# Classical Logic is better than Intuitionistic Logic: A Conjecture about Double-Negation Translations

Christian Urban and Diana Ratiu

{urban,ratiu}@math.lmu.de
University of Munich

**Abstract.** It is well-known that in terms of consistency classical logic and intuitionistic logic have equal strength: every intuitionistic proof is a classical proof and every classical proof can be embedded into intuitionistic logic via double negation translations. It is also well-known that intuitionistic proofs contain witnesses for existential statements, which is not always the case with classical proofs. However, here we study classical and intuitionistic logic as reduction systems. From this perspective, we conjecture that classical logic is more powerful than intuitionistic logic. The conjecture links double-negation translations to the colour-protocol introduced by Danos et al. for cut-elimination in classical logic. If the conjecture turns out to be true, then we can conclude that *not all* cut-reductions can be simulated by double-negation translations.

## 1 Introduction

Since the works on double-negation translations by Gentzen, Gödel and Kolmogorov, one knows that classical logic and intuitionistic logic have equal strength in terms of consistency: intuitionistic sequent-proofs can be seen as classical proofs where the right-hand side of the sequents is restricted to maximal one formula and every classical sequent-proof can be embedded into intuitionistic logic via double-negation translations. As a result, consistency of one logic implies consistency of the other. In this paper, however, we focus on the correspondence of intuitionistic and classical logic with respect to term-rewriting, proof-normalisation and cut-elimination.

According to the Curry-Howard correspondence, the simply-typed lambda-calculus can be viewed as a term-assignment for intuitionistic proofs formalised in Gentzen's natural deduction calculus NJ. Term-rewriting in the simply-typed lambda-calculus is a form of computation that converts a term to its simplest form, analogous to symbolic evaluation. On the other hand, normalisation is a method for eliminating certain redundancies in proofs. Applied iteratively, it transforms a proof to one in normalform. Using the Curry-Howard correspondence we see that the two notions coincide and consequently we can talk of a computational interpretation of intuitionistic proofs.

Two properties hold for term-rewriting in the simply-typed lambda-calculus and by the Curry-Howard correspondence also for normalisation in NJ: they are strongly normalising and Church-Rosser. With some limitations [8, 15, 18, 23], the Curry-Howard correspondence applies also to the sequent-calculus LJ and to the process of cut-elimination. For sake of simplicity, we shall ignore these limitations here and regard cut-elimination in intuitionistic logic as strongly normalising and as having (morally at least) the Church-Rosser property. This ignorance can be partially justified if one sees behind every LJ-proof an NJ-proof where the inessential differences present in sequent proofs “disappear”, and sees cut-elimination as an approximation of proof-normalisation.

Although it has been shown that some cut-elimination procdures for classical logic are strongly normalising as well, cut-elimination in classical logic is, given a sensible notion of cut-reductions, *not* Church-Rosser—not even morally. The lack of Church-Rosser in classical logic is a main theme running through the works [2, 3, 19, 22], which analyse what this means from a computational point of view. For example in [19] it has been shown that a lambda-calculus with a non-deterministic choice-operator can be embedded into a fragment of classical logic. A simple classical proof taken from [6, 9] shall illustrate the lack of Church-Rosser:

$$\dfrac{\dfrac{\dfrac{A\vdash A \quad A\vdash A}{A\vee A\vdash A,A}\,\vee_L}{A\vee A\vdash A}\,contr_R \qquad \dfrac{\dfrac{A\vdash A\quad A\vdash A}{A,A\vdash A\wedge A}\,\wedge_R}{A\vdash A\wedge A}\,contr_L}{A\vee A\vdash A\wedge A}\,cut \tag{1}$$

The cut in this proof can be eliminated by reducing it to one of the following two normalforms

$$\dfrac{\dfrac{\dfrac{\dfrac{A\vdash A \quad A\vdash A}{A,A\vdash A\wedge A}\,\wedge_R}{A\vdash A\wedge A}\,contr_L \qquad \dfrac{\dfrac{A\vdash A\quad A\vdash A}{A,A\vdash A\wedge A}\,\wedge_R}{A\vdash A\wedge A}\,contr_L}{A\vee A\vdash A\wedge A, A\wedge A}\,\vee_L}{A\vee A\vdash A\wedge A}\,contr_R \tag{2}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{A\vdash A \quad A\vdash A}{A\vee A\vdash A,A}\,\vee_L}{A\vee A\vdash A}\,contr_R \qquad \dfrac{\dfrac{A\vdash A\quad A\vdash A}{A\vee A\vdash A,A}\,\vee_L}{A\vee A\vdash A}\,contr_R}{A\vee A, A\vee A\vdash A\wedge A}\,\wedge_R}{A\vee A\vdash A\wedge A}\,contr_L \tag{3}$$

which are obtained, respectively, by either permuting the cut to the left over the $contr_R$-rule or to the right over the $contr_L$-rule. Another example showing that in classical logic one can, in general, reach more than one normalform is Lafont’s proof [10, Page 151].

In light of the absence of the Church-Rosser property for cut-elimination in classical logic and in light of the work on double-negation translations, there seem to be obvious questions: What is the correspondence between cut-elimination in classical logic and the embeddings of classical proofs into intuitionistic logic via double-negation translations? Since cut-elimination in intuitionistic logic is Church-Rosser,[1] which restriction is tacitly enforced by a double-negation translation so that eliminating cuts in the double-negated version of a classical proof leads to only a single normalform? Or

[1] As mentioned earlier we ignore what we believe to be superficial variations between different normalforms reachable from an intuitionistic sequent-proof. However see [18] for a more thorough analysis of this aspect.

more concisely asked, do double-negation translations correspond to particular strategies of how to eliminate cuts? Does every double-negation translation lead to the same normalform (by *same* we mean corresponding to one particular normalform obtained by cut-elimination in classical logic)? If not, then can one find for every normalform of a classical proof a corresponding double-negation translation that will produce the double-negated version of this normalform—that is, can every reduction sequence in classical logic be simulated by a (probably carefully chosen) double negation translation and performing cut-elimination in intuitionistic logic? Are there any double-negation translations that lead to normalforms that have no equivalent amongst the normalforms reachable by cut-elimination in classical logic? Can one characterise somehow, which normalforms can be reached by double-negation translations and which can not? In this paper we conjecture answers for all these questions.

Although some special cases seem to be answered by existing work, for example [5, 6, 14], we are unaware of any work that treat these questions in full generality. The answers we shall give to these questions are a lot inspired by the comments made in [6, Sec. 7]. However, there only one half of the correspondence is considered, namely how their version of classical logic and cut-elimination can be embedded via some specific double-negation translations into intuitionistic logic. We conjecture also a correspondence in the other direction, namely that every double-negation translation and corresponding reduction sequences can be simulated by their cut-elimination procedure. Since we shall use as "point of reference" a more general cut-elimination procedure for cut-elimination in classical logic than the one described in [6], we are also able to draw the conclusion that given our conjecture is true, then double-negation translations are not enough to describe the *full* computational meaning of a classical proof.

As can be seen to answer the correspondence questions, we first have to make precise what we mean by cut-elimination in classical logic. Most cut-elimination procedures, including Gentzen's original one, only terminate if a particular strategy for cut-elimination is employed. Common examples being an innermost reduction strategy, or the elimination of the cut with the highest rank. Using those cut-elimination procedures we cannot characterise what the set of *all* normalforms of a classical proof is—they would produce only one or a limited number of normalforms. We shall therefore base our arguments on the cut-elimination procedure developed by Urban and Bierman [20, 21], which is like Gentzen's procedure except it imposes one slight restriction on how commuting cuts need to be analysed. Since this cut-elimination procedure is strongly normalising, we can calculate all cut-free normalforms of a classical proof. Because this procedure is not Church-Rosser, the collection of normalforms for a classical proof contains in general more than one element—as can be seen for example with the proofs (2) and (3). As this cut-elimination procedure puts only very slight restrictions on the process of cut-elimination we believe a good case can be made that the collection of normalforms calculated by this procedure includes all "essential" normalforms.[2] However this is a point we shall not be concerned with in this paper.

[2] Making such a case is hopeless for other strongly-normalising cut-elimination procedures, like the one by Dragalin [7], because although they are strongly-normalising, they enforce quite strong restrictions on how cuts can be eliminated. For example Dragalin does not allow (multi)cuts to permute over other (multi)cuts, see [19].

The cut-elimination procedure of Urban and Bierman will be described in more detail in Sec. 3, together with a variant—the colour protocol—developed by Danos et al. [6, 12]. Beforehand, however, we present some preliminaries about double-negation translations in Sec. 2. We will state the conjecture in Sec. 4, give some evidence on why this conjecture is plausible and present some ideas on how to prove it. In Sec. 5 we shall draw some conclusions with respect to the computational interpretation of classical proofs.

## 2 Preliminaries on Double-Negation Translations

We assume the reader has acquaintance with sequent-calculus formulations of classical and intuitionistic logic. Because there exist sequents that are provable in classical logic, but unprovable in intuitionistic logic, the interesting point of double-negation translations is that one can embed classical logic into intuitionistic logic so that provability is preserved. For example the following translation defined over formulae

$$\begin{aligned} A^* &\stackrel{\text{def}}{=} \neg\neg A \text{ with } A \text{ being atomic} \\ (\neg B)^* &\stackrel{\text{def}}{=} \neg(B^*) \\ (B\wedge C)^* &\stackrel{\text{def}}{=} B^*\wedge C^* \\ (B\supset C)^* &\stackrel{\text{def}}{=} B^*\supset C^* \\ (B\vee C)^* &\stackrel{\text{def}}{=} \neg(\neg(B^*)\wedge\neg(C^*)) \end{aligned} \tag{4}$$

can be used to show that every classical proof with the end-sequent

$$\Gamma \vdash \Delta$$

can be translated to an intuitionistic proof with the end-sequent

$$\Gamma^*, \neg\Delta^* \vdash \ . \tag{5}$$

We use the convention that if $\Gamma$ is the sequent-context $\{B_1, \ldots, B_n\}$ then $\Gamma^*$ stands for the sequent-context $\{B_1^*, \ldots, B_n^*\}$. Similarly for $\neg\Delta^*$. We shall also use the convention that $A$ stands for an atomic formula and $B, \ldots$ for arbitrary formulae. A similar embedding can be obtained with the translation:

$$\begin{aligned} A^\circ &\stackrel{\text{def}}{=} \neg\neg A \\ (\neg B)^\circ &\stackrel{\text{def}}{=} \neg(B^\circ) \\ (B\wedge C)^\circ &\stackrel{\text{def}}{=} \neg\neg(B^\circ\wedge C^\circ) \\ (B\supset C)^\circ &\stackrel{\text{def}}{=} \neg\neg(B^\circ\supset C^\circ) \\ (B\vee C)^\circ &\stackrel{\text{def}}{=} \neg\neg(B^\circ\vee C^\circ) \end{aligned} \tag{6}$$

The usual proof (see for example [4]) for establishing that every classical proof can be translated to an intuitionistic proof proceeds inductively by translating stepwise every inference rule in a proof. For instance an axiom of the form

$$A \vdash A$$

is translated by $(-)^*$ to the proof

$$\dfrac{\neg\neg A \vdash \neg\neg A}{\neg\neg A, \neg\neg\neg A \vdash}\ \neg_L$$

which conforms with the desired property stated in (5). When translating a proof ending with an $\vee_L$-rule

$$\dfrac{\begin{matrix}\vdots \\ B, \Gamma_1 \vdash \Delta_1\end{matrix} \quad \begin{matrix}\vdots \\ C, \Gamma_2 \vdash \Delta_2\end{matrix}}{B \vee C, \Gamma_1, \Gamma_2 \vdash \Delta_1, \Delta_2}\ \vee_L \tag{7}$$

then by induction hypothesis we have two intuitionistic proofs ending with

$$\begin{matrix}\vdots \\ B^*, \Gamma_1^*, \neg\Delta_1^* \vdash\end{matrix} \quad \text{and} \quad \begin{matrix}\vdots \\ C^*, \Gamma_2^*, \neg\Delta_2^* \vdash\end{matrix}\ .$$

We can then form the intuitionistic proof

$$\dfrac{\dfrac{\dfrac{\begin{matrix}\vdots \\ B^*, \Gamma_1^*, \neg\Delta_1^* \vdash\end{matrix}}{\Gamma_1^*, \neg\Delta_1^* \vdash \neg B^*}\ \neg_R \quad \dfrac{\begin{matrix}\vdots \\ C^*, \Gamma_2^*, \neg\Delta_2^* \vdash\end{matrix}}{\Gamma_2^*, \neg\Delta_2^* \vdash \neg C^*}\ \neg_R}{\Gamma_1^*, \Gamma_2^*, \neg\Delta_1^*, \neg\Delta_2^* \vdash \neg B^* \wedge \neg C^*}\ \wedge_R}{\neg(\neg B^* \wedge \neg C^*), \Gamma_1^*, \Gamma_2^*, \neg\Delta_1^*, \neg\Delta_2^* \vdash}\ \neg_L \tag{8}$$

as the translation of (7). For the sake of more clarity we will omit in what follows the sequent-contexts whenever they are unimportant. Thus we shall give for the proof-fragment shown in (8) only the following simplified inference rules:

$$\dfrac{\dfrac{\dfrac{\begin{matrix}\vdots \\ B^* \vdash\end{matrix}}{\vdash \neg B^*}\ \neg_R \quad \dfrac{\begin{matrix}\vdots \\ C^* \vdash\end{matrix}}{\vdash \neg C^*}\ \neg_R}{\vdash \neg B^* \wedge \neg C^*}\ \wedge_R}{\neg(\neg B^* \wedge \neg C^*) \vdash}\ \neg_L$$

When translating a classical proof ending with an $\wedge_R$-rule

$$\dfrac{\begin{matrix}\vdots \\ \vdash B\end{matrix} \quad \begin{matrix}\vdots \\ \vdash C\end{matrix}}{\vdash B \wedge C}\ \wedge_R \tag{9}$$

we have by induction hypothesis two intuitionistic proofs ending in

$$\begin{matrix}\vdots \\ \neg B^* \vdash\end{matrix} \quad \text{and} \quad \begin{matrix}\vdots \\ \neg C^* \vdash\end{matrix}\ .$$

In order to form an intuitionistic proof ending with the sequent $\neg(B^* \wedge C^*) \vdash$, we need to exploit the property of $(-)^*$ that one can always prove intuitionistically the sequent $\neg\neg(-)^* \vdash (-)^*$. For example in the atomic case one has for $\neg\neg A^* \vdash A^*$ the intuitionistic proof:

$$\dfrac{\dfrac{\dfrac{\dfrac{\neg A \vdash \neg A}{\neg A, \neg\neg A \vdash}\ \neg_L}{\neg A \vdash \neg\neg\neg A}\ \neg_R}{\neg\neg\neg\neg A, \neg A \vdash}\ \neg_L}{\neg\neg\neg\neg A \vdash \neg\neg A}\ \neg_R$$

Using proofs for $\neg\neg B^* \vdash B^*$ and $\neg\neg C^* \vdash C^*$, we can construct the following translated proof for (9):

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\neg B^* \vdash}}{\vdash \neg\neg B^*}\,\neg_R \qquad \begin{array}{c}\vdots\\ \neg\neg B^* \vdash B^*\end{array}}{\vdash B^*}\,cut \qquad \dfrac{\dfrac{\dfrac{\vdots}{\neg C^* \vdash}}{\vdash \neg\neg C^*}\,\neg_R \qquad \begin{array}{c}\vdots\\ \neg\neg C^* \vdash C^*\end{array}}{\vdash C^*}\,cut}{\dfrac{\vdash B^* \wedge C^*}{\neg(B^* \wedge C^*) \vdash}\,\neg_L}\,\wedge_R$$

We shall refer to the cuts introduced by the double-negation translation as *auxiliary cuts*. For a number of reasons (one of them being to minimise the amount of writing) we shall use a new inference rule, namely

$$\dfrac{\vdash \neg\neg B}{\vdash B}\,\neg\neg_R$$

to stand for auxiliary cuts, which have always the form:

$$\dfrac{\begin{array}{c}\vdots\\ \Gamma \vdash \neg\neg B\end{array} \qquad \begin{array}{c}\vdots\\ \neg\neg B \vdash B\end{array}}{\Gamma \vdash B}\,cut\ .$$

Clearly, this new rule does not affect the provability of sequents. Issues whether the $\neg\neg_R$-rule (or an auxiliary cut) affects the behaviour under cut-elimination are delayed until Sec. 4. With this new inference rule we can give the translation of (9) more compactly as:

$$\dfrac{\dfrac{\dfrac{\dfrac{\vdots}{\neg B^* \vdash}}{\vdash \neg\neg B^*}\,\neg_R}{\vdash B^*}\,\neg\neg_R \qquad \dfrac{\dfrac{\dfrac{\vdots}{\neg C^* \vdash}}{\vdash \neg\neg C^*}\,\neg_R}{\vdash C^*}\,\neg\neg_R}{\dfrac{\vdash B^* \wedge C^*}{\neg(B^* \wedge C^*) \vdash}\,\neg_L}\,\wedge_R$$

The translations for the rules $contr_L$, $contr_R$, $weak_L$, $weak_R$, $\neg_L$, $\neg_R$, $\vee_{R_i}$, $\wedge_{L_i}$, $\supset_L$ and $\supset_R$ are left as exercises to the reader.

We can now give the double-negation translations of the two subproofs

$$\dfrac{\dfrac{A \vdash A \quad A \vdash A}{A \vee A \vdash A, A}\,\vee_L}{A \vee A \vdash A}\,contr_R \qquad\qquad \dfrac{\dfrac{A \vdash A \quad A \vdash A}{A, A \vdash A \wedge A}\,\wedge_R}{A \vdash A \wedge A}\,contr_L$$

shown in (1). The translations are:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg\neg A \vdash \neg\neg A}{\neg\neg A, \neg\neg\neg A \vdash}\,\neg_L}{\neg\neg\neg A \vdash \neg\neg\neg A}\,\neg_R \qquad \dfrac{\dfrac{\neg\neg A \vdash \neg\neg A}{\neg\neg A, \neg\neg\neg A \vdash}\,\neg_L}{\neg\neg\neg A \vdash \neg\neg\neg A}\,\neg_R}{\neg\neg\neg A, \neg\neg\neg A \vdash \neg\neg\neg A \wedge \neg\neg\neg A}\,\wedge_R}{\neg(\neg\neg\neg A \wedge \neg\neg\neg A), \neg\neg\neg A, \neg\neg\neg A \vdash}\,\neg_L}{\neg(\neg\neg\neg A \wedge \neg\neg\neg A), \neg\neg\neg A \vdash}\,contr_L \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg\neg A \vdash \neg\neg A}{\neg\neg A, \neg\neg\neg A \vdash}\,\neg_L}{\neg\neg A \vdash \neg\neg\neg\neg A}\,\neg_R}{\neg\neg A \vdash \neg\neg A}\,\neg\neg_R \qquad \dfrac{\dfrac{\dfrac{\neg\neg A \vdash \neg\neg A}{\neg\neg A, \neg\neg\neg A \vdash}\,\neg_L}{\neg\neg A \vdash \neg\neg\neg\neg A}\,\neg_R}{\neg\neg A \vdash \neg\neg A}\,\neg\neg_R}{\neg\neg A, \neg\neg A \vdash \neg\neg A \wedge \neg\neg A}\,\wedge_R}{\neg\neg A, \neg\neg A, \neg(\neg\neg A \wedge \neg\neg A) \vdash}\,\neg_L}{\neg\neg A, \neg(\neg\neg A \wedge \neg\neg A) \vdash}\,contr_L$$

To give a double-negation translation for the whole proof, we need to be able to translate instances of the cut-rule. While the logical inference rules and the structural rules have relatively canonical double-negation translations, there is a choice for how to translate cut-rules. Consider the following cut-instance:

$$\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash B\end{matrix} \quad \begin{matrix}\pi_2\\ \vdots\\ B\vdash\end{matrix}}{\vdash}\ cut \tag{10}$$

By induction hypothesis we have two intuitionistic proofs $\pi_1^*$ and $\pi_2^*$ with end-sequents:

$$\begin{matrix}\pi_1^*\\ \vdots\\ \neg B^* \vdash\end{matrix} \qquad \text{and} \qquad \begin{matrix}\pi_2^*\\ \vdots\\ B^* \vdash\end{matrix}\ .$$

To form an intuitionistic proof with the end-sequent $\vdash$ (remember we omit the sequent-contexts), we can translate (10) either as:

$$\dfrac{\begin{matrix}\pi_1^*\\ \vdots\\ \neg B^* \vdash\end{matrix} \quad \dfrac{\begin{matrix}\pi_2^*\\ \vdots\\ B^* \vdash\end{matrix}}{\vdash \neg B^*}\ \neg_R}{\vdash}\ cut \qquad \text{or} \qquad \dfrac{\dfrac{\dfrac{\begin{matrix}\pi_1^*\\ \vdots\\ \neg B^* \vdash\end{matrix}}{\vdash \neg\neg B^*}\ \neg_R}{\vdash B^*}\ \neg\neg_R \quad \begin{matrix}\pi_2^*\\ \vdots\\ B^* \vdash\end{matrix}}{\vdash}\ cut$$

We refer to these choices as *left*- and *right-translation* of a cut, respectively.

Coming back to the proof given in (1), let us call the left- and right-translation of this proof $\pi_L^*$ and $\pi_R^*$, respectively. Eliminating all cuts (including auxiliary cuts) from $\pi_L^*$ and $\pi_R^*$, we obtain the two cut-free proofs $\pi_L'^*$ and $\pi_R'^*$ shown in Fig. 1. It turns out (we however leave out the calculations) that had we double-negation translated the two normalforms of (1) and then eliminated all auxiliary cuts from the double-negated proofs, we would have also obtained the proofs $\pi_L'^*$ and $\pi_R'^*$: The normalform (2) obtained from (1) by commuting the cut to the left leads to $\pi_L'^*$—the normalform of the left-translation of (1), while (3) obtained by commuting the cut to the right leads to $\pi_R'^*$—the normalform of the right-translation of (1). We take this as a first hint that double-negation translations seem to be able to simulate the behaviour of cut-elimination in classical logic.

To sum up this section, let us remark that a similar "story" can be told for the translation $(-)^\circ$ given in (6). In fact, it can be told for any sensible notion of double-negation translation. For example it would be completely immaterial to our "story" if we had translated atomic formulae $A$ as $\neg\neg A$, $\neg\neg\neg\neg A$ or $\neg\neg\neg\neg\neg\neg A$. The most important property we distill from the arguments above is that we regard a double-negation translation, say $(-)^x$, as a translation of a classical proof having an end-sequent

$$\begin{matrix}\vdots\\ \Gamma \vdash \Delta\end{matrix}$$

to an intuitionistic proof with the end-sequent

$$\begin{matrix}\vdots\\ \Gamma^x, \neg\Delta^x \vdash\end{matrix}$$

$$\pi_L'^{*} = \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg A \vdash \neg A}{\neg A, \neg\neg A \vdash}\,\neg_L}{\neg A \vdash \neg\neg\neg A}\,\neg_R \quad \dfrac{\dfrac{\neg A \vdash \neg A}{\neg A, \neg\neg A \vdash}\,\neg_L}{\neg A \vdash \neg\neg\neg A}\,\neg_R}{\neg A, \neg A \vdash \neg\neg\neg A\wedge\neg\neg\neg A}\,\wedge_R}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg A, \neg A \vdash}\,\neg_L}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg A \vdash}\,contr_L}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A) \vdash \neg\neg A}\,\neg_R \quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg A \vdash \neg A}{\neg A, \neg\neg A \vdash}\,\neg_L}{\neg A \vdash \neg\neg\neg A}\,\neg_R \quad \dfrac{\dfrac{\neg A \vdash \neg A}{\neg A, \neg\neg A \vdash}\,\neg_L}{\neg A \vdash \neg\neg\neg A}\,\neg_R}{\neg A, \neg A \vdash \neg\neg\neg A\wedge\neg\neg\neg A}\,\wedge_R}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg A, \neg A \vdash}\,\neg_L}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg A \vdash}\,contr_L}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A) \vdash \neg\neg A}\,\neg_R}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg(\neg\neg\neg A\wedge\neg\neg\neg A) \vdash \neg\neg A\wedge\neg\neg A}\,\wedge_R}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg(\neg\neg A\wedge\neg\neg A) \vdash}\,\neg_L}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg(\neg\neg A\wedge\neg\neg A) \vdash}\,contr_L$$

$$\pi_R'^{*} = \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg A \vdash \neg A}{\neg\neg A, \neg A \vdash}\,\neg_L}{\neg\neg A \vdash \neg\neg A}\,\neg_R \quad \dfrac{\dfrac{\neg A \vdash \neg A}{\neg\neg A, \neg A \vdash}\,\neg_L}{\neg\neg A \vdash \neg\neg A}\,\neg_R}{\neg\neg A, \neg\neg A \vdash \neg\neg A\wedge\neg\neg A}\,\wedge_R}{\neg(\neg\neg A\wedge\neg\neg A), \neg\neg A, \neg\neg A \vdash}\,\neg_L}{\neg(\neg\neg A\wedge\neg\neg A), \neg\neg A \vdash}\,contr_L}{\neg(\neg\neg A\wedge\neg\neg A) \vdash \neg\neg\neg A}\,\neg_R \quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg A \vdash \neg A}{\neg\neg A, \neg A \vdash}\,\neg_L}{\neg\neg A \vdash \neg\neg A}\,\neg_R \quad \dfrac{\dfrac{\neg A \vdash \neg A}{\neg\neg A, \neg A \vdash}\,\neg_L}{\neg\neg A \vdash \neg\neg A}\,\neg_R}{\neg\neg A, \neg\neg A \vdash \neg\neg A\wedge\neg\neg A}\,\wedge_R}{\neg(\neg\neg A\wedge\neg\neg A), \neg\neg A, \neg\neg A \vdash}\,\neg_L}{\neg(\neg\neg A\wedge\neg\neg A), \neg\neg A \vdash}\,contr_L}{\neg(\neg\neg A\wedge\neg\neg A) \vdash \neg\neg\neg A}\,\neg_R}{\neg(\neg\neg A\wedge\neg\neg A), \neg(\neg\neg A\wedge\neg\neg A) \vdash \neg\neg\neg A\wedge\neg\neg\neg A}\,\wedge_R}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg(\neg\neg A\wedge\neg\neg A), \neg(\neg\neg A\wedge\neg\neg A) \vdash}\,\neg_L}{\neg(\neg\neg\neg A\wedge\neg\neg\neg A), \neg(\neg\neg A\wedge\neg\neg A) \vdash}\,contr_L$$

**Fig. 1.** One can obtain the first proof by double-negation translation of (1) using the left-translation for the cut, and then eliminating all cuts including the auxiliary cuts. Equally, one can first reduce (1) to (2), double-negate translate (2) and then eliminate all cuts. Similarly with the second proof: it can be obtained by right-translating the cut in (1) and then eliminate all cuts; or by double-negate translate (3) and then eliminate all cuts.

preserving the "structure" of the classical proof. In order to achieve this one needs the property that $\neg\neg(-)^x \vdash (-)^x$ is intuitionistically derivable. However, this leaves us with many possible double negation translations—clearly the ones given by Gentzen, Gödel and Kolmogorov are not the only ones that satisfy these constraints.

## 3 Cut-Elimination and Its Coloured Variant

Urban and Bierman have shown in [19, 21] that only a small restriction on the standard cut-elimination procedure for classical and intuitionistic logic is sufficient to obtain strongly normalising proof-transformations. The *logical cuts*, also sometimes called *key-cuts*, are transformed by this cut-elimination procedure in a completely standard fashion [10]. For example the logical cut

$$\dfrac{\dfrac{\begin{array}{c}\pi_1\\ \vdots\\ \vdash B\end{array} \quad \begin{array}{c}\pi_2\\ \vdots\\ \vdash C\end{array}}{\vdash B\wedge C}\,\wedge_R \quad \dfrac{\begin{array}{c}\pi_3\\ \vdots\\ B\vdash\end{array}}{B\wedge C\vdash}\,\wedge_{L_1}}{\vdash}\,cut \tag{11}$$

is transformed to

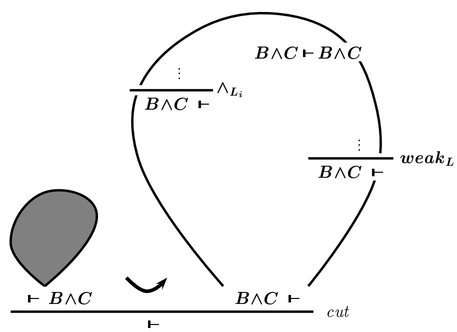
$$\dfrac{\begin{array}{c}\pi_1\\ \vdots\\ \vdash B\end{array} \quad \begin{array}{c}\pi_3\\ \vdots\\ B \vdash\end{array}}{\vdash}\ cut$$

and so on for the other connectives. As before we conveniently ignore all matters to do with how the sequent-contexts should be adjusted. A logical cut can be characterised as a cut where in both subproofs the cut-formulas are *freshly introduced* by logical rules directly above the cut. For example in the proof

$$\dfrac{\begin{array}{c}\pi\\ \vdots\end{array}}{\vdash B}\ r$$

we say the formula $B$ is freshly introduced if it is what usually is called the main formula of the logical inference rule $r$. Consequently, a logical cut is a cut where the cut-formula is freshly introduced in the two immediate subproofs of the cut. In all other cases we have a *commuting cut*. The cut in (1), for example, is a commuting cut because the cut-formula $A$ is in both subproofs introduced by a contraction-rule, which is not considered to be a logical inference rule.

Gentzen introduced proof-transformations that permute commuting cuts upwards in a stepwise fashion only by rewriting neighboring inference rules. In contrast, the cut-elimination procedure of Urban and Bierman contains proof-transformations that push commuting cuts upwards in a single "big" step towards all places where the cut-formula was introduced. Consider the following picture




where the cut-formula on the right-hand side is *not* freshly introduced, rather it is introduced somewhere deeper inside the subproof and because of contractions possibly in several places. We have indicated three cases for a cut-formula being introduced: by a logical inference rule, by an axiom and by a weakening-rule—in general we can have any mixture of these cases. To eliminate such commuting cuts the procedure of Urban and Bierman pushes up, roughly speaking, the cut-rule until it reaches *all* places where the cut-formula is introduced in *one* step. In case it reaches a logical inference-rule, then the proof on the left-hand side will be cut against this logical inference-rule; in case of an axiom, the proof on the left-hand side replaces the axiom and in case of a weakening-rule, the proof on the left-hand side is deleted. (Again all matters to do with adjusting the sequent-contexts are omitted in this paper. The work reported in [19, 21] formulates these proof-transformation as term-rewriting rules where such adjustments are built into the inference rules, very similar to term-rewriting in the simply-typed lambda-calculus.)

The important property of the cut-elimination procedure of Urban and Bierman is the fact that it is strongly-normalising. This property is not obvious: the reduction-rule for commuting cuts allows a cut-rule to "jump" over other cut-rules—a highly problematic reduction if one tries to construct a decreasing measure for cut-elimination. Also this rule might generate several copies of a subproof when a cut-formula is introduced in several places. Urban and Bierman, therefore, had to resort to a quite complicated logical relations argument to show strong normalisation.

Recall that the cut-elimination procedure of Urban and Bierman is *not* Church-Rosser: when on both sides of a commuting cut the cut-formula is not freshly introduced, then this cut can be moved either to the left or to the right, leading in general to two "non-joinable" proofs. For example, the proof shown in (1) can be reduced in one step to the normalform (2) or in one step to (3). Which choice is taken is left unspecified by the procedure. Nevertheless with this cut-elimination procedure we have some effective means to calculate for a classical proof its collection of *all* normalforms—for example by naïvely trying out all possible reductions.[3]

The cut-elimination procedure of Urban and Bierman was in part inspired by the work of Danos et al. [6]. The main difference is that their cut-elimination procedure *is* Church-Rosser. They achieve this by pre-determining every choice that can be made during cut-elimination. This pre-determination is done via *colours*, which are annotated to every formula and subformula in a proof. To see how colours work, consider again the proof shown in (1). The choice about which direction is taken for the commuting cut is determined by annotating the colour '$\leftharpoonup$' or '$\rightharpoonup$' to the cut-formula $A$. The colour-protocol of Danos et al. pre-scribes that in the former case it is first attempted to permute the cut to the left and in the second case to the right (hence the use of an arrow to denote a colour!). For example, if we want to reach from (1) the normalform (2) we need to orient the colour of the cut-formula $A$ to the left, as shown below

$$\dfrac{\dfrac{\dfrac{\overleftharpoon{A} \vdash \overleftharpoon{A} \quad \overleftharpoon{A} \vdash \overleftharpoon{A}}{\overleftharpoon{\overleftharpoon{A} \vee \overleftharpoon{A}} \vdash \overleftharpoon{A}, \overleftharpoon{A}}\,\vee_L}{\overleftharpoon{\overleftharpoon{A} \vee \overleftharpoon{A}} \vdash \overleftharpoon{A}}\,contr_R \qquad \dfrac{\dfrac{\overleftharpoon{A} \vdash \overleftharpoon{A} \quad \overleftharpoon{A} \vdash \overleftharpoon{A}}{\overleftharpoon{A}, \overleftharpoon{A} \vdash \overrightharpoon{\overleftharpoon{A} \wedge \overleftharpoon{A}}}\,\wedge_R}{\overleftharpoon{A} \vdash \overrightharpoon{\overleftharpoon{A} \wedge \overleftharpoon{A}}}\,contr_L}{\overleftharpoon{\overleftharpoon{A} \vee \overleftharpoon{A}} \vdash \overrightharpoon{\overleftharpoon{A} \wedge \overleftharpoon{A}}}\,cut$$

If the normalform (3) is to be reached, then accordingly we need to orient the colour of the cut-formula to the right. Note however that choosing colours has nothing to do with imposing a strategy for cut-elimination: it is not cuts that are selected by them, but rather the way how cuts are reduced. Important for our discussion is the fact that

[3] A less naïve method, which only tries out all possible reduction for outermost cuts, is described in [19].

one is, however, not completely free about how to annotate colours to a sequent proof. In fact once the colour '$\leftharpoonup$' is chosen for the cut-formula $A$ in (1), all occurrences of $A$ must have this colour. The only "free" choices in this proof are the colours for the formulae $A \vee A$ and $A \wedge A$—for them we can make any choice, but it has to be consistent throughout the proof. Danos et al. state this consistency requirement using the notion of an *identity class* in a proof (the following definition is slightly adapted from [16, Page 107]):

**Definition 1.** *Occurrences of (sub)formulae in a proof are* identified *whenever they are the corresponding occurrences of the same (sub)formula in*

- *the two formulae in an axiom,*
- *the cut-formulae in a cut and*
- *the up and down occurrences of a formula in an inference rule (this includes the contracted occurrences in contractions rules).*

*An* identity class *in a proof is the reflexive, symmetric and transitive closure of the identification relation.* □

The consistency requirement can then be stated as follows: Whenever colours are annotated to a proof, then every formula in an identity class must receive the same colour.

The interesting point of colour-annotations is the fact that they determine uniquely a normalform. In light of this, it seems reasonable to regard as the collection of normalforms reachable from a classical proof all those for which a colour annotation exists that makes them reachable. Then the question arises: Can we find for every normalform reachable by the (un-coloured) cut-elimination procedure of Urban and Bierman a colour-annotation that makes them reachable by the cut-elimination procedure of Danos et al.? The answer is no and for deep reasons! Consider the following classical proof

$$\dfrac{\begin{array}{c}(1)\\ \vdots \\ A \vee A \vdash A \wedge A\end{array} \qquad \dfrac{\dfrac{A \wedge A \vdash A \wedge A \qquad A \wedge A \vdash A \wedge A}{A \wedge A, A \wedge A \vdash (A \wedge A) \wedge (A \wedge A)}\ \wedge_R}{A \wedge A \vdash (A \wedge A) \wedge (A \wedge A)}\ contr_L}{A \vee A \vdash (A \wedge A) \wedge (A \wedge A)}\ cut \tag{12}$$

where we cut the proof from (1) against a proof whose cut-formula, $A \wedge A$, is contracted in the right-subproof. We can reduce the lower cut so that we obtain two copies of the proof (1):

$$\dfrac{\dfrac{\begin{array}{c}(1)\\ \vdots \\ A \vee A \vdash A \wedge A\end{array} \qquad \begin{array}{c}(1)\\ \vdots \\ A \vee A \vdash A \wedge A\end{array}}{A \vee A, A \vee A \vdash (A \wedge A) \wedge (A \wedge A)}\ \wedge_R}{A \vee A \vdash (A \wedge A) \wedge (A \wedge A)}\ contr_L \tag{13}$$

Without colours, we can then reduce each copy completely independently as follows:

$$\dfrac{\dfrac{\begin{array}{c}(2)\\ \vdots \\ A \vee A \vdash A \wedge A\end{array} \qquad \begin{array}{c}(3)\\ \vdots \\ A \vee A \vdash A \wedge A\end{array}}{A \vee A, A \vee A \vdash (A \wedge A) \wedge (A \wedge A)}\ \wedge_R}{A \vee A \vdash (A \wedge A) \wedge (A \wedge A)}\ contr_L \tag{14}$$

Such a behaviour cannot be achieved by using colours: the colours must be annotated before cut-elimination commences and is invariant under cut-reductions. Consequently, whenever a cut is duplicated in a reduction sequence (as in the reduction (12)→(13)), the colour-annotation prevents both instances from reducing differently. (The deeper reason mentioned earlier is that one just cannot pre-determine the choices in a completely non-deterministic reduction system.)

Comparing the cut-elimination procedure of Urban and Bierman with the one of Danos et al., two points stand out: Both cut-elimination procedures are strongly normalising[4] and also determine a collection of normalforms reachable from a sequent-proof in classical logic. As shown by example, these collections contain in general more than one element. Also as shown by example, the collection determined by the procedure of Danos et al. is generally a proper subset of the collection determined by the procedure of Urban and Bierman. The colour-annotations in the procedure of Danos et al. cannot fully account for the non-determinism present in classical logic.

Both cut-elimination procedures can also be used for reducing intuitionistic proofs. Because of the restrictions imposed upon intuitionistic sequents, non-deterministic reduction sequences such as (12)→(13)→(14) cannot be constructed. But still the procedure of Urban and Bierman is *not* Church-Rosser in the intuitionistic case, and also different colour-annotations of an intuitionistic proof might lead to different normalforms. However, as mentioned earlier, we regard the differences between the normalforms reachable from an intuitionistic sequent-proof as inessential and regard cut-elimination as morally Church-Rosser. That in turn means that in the intuitionistic case there is no difference between coloured and un-coloured cut-elimination—at least morally.

## 4 Conjecture

We have already seen that by translating the cut in (1) using a left- and right-translation, we can simulate the reductions (1)→(2) and (1)→(3) by double-negations. However in general, a left- or right-translation of a cut is not sufficient to simulate all cut-reduction sequences in classical logic. Consider the following instance of a logical cut:

$$\dfrac{\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash B\end{matrix}}{\vdash B\vee C}\,\vee_{R_1} \qquad \dfrac{\begin{matrix}\pi_2\\ \vdots\\ B\vdash\end{matrix}\quad\begin{matrix}\pi_3\\ \vdots\\ C\vdash\end{matrix}}{B\vee C\vdash}\,\vee_L}{\vdash}\,cut \tag{15}$$

We can reduce this cut to

$$\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash B\end{matrix}\quad\begin{matrix}\pi_2\\ \vdots\\ B\vdash\end{matrix}}{\vdash}\,cut$$

[4] Danos et al. showed strong normalisation of their cut-elimination procedure by translating reduction sequences in classical logic to reduction sequences of proof-nets in linear logic.

and assuming that the cut-formula $B$ is not freshly introduced in $\pi_1$, we can further permute $\pi_2$ inside $\pi_1$. This behaviour correspond to the colour-annotation

$$\dfrac{\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash \overleftharpoon{B}\end{matrix}}{\vdash \overleftharpoon{B}\vee C}\vee_{R_1} \qquad \dfrac{\begin{matrix}\pi_2\\ \vdots\\ \overleftharpoon{B}\vdash\end{matrix} \quad \begin{matrix}\pi_3\\ \vdots\\ C\vdash\end{matrix}}{\overleftharpoon{B}\vee C\vdash}\vee_L}{\vdash}\,cut \tag{16}$$

where we leave the colour annotation for $C$ and $B{\vee}C$ unspecified, since it is not important for the argument at hand. The behaviour of (16) can be simulated by the double-negation translation $(-)^*$. The double-negated version of (15) is as follows:

$$\dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\pi_1^*\\ \vdots\\ \neg B^*\vdash\end{matrix}}{\neg B^*\wedge\neg C^*\vdash}\wedge_{L_1}}{\vdash\neg(\neg B^*\wedge\neg C^*)}\neg_R}{\neg\neg(\neg B^*\wedge\neg C^*)\vdash}\neg_L \qquad \dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\pi_2^*\\ \vdots\\ B^*\vdash\end{matrix}}{\vdash\neg B^*}\neg_R \quad \dfrac{\begin{matrix}\pi_3^*\\ \vdots\\ C^*\vdash\end{matrix}}{\vdash\neg C^*}\neg_R}{\vdash\neg B^*\wedge\neg C^*}\wedge_R}{\neg(\neg B^*\wedge\neg C^*)\vdash}\neg_L}{\vdash\neg\neg(\neg B^*\wedge\neg C^*)}\neg_R}{\vdash}\,cut$$

which reduces in three steps to the proof

$$\dfrac{\begin{matrix}\pi_1^*\\ \vdots\\ \neg B^*\vdash\end{matrix} \qquad \dfrac{\begin{matrix}\pi_2^*\\ \vdots\\ B^*\vdash\end{matrix}}{\vdash\neg B^*}\neg_R}{\vdash}\,cut$$

Because the $\neg_R$-rule introduces freshly the cut-formula $\neg B^*$, the cut is “blocked” from reducing to the right. It must first reduce to the left just as the colour annotation in (16) prescribed. If we wanted to simulated the opposite colouring for $B$, namely

$$\dfrac{\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash \overrightharpoon{B}\end{matrix}}{\vdash \overrightharpoon{B}\vee C}\vee_{R_1} \qquad \dfrac{\begin{matrix}\pi_2\\ \vdots\\ \overrightharpoon{B}\vdash\end{matrix} \quad \begin{matrix}\pi_3\\ \vdots\\ C\vdash\end{matrix}}{\overrightharpoon{B}\vee C\vdash}\vee_L}{\vdash}\,cut \tag{17}$$

it turns out we have to double-negate translate (15) using the translation $(-)^\circ$ given in (6). The resulting intuitionistic proof is:

$$
\begin{prooftree}
\AxiomC{$\pi_1^\circ$}
\noLine
\UnaryInfC{$\vdots$}
\noLine
\UnaryInfC{$\neg B^\circ \vdash$}
\RightLabel{$\neg_R$}
\UnaryInfC{$\vdash \neg\neg B^\circ$}
\RightLabel{$\neg\neg_R$}
\UnaryInfC{$\vdash B^\circ$}
\RightLabel{$\vee_{R_1}$}
\UnaryInfC{$\vdash B^\circ \vee C^\circ$}
\RightLabel{$\neg_L$}
\UnaryInfC{$\neg(B^\circ \vee C^\circ) \vdash$}
\RightLabel{$\neg_R$}
\UnaryInfC{$\vdash \neg\neg(B^\circ \vee C^\circ)$}
\RightLabel{$\neg_L$}
\UnaryInfC{$\neg\neg\neg(B^\circ \vee C^\circ) \vdash$}
\AxiomC{$\pi_2^\circ$}
\noLine
\UnaryInfC{$\vdots$}
\noLine
\UnaryInfC{$B^\circ \vdash$}
\AxiomC{$\pi_3^\circ$}
\noLine
\UnaryInfC{$\vdots$}
\noLine
\UnaryInfC{$C^\circ \vdash$}
\RightLabel{$\vee_L$}
\BinaryInfC{$B^\circ \vee C^\circ \vdash$}
\RightLabel{$\neg_R$}
\UnaryInfC{$\vdash \neg(B^\circ \vee C^\circ)$}
\RightLabel{$\neg_L$}
\UnaryInfC{$\neg\neg(B^\circ \vee C^\circ) \vdash$}
\RightLabel{$\neg_R$}
\UnaryInfC{$\vdash \neg\neg\neg(B^\circ \vee C^\circ)$}
\RightLabel{$cut$}
\BinaryInfC{$\vdash$}
\end{prooftree}
$$

which after four steps reduces to the proof

$$
\begin{prooftree}
\AxiomC{$\pi_1^\circ$}
\noLine
\UnaryInfC{$\vdots$}
\noLine
\UnaryInfC{$\neg B^\circ \vdash$}
\RightLabel{$\neg_R$}
\UnaryInfC{$\vdash \neg\neg B^\circ$}
\RightLabel{$\neg\neg_R$}
\UnaryInfC{$\vdash B^\circ$}
\AxiomC{$\pi_2^\circ$}
\noLine
\UnaryInfC{$\vdots$}
\noLine
\UnaryInfC{$B^\circ \vdash$}
\RightLabel{$cut$}
\BinaryInfC{$\vdash$}
\end{prooftree}
\tag{18}
$$

What happens next, however, is not clear at first sight. If we expand the $\neg\neg_R$-rule to an auxiliary cut, then the cut can reduce into both directions. If we regard the $\neg\neg_R$-rule as an inference rule in its own right, then the cut-formula $B^\circ$ is freshly introduced in the subproof on the left-hand side and therefore the cut can only move to the right—just as prescribed by the colour-annotation in (17). Although we do not have a proof of this fact, experiments with [17] have convinced us that when cut-elimination is concerned, we can indeed regard the $\neg\neg_R$-rule as a proper inference rule with the consequence that in the proof above $B^\circ$ is freshly introduced. (Roughly speaking, if we had expanded the $\neg\neg_R$-rule to an auxiliary cut in the proof (18), then moving the cut to the left means it cannot move very far, namely only to the place where $B^\circ$ is introduced in the proof of the sequent $\neg\neg B^\circ \vdash B^\circ$ and then the cut has to move right. In effect we obtain a behaviour which is almost identical to moving the cut to the right in the first place.)

Since the colour-protocol of Danos et al. allows us to annotate in many circumstances either colour '$\leftharpoonup$' or '$\rightharpoonup$' to the formulae in a classical proof, we need however to depart from the traditional double-negation technique that translates a classical proof uniformly using a single double-negation translation. To simulate the coloured cut-elimination procedure in a meaningful way, we need to allow more than one double-negation translation. Let us explain this fact with a classical proof $\pi$ ending in a cut with the cut-formula

$$\overleftharpoon{\overrightharpoon{B} \vee \overleftharpoon{C}}\ .$$

We will show that the behaviour of this (coloured) cut can be simulated by a double-negation translation with the clause $(B \vee C)^\bullet \stackrel{\text{def}}{=} \neg\neg(B^\bullet \vee \neg\neg C^\bullet)$. To show this we analyse all cases how the cut in $\pi$ could have arisen. Consider first the case where $\pi$

ends with the following logical cut

$$\dfrac{\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash \overrightarrow{B}\end{matrix}}{\vdash \overleftarrow{\overrightarrow{B}\vee\overleftarrow{C}}}\vee_{R_1} \qquad \dfrac{\begin{matrix}\pi_2\\ \vdots\\ \overrightarrow{B}\vdash\end{matrix}\quad\begin{matrix}\pi_3\\ \vdots\\ \overleftarrow{C}\vdash\end{matrix}}{\overleftarrow{\overrightarrow{B}\vee\overleftarrow{C}}\vdash}\vee_L}{\vdash}\,cut \tag{19}$$

which can reduce to

$$\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash \overrightarrow{B}\end{matrix}\quad\begin{matrix}\pi_2\\ \vdots\\ \overrightarrow{B}\vdash\end{matrix}}{\vdash}\,cut \tag{20}$$

The $(-)^{\bullet}$-translated version of $\pi$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\pi_1^{\bullet}\\ \vdots\\ \neg B^{\bullet}\vdash\end{matrix}}{\vdash\neg\neg B^{\bullet}}\neg_R}{\vdash B^{\bullet}}\neg\neg_R}{\vdash B^{\bullet}\vee\neg\neg C^{\bullet}}\vee_{R_1}}{\neg(B^{\bullet}\vee\neg\neg C^{\bullet})\vdash}\neg_L}{\vdash\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})}\neg_R}{\neg\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})\vdash}\neg_L \qquad \dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\pi_2^{\bullet}\\ \vdots\\ B^{\bullet}\vdash\end{matrix}\quad\dfrac{\dfrac{\begin{matrix}\pi_3^{\bullet}\\ \vdots\\ C^{\bullet}\vdash\end{matrix}}{\vdash\neg C^{\bullet}}\neg_R}{\neg\neg C^{\bullet}\vdash}\neg_L}{B^{\bullet}\vee\neg\neg C^{\bullet}\vdash}\vee_L}{\vdash\neg(B^{\bullet}\vee\neg\neg C^{\bullet})}\neg_R}{\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})\vdash}\neg_L}{\vdash\neg\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})}\neg_R}{\vdash}\,cut$$

reduces to

$$\dfrac{\dfrac{\dfrac{\begin{matrix}\pi_1^{\bullet}\\ \vdots\\ \neg B^{\bullet}\vdash\end{matrix}}{\vdash\neg\neg B^{\bullet}}\neg_R}{\vdash B^{\bullet}}\neg\neg_R \qquad \begin{matrix}\pi_2^{\bullet}\\ \vdots\\ B^{\bullet}\vdash\end{matrix}}{\vdash}\,cut$$

where (remember we regard the $\neg\neg_R$-rule as proper inference rule) the proof $\pi_1^{\bullet}$ has to move inside $\pi_2^{\bullet}$ just like the behaviour of (20). If $\pi$ ends with the logical cut

$$\dfrac{\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash \overleftarrow{C}\end{matrix}}{\vdash \overleftarrow{\overrightarrow{B}\vee\overleftarrow{C}}}\vee_{R_2} \qquad \dfrac{\begin{matrix}\pi_2\\ \vdots\\ \overrightarrow{B}\vdash\end{matrix}\quad\begin{matrix}\pi_3\\ \vdots\\ \overleftarrow{C}\vdash\end{matrix}}{\overleftarrow{\overrightarrow{B}\vee\overleftarrow{C}}\vdash}\vee_L}{\vdash}\,cut \tag{21}$$

which can reduce to

$$\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash \overleftarrow{C}\end{matrix}\quad\begin{matrix}\pi_3\\ \vdots\\ \overleftarrow{C}\vdash\end{matrix}}{\vdash} \tag{22}$$

then, the $(-)^{\bullet}$-translated version of $\pi$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\pi_1^{\bullet}\\ \vdots\end{matrix}}{\neg C^{\bullet}\vdash}}{\vdash\neg\neg C^{\bullet}}\,\neg_R}{\vdash B^{\bullet}\vee\neg\neg C^{\bullet}}\,\vee_{R_1}}{\neg(B^{\bullet}\vee\neg\neg C^{\bullet})\vdash}\,\neg_L}{\vdash\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})}\,\neg_R}{\neg\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})\vdash}\,\neg_L \quad \dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}\pi_2^{\bullet}\\ \vdots\\ B^{\bullet}\vdash\end{matrix} \quad \dfrac{\dfrac{\begin{matrix}\pi_3^{\bullet}\\ \vdots\\ C^{\bullet}\vdash\end{matrix}}{\vdash\neg C^{\bullet}}\,\neg_R}{\neg\neg C^{\bullet}\vdash}\,\neg_L}{B^{\bullet}\vee\neg\neg C^{\bullet}\vdash}\,\vee_L}{\vdash\neg(B^{\bullet}\vee\neg\neg C^{\bullet})}\,\neg_R}{\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})\vdash}\,\neg_L}{\vdash\neg\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})}\,\neg_R$$
$$\overline{\qquad\qquad\vdash\qquad\qquad}\ cut$$

reduces to

$$\dfrac{\begin{matrix}\pi_1^{\bullet}\\ \vdots\\ \neg C^{\bullet}\vdash\end{matrix} \quad \dfrac{\begin{matrix}\pi_3^{\bullet}\\ \vdots\\ C^{\bullet}\vdash\end{matrix}}{\vdash\neg C^{\bullet}}\,\neg_R}{\vdash}\ cut$$

where proof $\pi_3^{\bullet}$ has to move inside $\pi_1^{\bullet}$ just like in the proof (22). The only case we still need to consider is when $\pi$ ends in a commuting cut of the form

$$\dfrac{\begin{matrix}\pi_1\\ \vdots\\ \vdash \overleftharpoon{\overrightharpoon{B}\vee\overleftharpoon{C}}\end{matrix} \quad \begin{matrix}\pi_2\\ \vdots\\ \overleftharpoon{\overrightharpoon{B}\vee\overleftharpoon{C}}\vdash\end{matrix}}{\vdash}\ cut \tag{23}$$

The behaviour of this cut is determined by the outermost colour '$\leftharpoonup$'. This behaviour can be simulated by the $(-)^{\bullet}$-translation, provided we use a left-translation for the cut in (23). The translated proof is then as follows:

$$\dfrac{\begin{matrix}\pi_1^{\bullet}\\ \vdots\\ \neg\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})\vdash\end{matrix} \quad \dfrac{\begin{matrix}\pi_2^{\bullet}\\ \vdots\\ \neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})\vdash\end{matrix}}{\vdash\neg\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})}\,\neg_R}{\vdash}\ cut$$

Now $\pi_1^{\bullet}$ will freshly introduce the cut-formula $\neg\neg\neg(B^{\bullet}\vee\neg\neg C^{\bullet})$ only if $\pi_1$ freshly introduces the formula $B\vee C$ (recall that double-negation translations need to preserve the structure of a classical proof). Consequently, the translated proof simulates exactly the behaviour of (23).

What this example shows is that the $(-)^{\bullet}$-translation of $\pi$ can simulate the behaviour where the cut-formula is annotated with the colours

$$\overleftharpoon{\overrightharpoon{B}\vee\overleftharpoon{C}}\ .$$

Note that there are other double negation translation which can be used for a similar simulation of this particular colour-annotation. From our discussion it seems reasonable to expect that one can find corresponding double-negation translations for every possible colour-annotation. Since the formulae $B$ and $C$ can be compound with further colour-annotations inside, we need some flexibility of how to double-negate translate formulae. One has to be able to build into the clause $(B \vee C)^{\bullet} \stackrel{\text{def}}{=} \neg\neg(B^{\bullet} \vee \neg\neg C^{\bullet})$ that the translations of $B$ and $C$ might follow a completely different double-negation scheme. How to do this elegantly is not known to us. On the other hand, we cannot expect complete "freedom" in a double-negation translation as we have to make sure, roughly speaking, that different double-negation translation still fit together in the translated proof. This means we have to make sure that double-negation translations respect the identity-class constraint from the colour-annotations. For example we cannot have an axiom in a translated proof where the double-negation translations disagree as in

$$B^{*} \vdash B^{\circ}$$

This point about fitting double-negation translations together and the identity-class constraint of colours we take as a further evidence that colours and double negation translation must have something to do with each other.

From the observations made above we conjecture that every colour-annotation determining a single normalform of a classical proof can be equally determined by a double negation translation, and every double-negation translation determining a normalform can be equally determined by a colour-annotation. In effect we conjecture that double-negation translations can be simulated by colours and vice versa.

While establishing the simulation properties is a first step for understanding the relation between double-negation translations and colour-annotations, we consider this as not yet giving the complete "picture". For this consider the collection of normalforms of a classical proof determined by double-negation translations and by colour annotations. We conjecture that both are the "same" collection, whereby one needs a (yet unknown) very clever notion of "sameness". Clearly, there are more double-negation translations of a classical proof than there are colour-annotations (there are only finitely many colour-annotations, but there are infinitely many double negation translations as already the translations of atomic formulae as $\neg\neg A$, $\neg\neg\neg\neg A, \ldots$ indicate). For us it is, however, clear that one can group double-negation translations into different classes, where each class corresponds to a colour-annotation.

Let us consider what is necessary to turn these conjectures into theorems. First we have to make precise what we mean by *all* double-negation translation. As seen above, the notion of double-negation translation has to be a generalised version of the traditional notion—like the ones given by Gentzen, Gödel and Kolmogorov—because one needs to take into account the different colour-annotations by varying the double-negations translation when inductively descending a formula. Further, one would ideally like to have a rather general notion of double-negation translation so that one can meaningful state properties for all *possible* double-negation translation. However, we have been unable to formulate such a general notion. Then we have to categorise how the double-negation translations behave under cut-elimination. Finally, one has to establish the "simulation-property"—which unfortunately fails if one takes a very naïve view on sequent-proofs and cut-elimination. One problem is that double-negation translations might introduce auxiliary cuts. Such auxiliary cuts only occur in the double-negated proofs and are thus not taken account of by colour-annotations. Therefore we have to make sure that such auxiliary cuts cannot "mess up" the normalform reached by eliminating cuts in the double-negated proof. This can be relatively easily shown provided the auxiliary cut occurs as the lowermost inference in a proof—then a technique introduced in [19] shows that one can restrict attention to only outermost cuts when calculating the collection of normalforms. However in the general case we have for this not "messing up" only empirical evidence obtained from many calculations (the tools with which we do such calculations are given in [17]). These calculations indeed validate the assumption that the $\neg\neg_R$-rule

$$\frac{\vdash \neg\neg B}{\vdash B}\ \neg\neg_R$$

can be regarded as introducing the formula $B$. To give a proof of this fact, however, we guess one needs a "full-blown" context-lemma using Howe's method. The biggest problem we see in establishing a one-to-one correspondence between the collections of normalforms determined by double-negations translations and by colour-annotations is to find a meaningful equivalence relation on double-negation translations.

Further, the colour-annotations are given for sequent-proofs, for which it is well-known that they make "inessential" differences that would not materialise if we had a natural-deduction formulation or proof-net formulation. However, such formulations have not yet been developed far enough to be useful in getting rid of the inessential differences between classical sequent-proofs. This problem also shows up with intuitionistic sequent-proofs where cut-elimination, despite our working hypothesis stating the contrary, is *not* Church-Rosser. However, for fragments of intuitionistic logic there are already good tools—for example natural deduction and contraction-free sequent-calculi—which provide a canonical notion of what an intuitionistic proof is. Because of all these difficulties, we have, at the moment, to content ourselves with example calculations that, however, all seem to validate the conjecture.

## 5 Conclusion

Although there is plenty of literature on classical logic and double-negation translations, there is surprising little literature that studies the relation between double-negation translations and the process of normalising a classical proof. We find notable exceptions are [1, 5, 6, 11, 13], which however do not give much insight about this relation or consider only special cases. For example, they relate one kind of colour-annotation to one double-negation translation, or consider only specific double-negation translations.

If our conjecture turns out to be true, then one has a very simple characterisation of all double-negation translations in terms of colours. This is desirable, because we find it is rather mysterious how double-negation translation can turn a classical proof into an intuitionistic proof, whose computational interpretation as explained in the introduction is by the Curry-Howard correspondence well-understood, while it is not understood at all for the classical proof we started with (see [2, 19] for two interesting examples extracting computational meaning from a classical proof that because of the the non-determinism cannot be extracted by double negation translations).

The main point we take away from the conjecture is that double-negation translations are not enough to characterise the full computational content of classical proofs, where by *full* computational content we mean the collection of normalforms reachable by the cut-elimination procedure of Urban and Bierman. This is because there are some normalforms that can be reached by this cut-elimination procedure which cannot be reached by any colour-annotation, and by the conjecture also not by any double-negation translation. These "unreachable" normalforms embody a form of non-determinism present in classical logic, but not present in intuitionistic logic. Therefore the conclusion we draw from this is that intuitionistic logic and double-negation translations can only give some hints for understanding the *full* computational meaning of a classical proof.

Completely untouched by our treatment here is the correspondence between double-negation translations and linear logic. The colour annotation also connects to linear logic, where cut-elimination formulated with proof-nets is also Church-Rosser, to the cut-elimination procedure with colours (see [6, 12]).

## References


1. F. Barbanera and S. Berardi. A Symmetric Lambda Calculus for "Classical" Program Extraction. In *Theoretical Aspects of Computer Software*, volume 789 of *LNCS*, pages 495–515. Springer Verlag, 1994.
2. F. Barbanera, S. Berardi, and M. Schivalocchi. "Classical" Programming-with-Proofs in $\lambda^{sym}$: An Analysis of Non-Confluence. In *Theoretical Aspects of Computer Software*, volume 1281 of *LNCS*, pages 365–390. Springer Verlag, 1997.
3. G. Bellin, M. Hyland, E. Robinson, and C. Urban. Proof Theory of Classical Propositional Calculus. (Accepted for publication in Theoretical Computer Science), 2005.
4. S. R. Buss. An Introduction to Proof-Theory. In *Handbook of Proof Theory*, volume 137 of *Studies in Logic and the Foundations of Mathematics*, pages 1–78. North-Holland, 1998.
5. P.-L. Curien and H. Herbelin. The Duality of Computation. In *Conference on Functional Programming*, pages 233–243. ACM Press, 2000.
6. V. Danos, J.-B. Joinet, and H. Schellinx. A New Deconstructive Logic: Linear Logic. *Journal of Symbolic Logic*, 62(3):755–807, 1997.
7. A. G. Dragalin. *Mathematical Intuitionism: Introduction to Proof Theory*, volume 67 of *Translations of Mathematical Monographs*. American Mathematical Society, 1988.
8. R. Dyckhoff and L. Pinto. Permutability of Proofs in Intuitionistic Sequent Calculi. *Theoretical Computer Scince*, 212(1-2):141–155, 1999.
9. J. Gallier. Constructive Logics. Part I: A Tutorial on Proof Systems and Typed $\lambda$-calculi. *Theoretical Computer Science*, 110(2):249–239, 1993.
10. J.-Y. Girard, Y. Lafont, and P. Taylor. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1989.
11. T. Griffin. A Formulae-as-Types Notion of Control. In *Principles of Programming Languages*, pages 47–58. ACM Press, 1990.
12. J.-B. Joinet, H. Schellinx, and L. Tortora de Falco. SN and CR for Free-Style $\text{LK}^{tq}$: Linear Decorations and Simulation of Normalization. *Journal of Symbolic Logic*, 67(1):162–196, 2002.
13. J. Laird. A Deconstruction of Non-Deterministic Cut Elimination. In *Proceedings of the 5th International Conference on Typed Lambda Calculi and Applications*, volume 2044 of *LNCS*, pages 268–282. Springer Verlag, 2001.

14. M. Parigot. $\lambda\mu$-calculus: An Algorithmic Interpretation of Classical Logic. In *Logic Programming and Automated Deduction*, volume 624 of *LNCS*, pages 190–201. Springer Verlag, 1992.
15. G. Pottinger. Normalisation as Homomorphic Image of Cut-Elimination. *Annals of Mathematical Logic*, 12:323–357, 1977.
16. H. Schellinx. *The Noble Art of Linear Decorating*. PhD thesis, Institute for Logic, Language and Computation, University of Amsterdam, 1994. ILLC dissertation series.
17. C. Urban. http://www4.in.tum.de/~urbanc/cut/cutapplet.html.
18. C. Urban. Revisiting Zucker's Work on the Correspondence Between Cut-elimination and Normalisation. (To apear in Advances in Natural Deduction).
19. C. Urban. *Classical Logic and Computation*. PhD thesis, Cambridge University, October 2000.
20. C. Urban. Strong Normalisation for a Gentzen-like Cut-Elimination Procedure. In *Proceedings of the 5th International Conference on Typed Lambda Calculi and Applications*, volume 2044 of *LNCS*, pages 415–429. Springer Verlag, 2001.
21. C. Urban and G. M. Bierman. Strong Normalisation of Cut-Elimination in Classical Logic. *Fundamenta Informaticae*, 45(1–2):123–155, 2001.
22. S. van Bakel, S. Lengrand, and P. Lescanne. The Language X: Circuits, Computations and Classical Logic. In *Proc. of 9th Italian Conference on Theoretical Computer Science (ICTCS)*, volume 3701 of *LNCS*, pages 81–96, 2005.
23. J. Zucker. The Correspondence Between Cut-Elimination and Normalisation. *Annals of Mathematical Logic*, 7:1–112, 1974.