for $r = 0$ this also includes the case that R_i is an axiom scheme.

Let $\varphi_1, \dots, \varphi_n$ be tautologies and let $(\theta_1, \dots, \theta_k)$ be a proof of ψ of size m in the system $F \cup \{\varphi_1, \dots, \varphi_n\}$. Let $m' = \sum_{i=1}^{n} |\varphi_i|$. By induction on j we construct proofs of the implications

$$(\bigwedge_{i=1}^{n} \varphi_i) \to \theta_j \ .$$

We distinguish two cases on how the formula θ_j was derived.

If θ_j is one of the formulas from $\{\varphi_1, \dots, \varphi_n\}$, then we get $(\bigwedge_{i=1}^{n} \varphi_i) \to \theta_j$ in a proof of size $O(m')$.

If θ_j was inferred from $\theta_{j_1}, \dots, \theta_{j_r}$ by the F-rule R_i, then we can get from π_i an F-proof of size $O(m' + |\theta_j| + \sum_{l=1}^{r} |\theta_{j_l}|)$ of the tautology

$$(((\bigwedge_{i=1}^{n} \varphi_i) \to \theta_{j_1}) \wedge \dots \wedge ((\bigwedge_{i=1}^{n} \varphi_i) \to \theta_{j_r})) \to ((\bigwedge_{i=1}^{n} \varphi_i) \to \theta_j) \ .$$

Combining all the earlier proved implications $(\bigwedge_{i=1}^{n} \varphi_i) \to \theta_{j_l}$, $l = 1, \dots, r$ by conjunctions and using modus ponens we get the desired implication $(\bigwedge_{i=1}^{n} \varphi_i) \to \theta_j$ in a proof of size $O(m + m')$. □

A still weaker form of the deduction property is given in the next definition.

Definition 4. *A line based proof system P allows* weak deduction *if the following condition holds. For all printable sets $\Phi \subseteq \mathrm{TAUT}$ there exists a polynomial p such that for all finite subsets $\Phi_0 \subseteq \Phi$ we can infer from $P \cup \Phi_0 \vdash_{\leq m} \psi$ that $P \vdash_{\leq p(m+m')} (\bigwedge_{\varphi \in \Phi_0} \varphi) \to \psi$ where $m' = |\bigwedge_{\varphi \in \Phi_0} \varphi|$.*

In Definition 2 we allowed a fixed polynomial increase for the proof size in the transformation of a proof from ψ to the implication $(\bigwedge_{\varphi \in \Phi_0} \varphi) \to \psi$, whereas in the weak deduction property this polynomial might depend on the choice of the extra axioms Φ. This weakening of the deduction property allows us to show the following proposition.

Proposition 5. *Optimal line based proof systems have the weak deduction property.*

Proof. Let P be an optimal line based proof system and let Φ be a printable set of tautologies. Then $P \cup \Phi$ is a well defined proof system which by the optimality of P is simulated by P. Hence we have polynomial size P-proofs of all formulas from Φ. Given a finite set Φ_0 and a $P \cup \Phi_0$-proof π of a formula ψ we can therefore first derive all formulas from Φ_0 in polynomial size P-proofs and concatenate this with π. This results in a polynomial size P-proof of ψ from which we easily obtain a polynomial size P-proof of $\bigwedge_{\varphi \in \Phi_0} \varphi \to \psi$. □

It is, however, not clear if optimal proof system also have the seemingly stronger efficient deduction property.

4 Is the Canonical Pair of *EF* Complete?

In this section the proof systems $EF \cup \Phi$ for polynomial time computable sets Φ of tautologies will play an important role. As explained earlier, this notation means that in contrast to the usual axiom schemes of EF we are only allowed to use formulas from Φ but not their substitution instances in $EF \cup \Phi$-proofs.

For substitution Frege systems such a restriction does not make sense as the substitution rule immediately allows to produce all substitution instances of Φ. However, we make the following different and somewhat technical definition. Let $\Phi(\bar{p})$ be again a polynomial time computable set of tautologies in the possibly infinite sequence of variables $\bar{p}$. Let $\bar{q}$ be another infinite sequence of propositional variables such that $\bar{p}$ and $\bar{q}$ do not share any common elements. Then we denote by $SF \cup^{\bar{q}} \Phi(\bar{p})$ the following proof system P. The system P is the substitution Frege system SF augmented by the additional axioms Φ, but with the following restriction: each P-proof may only use the extra axioms from Φ, if the variables $\bar{q}$ do not appear in the P-proof.

Our first lemma shows that every disjoint NP-pair is reducible to the canonical pair of such an extension of SF.

Lemma 6. *For every disjoint* NP*-pair* (A, B) *there exists a polynomial time constructible sequence of formulas* $\varphi_n(\bar{p})$ *in the variables* $\bar{p}$*, such that different formulas* φ_n *do not share any variables. Let* $\bar{q}$ *be another infinite sequence of variables without any elements from* $\bar{p}$*. Then* (A, B) *is* $\leq_p$*-reducible to the canonical pair of the proof system* $SF \cup^{\bar{q}} \{\varphi_n(\bar{p}) \mid n \geq 0\}$*.*

Proof. Let (A, B) be a disjoint NP-pair. Similarly as in Cook's proof of the NP-completeness of SAT [Coo71], we can construct in polynomial time propositional formulas $\psi_n(\bar{x}, \bar{y})$ such that $\psi_n(\bar{a}, \bar{y})$ is satisfiable if and only if $\bar{a} \in A$. Similarly, we build such propositional formulas $\theta_n(\bar{x}, \bar{z})$ for B. We choose the variables of $\psi_n(\bar{x}, \bar{y})$ and $\theta_n(\bar{x}, \bar{z})$ in such a way that the input variables $\bar{x}$ are the common variables of ψ_n and θ_n, and the auxiliary variables $\bar{y}$ and $\bar{z}$ are distinct. Moreover, the variables of the formulas ψ_n and θ_n are chosen distinct for different n, and all variables $\bar{x}$, $\bar{y}$, and $\bar{z}$ are contained in the sequence of variables $\bar{p}$ which are distinct from the variables $\bar{q}$. We define the sequence φ_n as

$$\varphi_n = \psi_n(\bar{x}, \bar{y}) \rightarrow \neg\theta_n(\bar{x}, \bar{z}) \enspace .$$

Let P denote the system $SF \cup^{\bar{q}} \{\varphi_n(\bar{p}) \mid n \geq 0\}$. We claim that the reduction from (A, B) to $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ is given by

$$a \ \mapsto \ (\neg\theta_{|a|}(\bar{a}, \bar{z}), 1^{p(|a|)})$$

for some suitable polynomial p. To see the correctness of the reduction let first a be an element from A of length n. As ψ_n represents A there exists a witness $\bar{b}$ such that $\psi_n(\bar{a}, \bar{b})$ is a tautological formula. The P-proof of $\neg\theta_n(\bar{a}, \bar{z})$ proceeds as follows. First we use the axiom $\psi_n(\bar{x}, \bar{y}) \rightarrow \neg\theta_n(\bar{x}, \bar{z})$ and substitute the variables $\bar{x}$ and $\bar{y}$ by $\bar{a}$ and $\bar{b}$, respectively, obtaining

$$\psi_n(\bar{a}, \bar{b}) \rightarrow \neg\theta_n(\bar{a}, \bar{z}) \enspace .$$

As $\psi_n(\bar{a}, \bar{b})$ is a true propositional formula without variables we can provide a polynomial size Frege proof for it. An application of modus ponens gives a P-proof of $\neg\theta_n(\bar{a}, \bar{z})$ as desired.

Assume now $a \in B$. Then $\neg\neg\theta_{|a|}(\bar{a}, \bar{z}) = \theta_{|a|}(\bar{a}, \bar{z})$ is satisfiable and hence $(\neg\theta_{|a|}(\bar{a}, \bar{z}), 1^{p(|a|)}) \in \mathrm{SAT}^*$. □

The next lemma is an extension of the simulation of SF by EF as proved in [KP89].

Lemma 7. *Let* Φ *be a set of tautologies that is polynomial time decidable. Assume that* Φ *only contains the variables* $\bar{p}$ *such that different formulas from* Φ *use distinct variables from* $\bar{p}$*. Let further* $\bar{q}$ *be an infinite sequence of variables that is disjoint from* $\bar{p}$*. Then the system* $SF \cup^{\bar{q}} \Phi(\bar{p})$ *is p-simulated by the system* $EF \cup \Phi(\bar{q})$*.*

Proof. Let $\pi = (\varphi_1, \ldots, \varphi_k)$ be a proof of φ_k in the system $SF \cup^{\bar{q}} \Phi(\bar{p})$. If π uses the variables $\bar{q}$, then π is an ordinary SF-proof which can be translated into a polynomially longer EF-proof. If π proves a formula $\varphi(\bar{p})$ from $\Phi(\bar{p})$, then we construct an $EF \cup \Phi(\bar{q})$-proof of $\varphi(\bar{p})$ as follows. First we introduce all variables $p_1, \ldots, p_m$ occurring in $\varphi(\bar{p})$ as extension variables via the rules $p_i \leftrightarrow q_i$ for $i = 1, \ldots, m$. Using the formulas $p_i \leftrightarrow q_i$ we then prove the equivalence $\varphi(\bar{p}) \leftrightarrow \varphi(\bar{q})$ by induction on the logical complexity of φ. Finally, we include the axiom $\varphi(\bar{q})$ from $\Phi(\bar{q})$ and derive $\varphi(\bar{p})$ by modus ponens.

The general case follows the same paradigm as the last construction but is technically more involved. Without loss of generality we may now assume that the only axioms from Φ in the proof π are the first n formulas $\varphi_1(\bar{p}), \ldots, \varphi_n(\bar{p})$, and that these $n < k$ formulas are all distinct. Moreover, we may assume that the variables $\bar{q}$ do not occur in π and that apart from the axioms the system $SF \cup^{\bar{q}} \Phi(\bar{p})$ only uses the substitution rule and modus ponens.

Similarly as in [KP89] we will now construct a proof of φ_k in the system $EF \cup \Phi(\bar{q})$. Let $\bar{r}$ be all variables occurring in π. We choose tuples of mutually distinct variables $\bar{q}_1, \ldots, \bar{q}_k$ that have the same length as $\bar{r}$. For $\bar{q}_1, \ldots, \bar{q}_n$ we choose those variables from $\bar{q}$ that correspond to the respective variables from $\bar{p}$ in $\varphi_1(\bar{p}), \ldots, \varphi_n(\bar{p})$. By hypothesis all formulas from Φ use different variables, and therefore $\bar{q}_1, \ldots, \bar{q}_n$ are pairwise distinct. For $\bar{q}_k$ we choose the variables $\bar{r}$. We denote the formulas $\varphi_j(\bar{q}_j)$ by ψ_j for $j = 1, \ldots, k$. Now we define tuples of formulas $\bar{\beta}_j$, $j = 1, \ldots, k$, as follows:

$$\bar{\beta}_j = \begin{cases} \bar{q}_j & \text{if } \varphi_j \text{ is an axiom or has been derived by modus ponens,} \\ \alpha(\bar{q}_j) & \text{if } \varphi_j \text{ was derived from } \varphi_i,\ i < j, \text{ by the substitution } \alpha. \end{cases}$$

Here $\alpha(\bar{q}_j)$ means that the variables $\bar{q}_j$ are substituted by α in the same way as the variables $\bar{r}$ in φ_i.

We further use the abbreviations $\Psi_{i,j} = \psi_i \wedge \ldots \wedge \psi_j$ with the convention $\Psi_{i+1,i} = 1$. The $EF \cup \Phi(\bar{q})$-proof starts with the following applications of the extension rule

$$q_{i,l} \leftrightarrow (\Psi_{i+1,i} \wedge \neg\psi_{i+1} \wedge \beta_{i+1,l}) \vee \ldots \vee (\Psi_{i+1,k-1} \wedge \neg\psi_k \wedge \beta_{k,l})$$

for $i = k-1, \ldots, 1$ and $l = 1, \ldots, |\bar{r}|$. In particular, all variables in the new proof, except for $\bar{q}_k = \bar{r}$ in ψ_k, are extension variables. Apparently, the formulas

$$\Psi_{i+1,j-1} \wedge \neg\psi_j \rightarrow (q_{i,l} \leftrightarrow \beta_{j,l}) \ , \quad i < j$$

are tautologies. Therefore also the formulas

$$\Psi_{i+1,j-1} \wedge \neg\psi_j \rightarrow (\varphi_i(\bar{q}_i) \leftrightarrow \varphi_i(\bar{\beta}_j)) \ , \quad i < j$$

are tautological. Moreover, by induction on the formulas φ_i it can be shown that the above formulas admit polynomial size Frege proofs. By definition these formulas can also be written as

$$\Psi_{i+1,j-1} \wedge \neg\psi_j \rightarrow (\psi_i \leftrightarrow \psi_i(\bar{\beta}_j)) \ , \quad i < j \ . \tag{1}$$

Once these formulas have been derived the $EF \cup \Phi(\bar{q})$-proof proceeds by successively deriving the formulas $\psi_1, \ldots, \psi_k$, thus yielding with $\psi_k = \varphi_k$ the desired formula. Proving the formulas ψ_j for $j = 1, \ldots, k$ is done by induction on j. We have to distinguish three cases on how φ_j was derived.

If φ_j is an axiom, then this is also true for ψ_j. In particular, this holds for the formulas $\psi_1(\bar{q}_1), \dots, \psi_n(\bar{q}_n)$, as these are formulas from $\Phi(\bar{q})$.

For the second case we assume that φ_j was derived by modus ponens from formulas φ_u and $\varphi_v = \varphi_u \to \varphi_j$ with $u, v < j$. By Formula 1 above we have

$$\Psi_{u+1,j-1} \wedge \neg\psi_j \to (\psi_u \leftrightarrow \psi_u(\bar{\beta}_j)) \ .$$

By induction hypothesis we have already derived $\Psi_{u+1,j-1}$ and therefore obtain

$$\neg\psi_j \to (\psi_u \leftrightarrow \psi_u(\bar{\beta}_j)) \ .$$

Similarly, we get $\neg\psi_j \to (\psi_v \leftrightarrow \psi_v(\bar{\beta}_j))$. As ψ_u and ψ_v have already been derived by induction hypothesis we get

$$\neg\psi_j \to \psi_u(\bar{\beta}_j) \wedge \psi_v(\bar{\beta}_j) \ ,$$

i.e., $\neg\psi_j \to \psi_u(\bar{\beta}_j) \wedge (\psi_u(\bar{\beta}_j) \to \psi_j(\bar{\beta}_j))$. Applying modus ponens gives

$$\neg\psi_j \to \psi_j(\bar{\beta}_j)$$

which is by definition $\neg\psi_j \to \psi_j$. Hence we get the formula ψ_j.

For the last case assume that φ_j is obtained by the substitution α from φ_i for some $i < j$. Formula 1 yields

$$\Psi_{i+1,j-1} \wedge \neg\psi_j \to (\psi_i \leftrightarrow \psi_i(\bar{\beta}_j)) \ .$$

By induction hypothesis we have both $\Psi_{i+1,j-1}$ and ψ_i which gives

$$\neg\psi_j \to \psi_i(\bar{\beta}_j) \ .$$

But now we have by definition

$$\psi_i(\bar{\beta}_j) = \psi_i(\alpha(\bar{q}_j)) = \psi_j(\bar{q}_j) = \psi_j \ ,$$

hence we get again $\neg\psi_j \to \psi_j$ and therefore the formula ψ_j. □

Augmenting line based proof systems P by additional axioms Φ will usually enhance the power of the proof system. The following lemma shows, however, that if P has the weak deduction property, then the canonical pair of $P \cup \Phi$ will not be more difficult than the canonical P-pair. In particular, combined with Theorem 3 the next lemma shows that the canonical pairs of F and its extensions $F \cup \Phi$ are equivalent for printable sets $\Phi \subseteq \mathrm{TAUT}$.

Lemma 8. *Let Φ be a printable set of tautologies and let P be a proof system with the weak deduction property. Then* $(\mathrm{Ref}(P \cup \Phi), \mathrm{SAT}^*) \leq_p (\mathrm{Ref}(P), \mathrm{SAT}^*)$.

Proof. Let Φ be printable and let p be the polynomial from the weak deduction property for P and Φ. Because Φ is printable there exists a polynomial q such that for each number m the set Φ contains at most $q(m)$ tautologies of length $\leq m$. Let $\Phi_m = \Phi \cap \Sigma^{\leq m}$ be the set of these tautologies.

Then $(\mathrm{Ref}(P \cup \Phi), \mathrm{SAT}^*)$ reduces to $(\mathrm{Ref}(P), \mathrm{SAT}^*)$ via the function

$$(\psi, 1^m) \ \mapsto \ (\,(\bigwedge_{\varphi \in \Phi_m} \varphi) \to \psi, 1^{p(mq(m)+m)}) \ .$$

To verify the claim assume that $(\psi, 1^m) \in \mathrm{Ref}(P \cup \Phi)$. Let π be a $P \cup \Phi$-proof of ψ of length $\leq m$. This proof π can use only formulas of length $\leq m$ from Φ of which there are only $\leq q(m)$ many. Hence the tautologies used in the proof π are contained in $\bigwedge_{\varphi \in \Phi_m} \varphi$. Therefore we know that π is also a proof for ψ in the proof system $P \cup \Phi_m$. Using the weak deduction property of P we get a P-proof of size $\leq p(mq(m)+m)$ of $(\bigwedge_{\varphi \in \Phi_m} \varphi) \to \psi$.

Now assume $(\psi, 1^m) \in \mathrm{SAT}^*$. Then $\neg\psi$ is satisfiable and therefore

$$\neg((\bigwedge_{\varphi \in \Phi_m} \varphi) \to \psi) \ = \ (\bigwedge_{\varphi \in \Phi_m} \varphi) \wedge \neg\psi$$

is also satisfiable because $(\bigwedge_{\varphi \in \Phi_m} \varphi)$ is a tautology. □

Combining the above lemmas we can now prove our main results.

Theorem 9. *If EF satisfies the weak deduction property, then the canonical pair of EF is $\leq_p$-complete for the class of all disjoint* NP*-pairs.*

Proof. Let (A, B) be a disjoint NP-pair and let $\varphi_n(\bar{p})$ be the polynomial time constructible sequence in the variables $\bar{p}$ which is guaranteed by Lemma 6. Let $\Phi(\bar{p})$ denote the set $\{\varphi_n(\bar{p}) \mid n \geq 0\}$. By Lemma 6 we have

$$(A, B) \leq_p (\mathrm{Ref}(SF \cup^{\bar{q}} \Phi(\bar{p})), \mathrm{SAT}^*) \ .$$

As $\Phi(\bar{p})$ uses only the variables $\bar{p}$ and avoids the variables $\bar{q}$ we get by Lemma 7

$$SF \cup^{\bar{q}} \Phi(\bar{p}) \leq_p EF \cup \Phi(\bar{q}) \ .$$

By Proposition 1 this implies that (A, B) is $\leq_p$-reducible to the canonical pair of $EF \cup \Phi(\bar{q})$. Finally, we use the weak deduction property of EF to reduce the canonical pair of $EF \cup \Phi(\bar{q})$ to the canonical pair of EF. By combining the last two reductions we have reduced (A, B) to $(\mathrm{Ref}(EF), \mathrm{SAT}^*)$. □

As we know that every proof system P is simulated by a proof system of the form $EF + \Phi$ with printable $\Phi \subset \mathrm{TAUT}$ (for instance we can take Φ as translations of the reflection principle of P), we can easily adapt the above arguments to prove a somewhat more general version of Theorem 9, thus demonstrating the importance of the question whether EF or its extensions satisfy the weak deduction property.

Theorem 10. *1. If optimal proof systems exist, then $EF + \Phi$ has the weak deduction property for some printable set $\Phi \subset \mathrm{TAUT}$.*
2. If $EF + \Phi$ has the weak deduction property for some printable $\Phi \subset \mathrm{TAUT}$, then complete disjoint NP*-pairs exist.*

Now we will exhibit two conditions which lead to the stronger consequence of the completeness of the canonical pair of Frege systems. Although these conditions are seemingly unrelated with the deduction property, the results follow easily from the chain of the above lemmas. In the next theorem we will show that the existence of complete NP-pairs is tightly connected with the question whether F and EF are indeed proof systems of different strength.

Theorem 11. *Assume that for all printable sequences Φ of tautologies the proof systems $F \cup \Phi$ and $EF \cup \Phi$ are equivalent. Then the canonical pair of the Frege proof system is complete for the class of all disjoint NP-pairs.*

Proof. By Lemmas 6 and 7 we can reduce every NP-pair to the canonical pair of a proof system $EF \cup \Phi$ with printable $\Phi \subset \mathrm{TAUT}$. If $EF \cup \Phi \leq F \cup \Phi$, then EF has efficient deduction, by which the result follows with Lemma 8. □

The next theorem asks, in principle, whether the systems $F \cup \Phi$ and $F + \Phi$ are equivalent.

Theorem 12. *Assume that for all printable sets of tautologies Φ the system $F \cup \Phi$ is closed under substitutions by constants. Then the canonical F-pair is a complete disjoint NP-pair.*

Proof. Assume that all systems $F \cup \Phi$ are closed under substitutions by constants, i.e., from an $F \cup \Phi$-proof of some formula $\varphi(\bar{x}, \bar{y})$ we can construct an $F \cup \Phi$-proof of $\varphi(\bar{x}, \bar{a})$ where some variables $\bar{y}$ of φ are substituted by the constants $\bar{a}$. Then we can reduce every disjoint NP-pair to the canonical pair of such a proof system, analogously as in Lemma 6. Together with the deduction theorem for F and Lemma 8 this yields the result. □

5 Conclusion

From Proposition 5 it follows that the optimality of EF implies the weak deduction property for EF. In turn, this weak deduction property for EF gives us a complete disjoint NP-pair in form of the canonical pair of EF. These results show the importance of the following problem:

Problem 13. Does EF have the weak or even the efficient deduction property[1]?

Given the implications above, we expect, however, that neither proving nor disproving this question will be an easy task.

A hopefully more accessible question is to determine whether the deduction property is robust in the sense that it is preserved inside a degree of equivalent proof systems. A positive answer would imply, for instance, that deduction for EF implies deduction for SF and vice versa. It would also allow us to weaken the hypothesis of Theorem 11 to $F \equiv EF$.

[1] In [Bey06a] I claimed that efficient deduction holds for EF. The proof that I had in mind was an easy modification of the proof of the deduction theorem for Frege systems. This, however, works only if the $EF \cup \Phi$-proof does not contain any applications of the extension rule involving the variables of Φ (as it does e.g. in the proof of Lemma 7).

References

[BB93] Maria Luisa Bonet and Samuel R. Buss. The deduction rule and linear and near-linear proof simulations. *The Journal of Symbolic Logic*, **58**(2):688–709, 1993.

[Bey04] Olaf Beyersdorff. Representable disjoint NP-pairs. In *Proc. 24th Conference on Foundations of Software Technology and Theoretical Computer Science*, Lecture Notes in Computer Science #3328, 122–134. Springer-Verlag, Berlin Heidelberg, 2004.

[Bey06a] Olaf Beyersdorff. Disjoint NP-pairs from propositional proof systems. In *Proc. 3rd Conference on Theory and Applications of Models of Computation*, Lecture Notes in Computer Science #3959, 236–247. Springer-Verlag, Berlin Heidelberg, 2006.

[Bey06b] Olaf Beyersdorff. Tuples of disjoint NP-sets. In *Proc. 1st International Computer Science Symposium in Russia*, Lecture Notes in Computer Science #3967, 80–91. Springer-Verlag, Berlin Heidelberg, 2006.

[Bon93] Maria Luisa Bonet. Number of symbols in Frege proofs with and without the deduction rule. In P. Clote and J. Krajíček, editors, *Arithmetic, Proof Theory and Computational Complexity*, 61–95. Oxford University Press, Oxford, 1993.

[Coo71] Stephen A. Cook. The complexity of theorem proving procedures. In *Proc. 3rd Annual ACM Symposium on Theory of Computing*, 151–158, 1971.

[CR79] Stephen A. Cook and Robert A. Reckhow. The relative efficiency of propositional proof systems. *The Journal of Symbolic Logic*, **44**:36–50, 1979.

[Dow85] Martin Dowd. Model-theoretic aspects of P≠NP. Unpublished manuscript, 1985.

[GS88] Joachim Grollmann and Alan L. Selman. Complexity measures for public-key cryptosystems. *SIAM Journal on Computing*, **17**(2):309–335, 1988.

[GSS05] Christian Glasser, Alan L. Selman, and Samik Sengupta. Reductions between disjoint NP-pairs. *Information and Computation*, **200**(2):247–267, 2005.

[GSSZ04] Christian Glasser, Alan L. Selman, Samik Sengupta, and Liyu Zhang. Disjoint NP-pairs. *SIAM Journal on Computing*, **33**(6):1369–1416, 2004.

[GSZ06] Christian Glasser, Alan L. Selman, and Liyu Zhang. Survey of disjoint NP-pairs and relations to propositional proof systems. In Oded Reingold, Arnold L. Rosenberg, and Alan L. Selman, editors, *Essays in Theoretical Computer Science in Memory of Shimon Even*, 241–253. Springer-Verlag, Berlin Heidelberg, 2006.

[HS92] Steven Homer and Alan L. Selman. Oracles for structural properties: The isomorphism problem and public-key cryptography. *Journal of Computer and System Sciences*, **44**(2):287–301, 1992.

[KP89] Jan Krajíček and Pavel Pudlák. Propositional proof systems, the consistency of first order theories and the complexity of computations. *The Journal of Symbolic Logic*, **54**:1963–1079, 1989.

[Kra95] Jan Krajíček. *Bounded Arithmetic, Propositional Logic, and Complexity Theory*, Encyclopedia of Mathematics and Its Applications #60. Cambridge University Press, Cambridge, 1995.

[Kra04] Jan Krajíček. Dual weak pigeonhole principle, pseudo-surjective functions, and provability of circuit lower bounds. *The Journal of Symbolic Logic*, **69**(1):265–286, 2004.

[Pud03] Pavel Pudlák. On reducibility and symmetry of disjoint NP-pairs. *Theoretical Computer Science*, **295**:323–339, 2003.

[Raz94] Alexander A. Razborov. On provably disjoint NP-pairs. Technical Report TR94-006, Electronic Colloquium on Computational Complexity, 1994.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de/