Chapter udf

Normalization

nor.1 Introduction

pt:nor:int:sec

If you have derived a conditional $\varphi \to \psi$, you might be in a position to use it in a derivation of ψ by using $\to$Elim. This would of course also require a derivation δ_1 of φ. Your derivation δ_1 of $\varphi \to \psi$, more likely than not, ends in $\to$Intro, which turns a derivation δ_2 of ψ from the open assumption φ into one of $\varphi \to \psi$. If it does, your final derivation has the form

$$\dfrac{\dfrac{\begin{array}{c}[\varphi]^x\\ \vdots\,\delta_2\\ \psi\end{array}}{\varphi \to \psi}\,x \to\text{Intro} \qquad \begin{array}{c}\vdots\,\delta_2\\ \varphi\end{array}}{\psi} \to\text{Elim}$$

Although your strategy is efficient in that you re-used something you already had (a derivation of $\varphi \to \psi$), your derivation isn't. It seems like a detour to introduce $\varphi \to \psi$ only to immediately eliminate it. There should be a more direct way to derive ψ.

In fact, there always is. "Detours" like the above—an introduction rule immediately followed by an elimination rule—can always be removed from natural deduction derivations. This is called *normalization* and the result is a *normal* derivation (or a derivation *in normal form*).

The proof of this result is, like the cut-elimination theorem in the sequent calculus, somewhat involved and requires us to consider lots of cases. It is harder to prove for classical **N1c** than it is for intuitionistic **N1i**. Just like cut-elimination shows that you can't derive more using the Cut rule in the sequent calculus, normalization shows that you can't prove more things using detours than you can while avoiding them. There are other similarities: normal derivation can be much longer than the shortest derivation with detours of the same result, just like derivations in the sequent calculus can be much shorter with cuts than the shortest cut-free derivation. The sub-formula property,

that you can always derive a result using only sub-formulas of the result in your derivation, follows from normalization for natural deduction just as it follows from cut-elimination for the sequent calculus.

nor.2 Segments and Cuts

pt:nor:seg:sec

An introduction rule followed by an elimination rule is a detour in a derivation, and to normalize a derivation is to eliminate such detours. However, the ∨Elim and ∃Elim make even the definition of "detour we want to eliminate" a bit complicated. The problem is this. Because of these rules, it's not guaranteed that in a detour the elimination rule *immediately* follows the introduction rule. There could be an ∨Elim or ∃Elim in between, e.g.,

$$\dfrac{\dfrac{\dfrac{\exists x\,\chi(x)}{\varphi\to\psi}\,{\scriptstyle y}\quad\dfrac{\dfrac{[\varphi]^x}{\psi}}{\varphi\to\psi}\,{\scriptstyle x}\,\to\text{Intro}}{\psi}\,\exists\text{Elim}\qquad\vdots\;\varphi}{\psi}\,\to\text{Elim}$$

In fact, any number of ∨Elim or ∃Elim can separate the →Intro from the →Elim, e.g.,

$$\dfrac{\dfrac{\dfrac{\exists x\,\chi(x)}{\varphi\to\psi}\,{\scriptstyle y}\quad\dfrac{\dfrac{\theta\vee\alpha\qquad\dfrac{[\varphi]^x}{\psi}}{\varphi\to\psi}\,{\scriptstyle x}\,\to\text{Intro}\qquad\dfrac{[\varphi]^x}{\psi}}{\varphi\to\psi}\,{\scriptstyle x}\,\to\text{Intro}}{\varphi\to\psi}\,\vee\text{Elim}}{\psi}\,\exists\text{Elim}\qquad\vdots\;\varphi}{\psi}\,\to\text{Elim}$$

This example shows that the same →Elim might be involved in multiple detours, even, as it follows more than one →Intro. To account for all these possibilities, we define detours using certain kinds of sequences of formula occurrences in an **N1i**-proof called *segments*. In our example, it's a sequence of occurrences of $\varphi\to\psi$, where a minor premise of ∨Elim or ∃Elim follows its conclusion. In the example, there are two such segments, each containing three occurrences of $\varphi\to\psi$. Here is the official definition.

Definition nor.1. A *segment* in a **N1i** derivation δ is a sequence of occurrences $\varphi_1, \ldots, \varphi_n$ of the same formula φ in δ where

1. φ_1 is not the conclusion of ∨Elim or ∃Elim;

2. φ_n is not a minor premise of $\vee$Elim or $\exists$Elim;

3. For $i = 1, \ldots, n-1$, φ_i is a minor premise of $\vee$Elim or $\exists$Elim and A_{i+1} is the conclusion of that inference.

The *length* of that segment is n.

Not every segment is a target for elimination (a detour). The ones that are are called *cuts*.

Definition nor.2. A *cut* is a segment in which φ_n is the major premise of an elimination inference, and either $n = 1$ and $\varphi_1 = \varphi_n$ is the conclusion of an introduction inference or of $\bot_I$, or $n > 1$.

Note that a cut segment need not begin with the conclusion of an introduction rule. All that's required is that it ends with the major premise of an elimination rule. However, a segment of length 1 does have to be the conclusions of introduction rules to count as a cut.

Definition nor.3. The *rank* of a cut is $\mathrm{dp}(\varphi)$. The *cut rank* $\mathrm{cr}(\delta)$ of a derivation δ is the maximal rank of cuts in δ. A cut is *maximal* in δ if its rank is $\mathrm{cr}(\delta)$, i.e., it has maximal rank. The *cut length* of δ is the sum of the lengths of all its maximal cuts.

nor.3 Permutation Conversions

pt:nor:per:sec

Cuts in **N1i** are segments that end in the major premise of an elimination rule. One case in the normalization proof is to reduce the length of maximal cuts, and with it the cut length of the derivation. A cut of length > 1 can end in many ways depending on whether the last formula occurrence of φ in the cut is the conclusion of ($\vee$Elim or $\exists$Elim) and depending on the elimination inference it is the major premise of.

For instance, if the inferences involved are $\vee$Elim and $\wedge$Elim, $\varphi \equiv \psi \vee \chi$ and there is a sub-derivation δ_1 that ends in:

$$\dfrac{\begin{array}{ccc} & [\theta]^x & [\alpha]^y \\ \vdots\,\delta_2 & \vdots\,\delta_3 & \vdots\,\delta_4 \\ \theta \vee \alpha & \psi \wedge \chi & \psi \wedge \chi \end{array}}{\dfrac{\psi \wedge \chi}{\psi}\,\wedge\mathrm{Elim}_1}\,{}_{x\,y}\vee\mathrm{Elim}$$

The last formula occurrence φ_n in the cut is the conclusion of $\vee$Elim, the lower $\psi \wedge \chi$. The second-to-last formula occurrence φ_{n-1} is one of the minor premises of $\vee$Elim.

To reduce the length of this cut, we can switch the order ("permute") of the ∨Elim and ∧Elim inferences. That is, we replace the subproof δ_1 in δ by the following:

$$\dfrac{\theta \vee \alpha \quad \dfrac{\begin{array}{c}[\theta]^x \\ \vdots \\ \delta_3 \\ \vdots \\ \psi \wedge \chi\end{array}}{\psi}\wedge\text{Elim}_1 \quad \dfrac{\begin{array}{c}[\alpha]^y \\ \vdots \\ \delta_4 \\ \vdots \\ \psi \wedge \chi\end{array}}{\psi}\wedge\text{Elim}_1}{\psi}\vee\text{Elim}\ {\scriptstyle x\,y}$$

The cuts that ended in the premise of ∧Elim below ∨Elim now end in the premise of one of the ∧Elim inferences above the minor premises of ∨Elim, i.e., the length of each has decreased by 1.

Moreover, any cut that lies entirely within δ_2, δ_3, or δ_4, or entirely outside δ_1 in δ is unchanged: it still involves the same formulas, goes through the same inferences, so in particular has the same rank and length as the corresponding cut in δ. So the cut rank of the new derivation has not increased, and the cut length of the new derivation is less than that of the old derivation.

Problem nor.1. After a permutation conversion, the length of at least one cut is reduced by 1. If we permute ∧Elim across an ∨Elim, there are actually two cut segments the length of which decrease, so in this case the cut length of the derivation decreases by at least 2. Can the cut length of a derivation decrease by more than 2 by a single such permutation? Can the cut *rank* decrease?

The process of moving an elimination rule above an ∨Elim or ∃Elim rule is called a *permutation conversion*. The patterns for some of the rule pairs is given in Table nor.1. (The remaining are left as exercises.)

$$\dfrac{\theta \vee \alpha \quad \dfrac{\psi \wedge \chi}{\psi}\wedge\text{Elim}_1 \quad \dfrac{\psi \wedge \chi}{\psi}\wedge\text{Elim}_1}{\psi}\vee\text{Elim}\ {\scriptstyle x\,y} \quad\longrightarrow\quad \dfrac{\theta \vee \alpha \quad \dfrac{\begin{array}{c}[\theta]^x \\ \vdots \\ \psi \wedge \chi\end{array}}{\psi}\wedge\text{Elim}_1 \quad \dfrac{\begin{array}{c}[\alpha]^y \\ \vdots \\ \psi \wedge \chi\end{array}}{\psi}\wedge\text{Elim}_1}{\psi}\vee\text{Elim}\ {\scriptstyle x\,y}$$

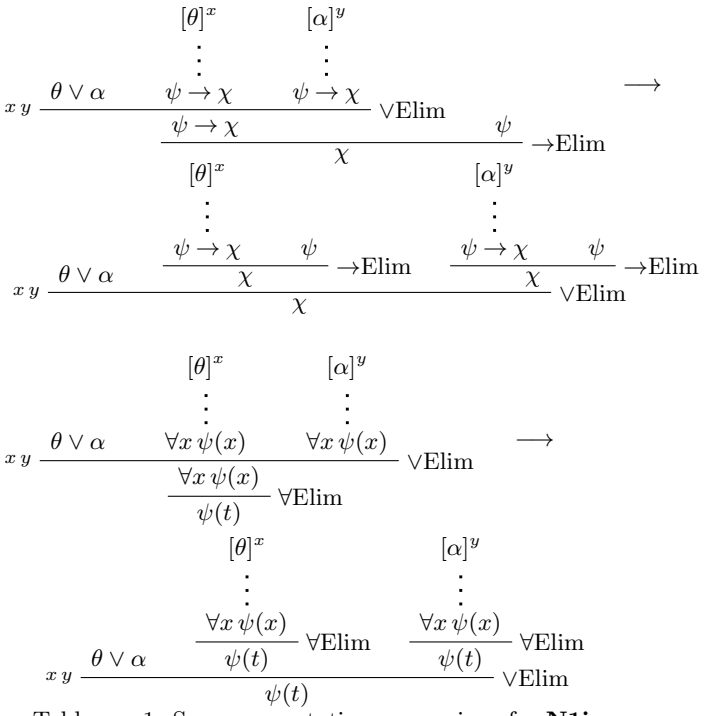


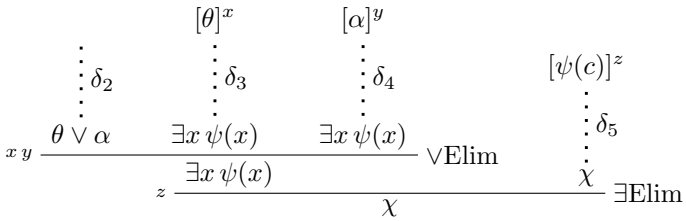
Table nor.1: Some permutation conversions for **N1i**.

pt:nor:per:
tab:permutation-conversions

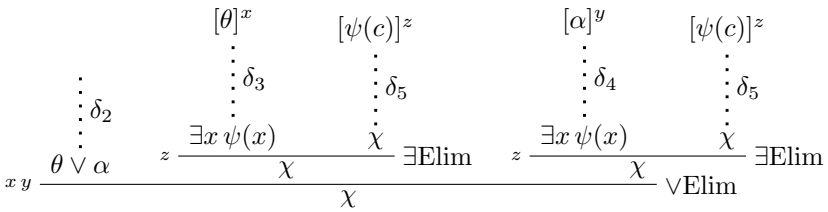
Problem nor.2. Give the permutation conversions for ∨Elim followed by ∨Elim or ¬Elim.

Problem nor.3. Give the permutation conversions for ∃Elim followed by ∃Elim. Explain why no eigenvariable condition is violated in the latter.

We should convince ourselves that when applying a permutation conversion, we don't perhaps violate an eigenvariable condition. Consider



If we replace this by



we might worry that the eigenvariable c might occur in, say, θ. After all, the assumption θ is discharged above the original $\exists$Elim and so in the original derivation it does not occur in an assumption that's open above the major premise of the $\exists$Elim inference, and so there the eigenvariable condition is satisfied. But in the new derivation, θ is not discharged until after the $\exists$Elim inferences, and so the eigenvariable condition would be violated. Recall, however, that we assume our derivations are regular: every eigenvariable is the eigenvariable of a single inference and only occurs above the minor premise of $\exists$Elim in the derivation. In particular, c cannot occur in δ_3 or δ_4.

This example shows something else: Note that the new derivation has two copies of δ_5. This means that if δ_5 contains a maximal cut, we have *not* reduced the cut length!Thus we must take care that we apply a permutation conversion only when we know that there is no maximal cut in δ_5. We will ensure this by always selecting a *topmost* maximal cut, that is a maximal cut such that the sub-derivation (in this case, δ_1) contains no maximal cut that ends above the last inference in the sub-derivation. This rules out the possibility that there are other maximal cuts in δ_5 (or indeed in δ_2, δ_3, or δ_4).

nor.4 Reduction Conversions

pt:nor:red:sec

If a cut in a derivation in **N1i** has length 1, it is a segment consisting of a single formula occurrence, and that formula is both the conclusion of an introduction inference and the major premise of an elimination inference. In the normalization proof, such cuts are "reduced" by replacing the sub-derivation ending in such an introduction/elimination pair by a new derivation in which the pair of inferences has been eliminated. Such a replacement may result in new cuts. However, we can convince ourselves that the new cuts have lower rank than the cut being reduced. Thus, the new derivation either has lower cut length (because a maximal cut of length 1 has been removed) or the cut rank of the derivation has been reduced (if the maximal cut in question was the only maximal cut).

For instance, if the inferences involved are $\to$Intro and $\to$Elim, $\varphi \equiv \psi \to \chi$ and there is a sub-derivation δ_1 that ends in:

$$\dfrac{\dfrac{\begin{array}{c}[\psi]^x\\ \vdots\,\delta_2\\ \chi\end{array}}{\psi \to \chi}\,{\scriptstyle x}\,{\to}\text{Intro} \qquad \begin{array}{c}\vdots\,\delta_3\\ \psi\end{array}}{\chi}\,{\to}\text{Elim}$$

To remove this cut we replace the subproof δ_1 in δ by the following:

$$\begin{array}{l}\vdots\\ \delta_3\\ \vdots\\ \psi\\ \vdots\\ \delta_2[\delta_3/\psi^x]\\ \vdots\\ \chi\end{array}$$

By $\delta_2[\delta_3/\psi^x]$ we mean the derivation that results from δ_2 by replacing every assumption ψ that's labelled x with the entire derivation δ_3. This derivation no longer has open assumptions of the form ψ with label x. It has no other assumptions labelled x, since we assume, as usual, that no two formulas occur as assumptions with the same label. Hence, the new derivation has the same open assumptions as δ_1, every inference in δ_2 that discharges assumptions discharges the same assumptions in $\delta_2[\delta_3/\psi^x]$, and every inference in δ_3 that discharges assumptions results in possibly multiple copies, which discharge the corresponding assumptions in $\delta_2[\delta_3/\psi^x]$.

Any cut that lies entirely within δ_2, δ_3, or δ_4, or entirely outside δ_1 in δ is unchanged: it still involves the same formulas, goes through the same inferences, so in particular has the same rank and length as the corresponding cut in δ. We have removed a maximal cut of length 1 and so either have reduced the cut length, or reduced the cut rank (if the cut length of δ was 1 and this was the only maximal cut).

However, two things may happen:

1. By replacing all assumptions ψ^x by δ_3, we have multiplied all the cuts in δ_3. If some of those were maximal cuts, the cut length of the new derivation may be greater than the cut length of the old derivation. We have to guarantee that this does not happen. We do so by applying reduction conversions only to *rightmost* maximal cuts, that is, to cuts which are such that no maximal cut lies on a branch through δ to the right of the cut being reduced. If there were a cut in δ_3 it would be to the right of the cut being reduced here, and so we wouldn't be working on a rightmost cut. By always selecting rightmost cuts, we guarantee that either we decrease the cut length of the derivation, or even reduce the cut rank.

2. If an assumption ψ^x in δ_2 is the major premise of an elimination inference, and the last inference of δ_3 is $\vee$Elim, $\exists$Elim, or an introduction inference, we have generated a new cut by grafting δ_3 onto δ_2 at such an assumption. What was just an assumption in δ_2 is now a cut of length 1 (the conclusion of an introduction inference and major premise of an elimination inference) or the last formula occurrence of a cut segment (major premise of an elimination inference and conclusion of $\vee$Elim or $\exists$Elim). If B^x was a minor premise of $\vee$Elim or $\exists$Elim, it already was the first

formula occurrence of a segment, and possibly of a cut segment. In the new derivation that segment may now be longer. However, observe that the formula making up this new cut or existing cut segment is ψ, and $\mathrm{dp}(\psi) < \mathrm{dp}(\psi \to \chi)$. Hence, although we have potentially added cuts or increased the length of cuts, none of these are maximal cuts. Thus we have either decreased the cut rank, or, if there are cuts of the same rank remaining, have decreased the cut length.

The process of reducing cut of length 1 is called a *reduction conversion* (or *detour conversion*). The patterns for some of the rule pairs is given in Table nor.2. (The remaining are left as exercises.)

Problem nor.4. Give the reduction conversion for ¬Intro followed by ¬Elim.

There is one case still to be considered. The definition of cut includes the case where the length of the cut is 1 and φ is the conclusion of $\bot_I$ as well as the major premise of an elimination inference. This case is dealt with in a similar way as when φ is the conclusion of an introduction rule. For instance, replace

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_2\\ \bot\end{array}}{\psi \to \chi} \quad \begin{array}{c}\vdots\,\delta_3\\ \psi\end{array}}{\chi} \to\text{Elim} \quad \text{by} \quad \begin{array}{c}\vdots\,\delta_2\\ \bot\\ \chi\end{array}$$

We must be a little careful in the cases where $\varphi \equiv (\psi \vee \chi)$ or $\varphi \equiv \exists x\, \psi(x)$. E.g., if we replaced

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_2\\ \bot\end{array}}{\psi \vee \chi}\,\bot_I \quad \begin{array}{c}[\psi]^x\\ \vdots\,\delta_3\\ \theta\end{array} \quad \begin{array}{c}[\chi]^x\\ \vdots\,\delta_4\\ \theta\end{array}}{\theta} \vee\text{Elim} \quad \text{by} \quad \dfrac{\begin{array}{c}\vdots\,\delta_2\\ \bot\end{array}}{\theta}\,\bot_I$$

we might be introducing a new cut with cut formula θ, and there is no guarantee that $\mathrm{dp}(\theta) < \mathrm{dp}(\psi \vee \chi)$! So we must proceed the same way here as we did for an ∨Intro/∨Elim reduction and replace with

$$\dfrac{\begin{array}{c}\vdots\,\delta_2\\ \bot\end{array}}{\psi}\,\bot_I \quad \begin{array}{c}\vdots\,\delta_3\\ \theta\end{array} \qquad \text{or} \qquad \dfrac{\begin{array}{c}\vdots\,\delta_2\\ \bot\end{array}}{\chi}\,\bot_I \quad \begin{array}{c}\vdots\,\delta_4\\ \theta\end{array}$$

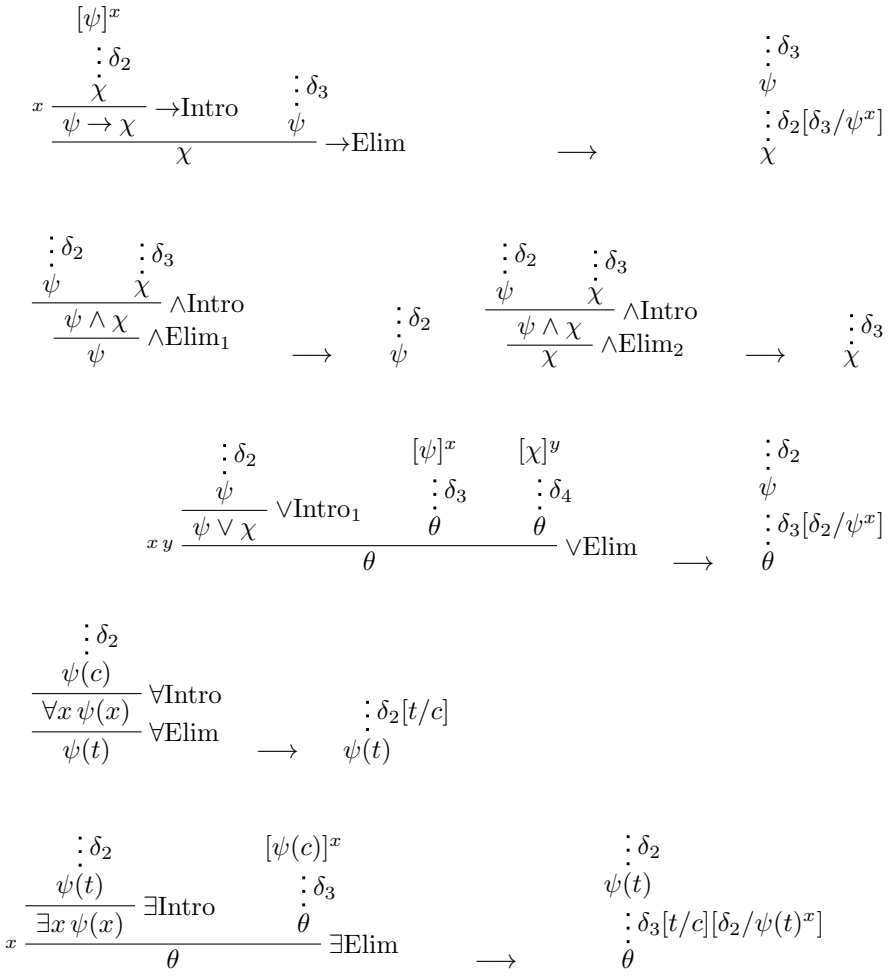


Table nor.2: Some reduction conversions for **N1i**.

pt:nor:red:
tab:reduction-conversions

nor.5 The Normalization Theorem

A derivation is in *normal* form if it has no cut segments. Clearly this is the case iff we can apply neither a permutation conversion nor a reduction conversion to it. To *normalize* a derivation means to transform it into a normal proof by applying permutation and reduction conversions until no cuts remain. That this is always possible is the *normalization theorem*. pt:nor:thm:sec

Theorem nor.4. *Any* **N1i**-*derivation* δ *normalizes, i.e., can be transformed into a derivation* δ' *without cuts.* pt:nor:thm:thm:normalization

Proof. By induction on the cut rank and length of δ. The induction hypothesis is that any derivation which either has lower cut rank, or the same cut rank but lower cut length, normalizes.

If the cut length is 0, there are no cuts, δ is already normal. So assume that cut rank and cut length of δ are > 0. Pick a topmost, rightmost maximal cut segment. If the cut segment has length > 1, apply a permutation conversion. If the cut segment has length 1, apply the corresponding reduction conversion. This results in a new derivation which either has the same cut rank but a lower cut length, or a lower cut rank, and the inductive hypothesis applies. □

The strategy used in the above proof (always deal with a topmost rightmost maximal cut) suggests a specific order in which conversions must be applied to a derivation in order to result in a normal derivation. Surprisingly enough, the order actually does not matter. Applying permutation and reduction conversions in any order eventually results in a normal proof.

The definition of segment and cut can easily be formulated for **N2i**, and the permutation and reduction conversions for **N1i** directly translate to **N2i**. Consequently, the normalization theorem also holds for **N2i**.

Theorem nor.5. *Any* **N2i**-*derivation* δ *normalizes, i.e., can be transformed into a derivation* δ' *without cuts.* pt:nor:thm:thm:normalization-N2i

nor.6 Translating Between Normal N2i-Derivations and G2i

Derivations in **N2i** can be translated to derivations in **G2i** and vice versa. First of all, the result of translating cut-free **2i** derivations results in normal **N2i**-derivations. pt:nor:tr:sec

To state the result we need a bit of terminology to let us relate antecendents of sequents in **N2i** and in **G2i** more easily: a set Γ' of labelled formulas *corresponds to* the multiset Γ of unlabelled formulas if, for every formula φ, the number of elements of Γ' of the form $x : \varphi$ is less than or equal to the multiplicity of φ in Γ.

Corollary nor.6. *If* $\mathbf{G2i} \vdash \Gamma \Rightarrow \Delta$ *then there is a normal* **N2c**-*derivation* δ *of* $\Gamma' \Rightarrow \chi$*, where the set* Γ' *of labelled formulas corresponds to the multiset* Γ*,*

i.e., for every formula φ, the number of elements of Γ' of the form $x : \varphi$ is less than or equal to the multiplicity of φ (i.e., the number of copies of φ) in Γ.

Proof. By inspection of the proof of **??**: we add introduction rules to the end of derivations, and elimination rules only at the top of derivations, so never introduce an introduction rule above an elimination rule. □

The translation of the Cut rule given in **??**, however, can introduce cuts in the **N2i**-derivation δ. If a **G2i** derivation ends in Cut with immediate subderivations π_1 and π_2 of $\Gamma_1 \Rightarrow \varphi$ and $\varphi, \Gamma_2 \Rightarrow \Delta_2$, respectively, we graft the translation δ_1 of π_1 onto the axioms $x : \varphi \Rightarrow \varphi$ occurring in the translation δ_2 of π_2. If δ_1 ends in an introduction rule with principal formula φ and the axiom $x : \varphi \Rightarrow \varphi$ in δ_2 is the major premise of an elimination rule, this would result in a cut in δ.

It is also possible to translate derivations in **N2i** to derivations in **G2i**. In the proof of **??**, the translation of ∧Elim, →Elim, ¬Elim, ∀Elim results in Cut inferences in the resulting **G2i** derivation. If the **N2i**-derivation we start from, this can, however, be avoided.

Let's call a list $S_1, \ldots, S_n$ of occurrences of sequents in a **N2i**-derivation δ a *branch* if S_1 is an axiom, and whenever S_i is a major premise of an inference, S_{i+1} is its conclusion. In other words, a branch traces sequents downwards in δ from axioms, but stops at minor premises of elimination rules. A *main branch* of δ is a branch where S_n is the end-sequent. Every derivation δ must have at least one main branch, since every rule has at least one major premise (only ∧Intro has two).

pt:nor:tr: prop:normal-N2i-to-G2i

Proposition nor.7. *If δ is a normal* **N2c** *or* **N2i**-*derivation of $\Gamma \Rightarrow \varphi$ then there is a (cut-free)* **G2c**-*derivation (***G2i**-*derivation) π of $\Gamma' \Rightarrow \Delta$, where Γ' is the multiset of formulas resulting from Γ by removing labels, and $\Delta = \{\varphi\}$ if φ is not $\bot$, and $\Delta = \emptyset$ if it is.*

Proof. This time, induction is on the number of inferences in the derivation δ of $\Gamma \Rightarrow \varphi$. If there is no inference in δ, it is an axiom $x : \varphi \Rightarrow \varphi$, and π is the corresponding axiom $\varphi \Rightarrow \varphi$ or $\bot \Rightarrow$ if $\varphi \equiv \bot$.

If δ ends in an inference rule, we distinguish cases:

1. If R is an introduction rule or $\bot_I$, the translations proceeds exactly as in the proof of **??**, and does not require Cut.

   If the last rule R of δ is an elimination rule, observe the following:

   a) A main branch of δ passes through no introduction rules, as otherwise there would be an introduction rule or $\bot_I$ followed by an elimination rule along the main branch, and δ would not be normal.

   b) Since there are no introduction rules along main branches of δ, there is only one main branch. (Only ∧Intro has two major premises, so multiple main branches would have to go through the premises of a ∧Intro rule.)

c) If the main branch includes the major premise of ∨Elim or ∃Elim, there is only one such rule and it is the last one, i.e., R. (Otherwise the main branch would contain an ∨Elim or ∃Elim rule where the conclusion of the rule is major premise of an elimination rule, and the proof would not be normal as we could apply a permutation conversion.)

d) No elimination rules other than ∨Elim and ∃Elim discharge assumptions, i.e., change the left context of sequents. ∨Elim and ∃Elim only discharge assumptions above a minor premise, i.e., change the left context only of their minor premises. In other words, if a sequent S_i on the main branch of δ has context Γ_1, the conclusion S_{i+1} of the inference of which S_i is the major premise has context $\Gamma_1' \supseteq \Gamma_1$.

e) Consequently, if $x : \chi \Rightarrow \chi$ is the topmost sequent S_1 in the main branch, $x : \chi \in \Gamma$.

We now distiguish cases according to the form of χ. Since every S_i in the main branch is major premise of an elimination rule, this also applies to $x : \chi \Rightarrow \chi$.

2. $\chi \equiv \theta \wedge \alpha$. Then δ has the form

$$\dfrac{\dfrac{x : \theta \wedge \alpha \;\Rightarrow\; \theta \wedge \alpha}{x : \theta \wedge \alpha \;\Rightarrow\; \theta}\,\wedge\text{Elim}_1}{\begin{array}{c}\vdots\\ x : \theta \wedge \alpha, \Gamma_1 \;\Rightarrow\; \varphi\end{array}}$$

The main branch includes $x : \theta \wedge \alpha \Rightarrow \theta$ contains no major premise of →Intro, ¬Intro nor a minor premise of ∨Elim or ∃Elim, the context formulas in $x : \theta \wedge \alpha$ remain unchanged along the main branch. This explains why the context of the end-sequent must contain $x : \theta \wedge \alpha$. Moreover, we can turn the derivation δ into a derivation δ_1 by replacing the sub-derivation ending in ∧Elim by $x : \theta \Rightarrow \theta$ and replacing $x : \theta \wedge \alpha$ in the context of every sequent along the main branch by $x : \theta$, i.e., turn

$$\dfrac{\dfrac{x : \theta \wedge \alpha \;\Rightarrow\; \theta \wedge \alpha}{x : \theta \wedge \alpha \;\Rightarrow\; \theta}\,\wedge\text{Elim}_1}{\begin{array}{c}\vdots\\ x : \theta \wedge \alpha, \Gamma_1 \;\Rightarrow\; \varphi\end{array}} \qquad \text{into} \qquad \dfrac{\begin{array}{c}x : \theta \;\Rightarrow\; \theta\\ \vdots\\ x : \theta, \Gamma_1 \;\Rightarrow\; \varphi\end{array}}{}$$

The induction hypothesis applies to δ_1, since the number of inferences in δ is the number of inferences in δ' plus 1.

By induction hypothesis, there is a **G2i**-derivation π_1 of $x : \theta, \Gamma_1 \Rightarrow \varphi$. We obtain the derivation π by applying $\wedge$L:

$$\dfrac{\begin{array}{c}\vdots\,\pi_1\\ \theta, \Gamma_1' \Rightarrow \varphi\end{array}}{\theta \wedge \alpha, \Gamma_1' \Rightarrow \varphi}\,\wedge\mathrm{L}$$

If $\varphi \equiv \bot$ we insert a WR, e.g,

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\pi_1\\ \theta, \Gamma_1' \Rightarrow\end{array}}{\theta \wedge \alpha, \Gamma_1' \Rightarrow}\,\wedge\mathrm{L}}{\theta \wedge \alpha, \Gamma_1' \Rightarrow \varphi}\,\mathrm{WR}$$

3. If $\chi \equiv \theta \vee \alpha$, it must be the major premise of $\vee$Elim, and this inference must be the last inference R on the main branch. In other words, δ has the form:

$$\dfrac{x : \theta \vee \alpha \Rightarrow \theta \vee \alpha \qquad \begin{array}{c}\vdots\,\delta_1\\ y : \theta, \Gamma_1 \Rightarrow \varphi\end{array} \qquad \begin{array}{c}\vdots\,\delta_2\\ z : \alpha, \Gamma_2 \Rightarrow \varphi\end{array}}{x : \theta \vee \alpha, \Gamma_1, \Gamma_2 \Rightarrow \varphi}\,\vee\mathrm{Elim}$$

The induction hypothesis applies to the derivations δ_1 and δ_2; we obtain **G2i**-derivations π_1 and π_2 of $\theta, \Gamma_1' \Rightarrow \varphi$ and $\alpha, \Gamma_2' \Rightarrow \varphi$. We apply $\vee$L to get π:

$$\dfrac{\begin{array}{c}\vdots\,\pi_1\\ \theta, \Gamma_1' \Rightarrow \varphi\end{array} \qquad \begin{array}{c}\vdots\,\pi_2\\ \alpha, \Gamma_2' \Rightarrow \varphi\end{array}}{\theta \vee \alpha, \Gamma_1', \Gamma_2' \Rightarrow \varphi}\,\vee\mathrm{L}$$

Should $\vee$Elim not discharge $y : \theta$ or $z : \alpha$ (i.e., these aren't present in the context of the minor premises), or should φ happen to be $\bot$, we have to add some WL and WR. For instance,

$$\dfrac{\dfrac{\dfrac{\begin{array}{c}\vdots\,\pi_1\\ \Gamma_1' \Rightarrow\end{array}}{\theta, \Gamma_1' \Rightarrow}\,\mathrm{WL} \qquad \dfrac{\begin{array}{c}\vdots\,\pi_2\\ \Gamma_2' \Rightarrow\end{array}}{\alpha, \Gamma_2' \Rightarrow}\,\mathrm{WL}}{\theta \vee \alpha, \Gamma_1', \Gamma_2' \Rightarrow}\,\vee\mathrm{L}}{\theta \vee \alpha, \Gamma_1', \Gamma_2' \Rightarrow \varphi}\,\mathrm{WR}$$

4. $\chi \equiv \theta \to \alpha$. Then δ has the form

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\delta_1\end{array}}{x:\theta\to\alpha \;\Rightarrow\; \theta\to\alpha \qquad \Gamma_1 \;\Rightarrow\; \theta}\;{\to}\text{Elim}}{\begin{array}{c}x:\theta\to\alpha,\Gamma_1 \;\Rightarrow\; \alpha\\ \vdots\,\delta_2\\ x:\theta\to\alpha,\Gamma_1,\Gamma_2 \;\Rightarrow\; \varphi\end{array}}$$

Here note that since the main branch, which includes $x:\theta\to\alpha,\Gamma_1\Rightarrow\alpha$, contains no major premise of $\to$Intro, $\neg$Intro nor a minor premise of $\vee$Elim or $\exists$Elim, the context formulas in $x:\theta\to\alpha,\Gamma_1$ remain unchanged along the main branch. This explains why the context of the end-sequent must contain $x:\theta\to\alpha,\Gamma_1$. Moreover, we can turn the derivation fragment δ_2 into a correct derivation δ_2' by replacing the sub-derivation ending in $\to$Elim by $x:\alpha\Rightarrow\alpha$ and replacing $x:\theta\to\alpha,\Gamma_1$ in the context of every sequent along the main branch by $x:\alpha$, i.e., turn

$$\begin{array}{c}x:\theta\to\alpha,\Gamma_1 \;\Rightarrow\; \alpha\\ \vdots\,\delta_2\\ x:\theta\to\alpha,\Gamma_1,\Gamma_2 \;\Rightarrow\; \varphi\end{array} \quad\text{into}\quad \begin{array}{c}x:\alpha \;\Rightarrow\; \alpha\\ \vdots\,\delta_2'\\ x:\alpha,\Gamma_2 \;\Rightarrow\; \varphi\end{array}$$

The induction hypothesis applies to δ_1 and δ_2', since the number of inferences in δ is the sum of the numbers of inferences in δ_1 and δ_2, plus 1 to account for the $\to$Elim inference at the start of the main branch. (If that inference is also the last inference R in δ then δ_1 contains 0 inferences.)

By induction hypothesis, there are **G2i**-derivations π_1 of $\Gamma_1\Rightarrow\theta$ and π_2 of $\alpha,\Gamma_1'\Rightarrow\varphi$. We obtain the derivation π by applying $\to$L:

$$\dfrac{\begin{array}{c}\vdots\,\pi_1\\ \Gamma_1' \;\Rightarrow\; \theta\end{array} \qquad \begin{array}{c}\vdots\,\pi_2\\ \alpha,\Gamma_2' \;\Rightarrow\; \varphi\end{array}}{\theta\to\alpha,\Gamma_1',\Gamma_2' \;\Rightarrow\; \varphi}\;{\to}\text{L}$$

If θ and/or $\varphi\equiv\bot$ we insert one or two WR, e.g,

$$\dfrac{\dfrac{\begin{array}{c}\vdots\,\pi_1\\ \Gamma_1' \;\Rightarrow\end{array}}{\Gamma_1' \;\Rightarrow\; \theta}\;\text{WR} \qquad \dfrac{\begin{array}{c}\vdots\,\pi_2\\ \alpha,\Gamma_2' \;\Rightarrow\end{array}}{\alpha,\Gamma_2' \;\Rightarrow\; \varphi}\;\text{WR}}{\theta\to\alpha,\Gamma_1',\Gamma_2' \;\Rightarrow\; \varphi}\;{\to}\text{L}$$

The remaining cases are treated similarly and are left as exercises. □

Corollary nor.8. *Every formula occuring in a normal* **N1i** *or* **N2i**-*derivation is a sub-formula of the end-formula or end-sequent.*

Proof. By Proposition nor.7, every normal **N2i**-derivation can be translated into a **G2i**-derivation. An inspection of the proof shows that every formula occurring in the **N2i**-derivation also occurs in the **G2i**-derivation. **G2i**, lacking the Cut rule, has the sub-formula property. □

Photo Credits

Bibliography