PAOLO PISTONE
LUCA TRANCHINI
MATTIA PETROLO

The Naturality of Natural Deduction (II): On Atomic Polymorphism and Generalized Propositional Connectives

Abstract. In a previous paper (of which this is a prosecution) we investigated the extraction of proof-theoretic properties of natural deduction derivations from their impredicative translation into System F. Our key idea was to introduce an extended equational theory for System F codifying at a syntactic level some properties found in parametric models of polymorphic type theory. A different approach to extract proof-theoretic properties of natural deduction derivations was proposed in a recent series of papers on the basis of an embedding of intuitionistic propositional logic into a predicative fragment of System F, called *atomic* System F. In this paper we show that this approach finds a general explanation within our equational study of second-order natural deduction, and a clear semantic justification in terms of parametricity.

Keywords: Second-order logic, Propositional quantification, Identity of proofs, Russell–Prawitz translation, Finite polynomial functor, Permutative conversions, Predicativity.

1. Introduction

Russell was the first to observe that propositional connectives like disjunction and conjunction can be defined using only implication and propositional quantification. Later Prawitz, in his monograph on natural deduction, showed how the natural deduction system for intuitionistic propositional logic (henceforth NI) can be embedded into the implicational fragment of second-order propositional intuitionistic logic (also known as *System F*, and here referred to as NI²). We will refer to this embedding as the Russell–Prawitz translation (shortly RP-translation).

Taking inspiration from this embedding, some recent work in the field of proof-theoretic semantics (see for instance [18] and [9]) has suggested the use of NI² as a convenient meta-language to investigate the proof theory

Presented by **Heinrich Wansing**; Received September 1, 2019

of propositional connectives. However, this way of looking at NI^2 faces two kinds of difficulties.

The equivalence-preservation problem In proof-theoretic semantics, one is not only concerned with *provability*—i.e. with whether there is a derivation of a given formula in a certain system—but also with *identity of proofs*—i.e. with whether two distinct derivations of the same formula can be viewed as different syntactic representations of *the same* proof (understood as an abstract object).

A common way to characterize identity of proofs is by declaring two derivations *equivalent* when they converge, under the usual conversions used for normalization, to the same normal derivation. Equivalent derivations are then taken to represent the same proof. This intuition can be made precise using the categorical semantics for natural deduction systems. For instance NI can be interpreted in any *bi-cartesian closed category*, with equivalent derivations being mapped onto the same morphism.

When not only provability but also identity of proofs is considered, the RP-translation appears as not entirely satisfactory. In fact, equivalent derivations in NI need not translate into equivalent derivations in NI^2 . Although the translation works for the equivalence induced by β -conversions only, it fails for the one induced by η -conversions and permutations, here referred to as γ -conversions. In categorical terms, the RP-translation of, say, a disjunction, is not interpreted as a co-product in every categorical model of NI^2 , but only as a “weak” variant of it. We will refer to this fact as the *equivalence-preservation* problem of the RP-translation.

In a previous paper, of which the present one is a follow-up, we explored a solution to the equivalence-preservation problem based on the fact that the RP-translation of conjunctions and disjunctions *does* yield categorical products/co-products in the class of *parametric models* of NI^2 [1, 14, 25]. With the goal of making this result accessible to the proof theory community at large, in [33] we provided a purely syntactic reconstruction of it: we introduced an equational theory extending the one arising from the usual β - and η -conversions for NI^2 -derivations using a new class of conversions—that we called ε -conversions—expressing a *naturality* condition for NI^2 -derivations which is satisfied by all parametric models of NI^2 . We then showed that the RP-translation does preserve the full equivalence of NI -derivations as soon as NI^2 -derivations are considered under this stronger equivalence.¹

¹In the functorial semantics of System F ([1, 14]), NI^2 -derivations are actually interpreted as *dinatural* transformations. The reason to focus on naturality, rather than on the more general notion of dinaturality, is briefly discussed in Section 6.

Impredicative versus predicative translations A second difficulty is of a foundational nature and stems from the fact that the RP-translation and, more generally, the second-order encoding of *inductive types* (e.g. the types of natural numbers and well-founded trees) inside NI^2 are *impredicative*. In fact, the embedding of NI into NI^2 requires the full power of second-order quantification: in the elimination rule for the second-order quantifier $\forall E$

$$\frac{\forall X.A}{A[B/X]} \forall E$$

no restriction can be imposed on the choice of the formula B (called the *witness* of the rule application).

A solution to this problem can be found in a recent series of papers by Fernando Ferreira and Gilda Ferreira, who proposed a variant of the RP-translation (to which we will refer to as FF-translation) which encodes NI in *atomic* System F (here referred to as NI_{at}^2), a weak predicative fragment of NI^2 in which the witnesses of $\forall E$ are required to be atomic formulas. A further refinement of the FF-translation was later proposed by José Espírito Santo and Gilda Ferreira in [5] (we will refer to it as the ESF-translation).

Besides being predicative, the FF- and ESF-translations have another significant advantage over the RP-translation: they do preserve the equivalence arising not only from β -conversions, but also from η - and γ -conversions for disjunction and $\perp$ [5, 8, 11].² For these reasons the predicative translations were advocated in [9] as evidence in favor of taking NI_{at}^2 as a convenient meta-language to investigate propositional connectives.

From impredicative to atomic polymorphism through ε -conversions In this paper we exploit the equational framework we developed in the first paper of this series to investigate the predicative translations of intuitionistic propositional logic into NI_{at}^2 . In particular, we show that the usual, impredicative, translation can be related to the predicative ones using the ε -conversion, thus providing a semantic explanation ultimately relying on parametricity for some syntactic results based on atomic polymorphism.

Our first observation is that the results of Ferreira and co-authors do not hold only for $\vee$ and $\perp$, but for the class of connectives that are definable in NI by arbitrarily composing $\wedge$, $\vee$, $\top$ and $\perp$ (to describe such connectives we borrow another concept from the toolbox of category theory, that of a *finite polynomial functor* [13]).

²Actually, preservation of η -conversions fails for conjunction.

By extending the RP-translation to a natural deduction system for this class of propositional connectives (called NI^p), we are led to consider another fragment of NI^2 , that we call the *Russell–Prawitz fragment* (NI_{RP}^2). Unlike the atomic fragment NI_{at}^2 , the fragment NI_{RP}^2 is impredicative, since no restriction is imposed on the witnesses of the applications of $\forall\text{E}$.

Nonetheless, we show that every derivation in NI_{RP}^2 can be “atomized”, i.e. it can be mapped onto a derivation in NI_{at}^2 with the same conclusion and the same assumptions, by applying instances of η -expansion and the ε -conversion. By composing the RP-translation from NI^p to NI_{RP}^2 with the atomization from NI_{RP}^2 to NI_{at}^2 one thereby obtains another predicative translation from NI^p to NI_{at}^2 (we call it the ε -translation) which only differs from the FF- and ESF-translation by some β -reduction steps.

An immediate consequence of this fact is that the RP-translation and its three predicative variants are all equivalent modulo β -, η - and ε -conversions. Thus, under the notion of identity of proofs induced by ε -conversions, all translations of a given propositional derivation are different syntactic descriptions of the same second-order proof (that is, all these translations interpret an NI^p -derivation as the same morphism in all parametric models of NI^2).

Limitations of the predicative translations The comparison of the impredicative and predicative translations leads us to highlight two limitations of the approach based on atomic polymorphism: first, the predicative translations do not preserve the full η -rule needed to interpret disjunction as a categorical co-product, and thus fail to provide a full solution to the equivalence-preservation problem. Moreover, we show that once $\forall\text{E}$ is restricted to atomic witnesses it is not possible to prove the logical equivalence between a propositional formula and its second-order translation (with the terminology of [22], this means that connectives are not *strongly* definable, but only *weakly* definable in NI_{at}^2).

Goals and plan of the paper One of the motivations for the previous and present papers is that of making some ideas underlying the categorical semantics of System F accessible to the proof-theoretic community at large, and to show that these ideas can be fruitfully connected with strands of research arisen within more philosophically-oriented areas of proof theory. This was the reason for reformulating in the first paper categorical notions such as functors and natural transformations in the language of natural deduction, at the expenses of typographic conciseness.

To keep the presentation compact and readable for the largest audience, we chose to present the main results of the paper using the natural deduction notation and restricting the attention only to the case of a particular ternary connective $\bullet(A, B, C)$. Full proofs for the whole class of connectives we consider are postponed to a (large) technical appendix written using the drastically more economical λ -calculus notation.

In Section 2, we introduce the natural deduction calculus NI^P for the class of propositional connectives we intend to investigate and a fragment of NI^2 , that we call the *Russell–Prawitz fragment* (noted NI_{RP}^2). Both NI^P and NI_{RP}^2 are inspired by the notion of finite polynomial functor from category theory, and we introduce a generalization of the usual RP-translation as a derivability-preserving embedding between these two systems. In Section 3 we recall the framework introduced in our previous paper to describe functors and natural transformations within natural deduction, based on the ε -conversion. In Section 4 we generalize the FF- and the ESF-translations to NI^P and we investigate their relationship to the RP-translation. To do this, we first show how the FF-translation can be analyzed as the composition of the RP-translation and of an embedding from the fragment NI_{RP}^2 into NI_{at}^2 that we call FF-atomization, and then defining an alternative embedding from NI_{RP}^2 into NI_{at}^2 using the ε -conversions, the ε -atomization. In Section 5 we discuss some limitations of the predicative translations for the study of identity of proofs of propositional connectives. In Section 6 we briefly summarize the results of the paper, we draw some connections with related work, and we suggest further directions of investigation. Finally, the rich appendix provides full proofs (in λ -calculus notation) of the results discussed or simply sketched in the main text.

2. Polynomial Connectives and Their RP-Translation

In this section we introduce a formal framework for natural deduction which extends the one from [33] to a more general class of propositional connectives.

2.1. Polynomial Connectives

As suggested in the previous paper (cf. [33] Section 4.3), the results we are concerned with are not limited to the standard intuitionistic connectives, but scale smoothly to a wider class of connectives investigated in proof-theoretic semantics (see e.g. [24, 26]). These are the connectives that are obtained by freely composing $\vee$, $\wedge$, $\top$ and $\perp$, such as the ternary connective $\bullet(A_1, A_2, A_3)$

definable as $(A_1 \wedge A_2) \vee A_3$, whose introduction and elimination rules are as follows:

$$\frac{A_1 \quad A_2}{\bullet(A_1, A_2, A_3)} \bullet I_1 \quad \frac{A_3}{\bullet(A_1, A_2, A_3)} \bullet I_2 \quad \frac{\bullet(A_1, A_2, A_3) \quad \frac{[A_1][A_2]}{C} \quad \frac{[A_3]}{C}}{C} \bullet E$$

In general, each such connective is definable in NI as $\bigvee_{p=1}^n \bigwedge_{q=1}^{m_p} A_{pq}$ for some choice of natural numbers n, m_p s and formulas A_{pq} s. We take here the viewpoint of treating such connectives as logical primitives.

Borrowing ideas from the theory of *finite polynomial functors* [13], each such connective can be described as determined by three finite lists $\mathcal{I}, \mathcal{J}, \mathcal{K}$ (to be thought of as lists of indices) and two functions $f : \mathcal{J} \mapsto \mathcal{I}$ and $g : \mathcal{J} \mapsto \mathcal{K}$ that we depict in a diagram as follows:³

$$\mathcal{I} \xleftarrow{f} \mathcal{J} \xrightarrow{g} \mathcal{K}$$

Any such diagram determines what we will call a *polynomial connective* to be indicated with $\dagger^{(f,g)}$, or simply $\dagger$ when f, g are clear from the context, in the following way:

- The length $|\mathcal{I}|$ of $\mathcal{I}$ measures the arity of $\dagger$, so that when $\dagger$ is applied to an $\mathcal{I}$ -indexed list of formulas $\langle A_i \rangle_{i \in \mathcal{I}}$ ⁴ one obtains a new formula $\dagger \langle A_i \rangle_{i \in \mathcal{I}}$.
- The length $|\mathcal{K}|$ of $\mathcal{K}$ is the number of distinct introduction rules of $\dagger$.
- Any element k in $\mathcal{K}$ determines a sublist of $\mathcal{J}$, namely the list of all $j \in \mathcal{J}$ such that $g(j) = k$, that we indicate with $g^{-1}(k)$ and whose length $|g^{-1}(k)|$ is the number of premises of the k -th introduction rule of $\dagger$.⁵

Using the functions f and g we can describe the introduction and elimination rules for $\dagger$ as follows. Given an $\mathcal{I}$ -indexed list of formulas $\langle A_i \rangle_{i \in \mathcal{I}}$, the k -th introduction rule $\dagger I_k$ for $\dagger$, allows us to infer $\dagger \langle A_i \rangle_{i \in \mathcal{I}}$ from the list of premises $\langle A_{f(j)} \rangle_{j \in g^{-1}(k)}$. Given $\dagger \langle A_i \rangle_{i \in \mathcal{I}}$ and a $\mathcal{K}$ -indexed list of derivations of an arbitrary formula C from (respectively) the premises of the k -th introduction rule $\langle A_{f(j)} \rangle_{j \in g^{-1}(k)}$, we can infer C thereby discharging in the

³In the language of category theory this configuration describes a unary finite polynomial functor, which is the reason for our terminological choice.

⁴ $\langle A_i \rangle_{i \in \mathcal{I}}$ abbreviates the sequence of formulas $A_{i_1} \dots A_{i_n}$ when $\mathcal{I} = \langle i_1, \dots, i_n \rangle$.

⁵Thus $\mathcal{J}$ can be seen as a family of lists indexed by the elements of $\mathcal{K}$.

k -th derivation of C the assumptions $\langle A_{f(j)} \rangle_{j \in g^{-1}(k)}$. We depict the rules as follows:

$$\left\langle \frac{\langle A_{f(j)} \rangle_{j \in g^{-1}(k)}}{\dagger \langle A_i \rangle_{i \in \mathcal{I}}} \dagger I_k \right\rangle_{k \in \mathcal{K}} \quad \frac{\dagger \langle A_i \rangle \quad \left\langle \frac{[\langle A_{f(j)} \rangle_{j \in g^{-1}(k)}]}{C} \right\rangle_{k \in \mathcal{K}}}{C} \dagger E$$

REMARK 2.1. Let $\mathbf{0}$ be the empty list, and $\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}, \mathbf{5}$ be the lists $\langle 1 \rangle, \langle 1, 2 \rangle, \langle 1, 2, 3 \rangle, \langle 1, 2, 3, 4 \rangle, \langle 1, 2, 3, 4, 5 \rangle$ respectively. The connective $\bullet(A_1, A_2, A_3) = (A_1 \wedge A_2) \vee A_3$ is shorthand for $\dagger(I_3, g) \langle A_1, A_2, A_3 \rangle$, given by $\mathbf{3} \xleftarrow{I_3} \mathbf{3} \xrightarrow{g} \mathbf{2}$, where I_A indicates the identity function on the list $\mathcal{A}$ and $g : \{1, 2 \mapsto 1; 3 \mapsto 2\}$. Similarly, the usual connectives $\vee, \wedge, \top, \perp$ are obtained through the configurations $\mathbf{2} \xleftarrow{I_2} \mathbf{2} \xrightarrow{I_2} \mathbf{2}, \mathbf{2} \xleftarrow{I_2} \mathbf{2} \xrightarrow{1} \mathbf{1}, \mathbf{0} \xleftarrow{\emptyset} \mathbf{0} \xrightarrow{\emptyset} \mathbf{1}$ and $\mathbf{0} \xleftarrow{\emptyset} \mathbf{0} \xrightarrow{\emptyset} \mathbf{0}$, respectively (where $\emptyset$ indicates the empty function and 1 the constant function with value 1). We observe that some arguments of a connective may play a “dummy” role or may be used more than once, such as in the connective $\blacktriangle(A_1, A_2, A_3, A_4, A_5) = (A_2 \wedge A_3) \vee (A_4 \wedge A_5)$ given by $\mathbf{5} \xleftarrow{f} \mathbf{4} \xrightarrow{g} \mathbf{2}$, where $f : \{1 \mapsto 2; 2, 4 \mapsto 3; 3 \mapsto 4\}$ and $g : \{1, 2 \mapsto 1; 3, 4 \mapsto 2\}$. Observe that, informally speaking, for every connective $\dagger$ determined by the diagram $\mathcal{I} \xleftarrow{f} \mathcal{J} \xrightarrow{g} \mathcal{K}$, the elements of $\mathcal{J}$ code the positions of the arguments of $\dagger$ in the formula of the language of NI that define $\dagger$.

REMARK 2.2. Treating $\vee$ and $\perp$ as polynomial connectives (see Remark 2.1) yields their usual introduction and elimination rules. This is not the case for $\wedge$ (and $\top$). When treated as a polynomial connective, conjunction has a unique elimination rule $\wedge E_p$ instead of the usual $\wedge E_1$ and $\wedge E_2$ (see also Remark 2.4 below):

$$\frac{A \wedge B}{C} \quad \frac{[A][B]}{C} \wedge E_p \quad \frac{A \wedge B}{A} \wedge E_1 \quad \frac{A \wedge B}{B} \wedge E_2$$

REMARK 2.3. We will adopt the convention of using i for indices in $\mathcal{I}$, k for indices in $\mathcal{K}$ and j for indices in $g^{-1}(k)$. Hence, to enhance readability, we will often omit the indication of the index set, so that for instance $\langle A_i \rangle_{i \in \mathcal{I}}$ will be abbreviated as $\langle A_i \rangle$.

Given a set of propositional variables $\mathcal{V}$, indicated as $X, Y, Z, \dots$, the formulas of the language $\mathcal{L}^{2p}$ will be constructed using implication, the universal quantifier and the family of polynomial connectives $\dagger^{(f, g)}$. Besides $\mathcal{L}^{2p}$, we will mainly be concerned with two restrictions thereof, $\mathcal{L}^p$ and $\mathcal{L}^2$:

$$\mathcal{L}^{2p} ::= X \mid A \supset B \mid \dagger^{(f, g)} \langle A_i \rangle_{i \in \mathcal{I}} \mid \forall X. A$$

$$\begin{aligned}\mathcal{L}^p &::= X \mid A \supset B \mid \dagger^{(f,g)} \langle A_i \rangle_{i \in \mathcal{I}} \\ \mathcal{L}^2 &::= X \mid A \supset B \mid \forall X.A\end{aligned}$$

We moreover indicate with $\mathcal{L}^\vee$ (respectively $\mathcal{L}^\bullet$) the restriction of $\mathcal{L}^p$ in which $\vee$ (resp. $\bullet$) is the only connective besides $\supset$, and similarly for $\mathcal{L}^{2\vee}$ and $\mathcal{L}^{2\bullet}$.

The natural deduction system NI^{2p} over the language $\mathcal{L}^{2p}$ is obtained by adding to the rules of the standard system NI^2 for $\mathcal{L}^2$ (recalled in Definition 1.1 and Table 1 on pages 197–198 of [33]), all introduction and elimination rules for the connectives $\dagger^{(f,g)}$. The system NI^{2p} and its equational theory are described in detail in λ -notation in Appendices A and D respectively. Finally, by restricting NI^{2p} to the languages $\mathcal{L}^p, \mathcal{L}^\bullet, \dots$ we obtain the subsystems $\text{NI}^p, \text{NI}^\bullet, \dots$.

REMARK 2.4. Observe that the standard intuitionistic natural deduction system NI is *not* a fragment of NI^p since the elimination rule for conjunction in the latter system is in general form (see Remark 2.2). Although replacing $\wedge E_1$ and $\wedge E_2$ with $\wedge E_p$ does not alter derivability, the two forms of elimination rules behave differently with respect to identity of proofs. More on this below in Section 5.

The system NI_{at}^2 (as well as $\text{NI}_{at}^{2p}, \text{NI}_{at}^{2\bullet}, \text{NI}_{at}^{2\vee}$) is obtained by replacing the rule $\forall E$ with the following:

$$\frac{\forall X.A}{A[\![Y/X]\!]} \forall E_{at}$$

in which the witness must be an atomic formula.

We indicate derivability in $\text{NI}^p, \text{NI}^2, \dots$ with $\vdash_{\text{NI}^p}, \vdash_{\text{NI}^2}, \dots$

In the following we will often indicate an arbitrary formula $A \in \mathcal{L}^2$ as:

$$\forall \langle Y_1 \rangle (F_1 \supset \forall \langle Y_2 \rangle (F_2 \supset \dots \supset \forall \langle Y_n \rangle (F_n \supset \forall \langle Y_{n+1} \rangle (X)) \dots)) \quad (*)$$

where $\forall \langle Y_i \rangle$ (for $1 \leq i \leq n+1$) indicates a (possibly empty) list of consecutive quantifications. Moreover, we let $\text{at}(A) = X$ indicate the rightmost atom of A . Finally, we will often abbreviate $A_1 \supset (\dots \supset (A_{n-1} \supset A_n) \dots)$ with $A_1 \supset \dots \supset A_{n-1} \supset A_n$ (i.e. we assume right-associativity for implication).

2.2. Polynomial Formulas and the RP-Fragment of NI^2

The RP-translation of standard connectives $\wedge, \vee, \perp, \top$ is recalled in Table 1. Although the connective $\bullet$ can be translated by composing the translations of $\wedge$ and $\vee$, a natural and more economical way to encode $\bullet$ in NI^2 is given

Table 1. RP-Translation of Standard Connectives

$(A \wedge B)^* = \forall X.(A^* \supset B^* \supset X) \supset X$	(2.1)
$(A \vee B)^* = \forall X.(A^* \supset X) \supset (B^* \supset X) \supset X$	(2.2)
$\perp^* = \forall X.X$	(2.3)
$\top^* = \forall X.X \supset X$	(2.4)

by

$$\bullet(A, B, C)^* = \forall X.(A^* \supset (B^* \supset X)) \supset (C^* \supset X) \supset X$$

The universal formula above shares a common structure with those in Table 1: all such formulas are of the form $\forall X.A_1 \supset \cdots \supset A_n \supset X$, where each A_i has a unique occurrence of X in rightmost position. This suggests the following definition:

DEFINITION 2.5. A formula $A \in \mathcal{L}^2$ is *strongly positive in X* (short *sp- X*) when it is of the form $A_1 \supset \cdots \supset A_n \supset X$ for some $n \in \mathbb{N}$ and X does not occur in any of the A_i .

A formula $A \in \mathcal{L}^2$ is *polynomial in X* when it is of the form $A_1 \supset \cdots \supset A_n \supset X$ for some $n \in \mathbb{N}$ and all A_i are *sp- X* .

A formula $\forall X.A$ is called *universal polynomial* if A is polynomial in X .

REMARK 2.6. We introduce the following compact notation for NI²-formulas. Given a finite list $\mathcal{A} = \langle a_1, \dots, a_k \rangle$, an $\mathcal{A}$ -indexed list of formulas $A_{a_1}, \dots, A_{a_k}$ and a formula B , we let

$$\langle A_a \rangle_{a \in \mathcal{A}} \supset B$$

be shorthand for the formula $A_{a_1} \supset \cdots \supset A_{a_k} \supset B$ if $\mathcal{A}$ is non-empty, otherwise for B . When $\mathcal{A}$ is clear from the context the index $a \in \mathcal{A}$ will be omitted so we simply write $\langle A_a \rangle \supset B$.

A universal polynomial formula A can thus be written as $\forall X.\langle B_b \rangle_{b \in \mathcal{B}} \supset X$, where $\mathcal{B}$ is some list and the B_b s are *sp- X* . Moreover, any *sp- X* formula B can in turn be written as $\langle A_a \rangle_{a \in \mathcal{A}} \supset X$, for some list $\mathcal{A}$ such that each A_a has no occurrence of X . Hence, a universal polynomial formula can be written as $\forall X.\langle \langle A_a \rangle_{a \in \mathcal{A}_b} \supset X \rangle_{b \in \mathcal{B}} \supset X$ for some list $\mathcal{B}$ and family of lists $\mathcal{A}_b$ indexed by $\mathcal{B}$. Since a diagram $\mathcal{I} \xleftarrow{f} \mathcal{J} \xrightarrow{g} \mathcal{K}$ describes a family of lists indexed by the elements of a list (see footnote 5 above), it can be used to associate to each $\mathcal{I}$ -indexed family of formulas $\langle A_i \rangle_{i \in \mathcal{I}}$ the universal polynomial formula $\forall X.\langle \langle A_{f(j)}^* \rangle_{j \in g^{-1}(k)} \supset X \rangle_{k \in \mathcal{K}} \supset X$, which in turn we propose to take as the RP-translation of the propositional formula $\dagger^{f,g} \langle A_i \rangle_{i \in \mathcal{I}}$.

The Russell–Prawitz translation can thus be generalized to the whole of $\mathcal{L}^{2p}$ by translating the formulas of $\mathcal{L}^{2p}$ whose outermost connective is polynomial with a universal polynomial formula of $\mathcal{L}^2$:

DEFINITION 2.7. (*RP-Translation of Formulas*) We define a translation $*$ from formulas of $\mathcal{L}^{2p}$ to formulas of $\mathcal{L}^2$ as follows:

$$\begin{aligned} X^* &= X & (A \supset B)^* &= A^* \supset B^* & (\forall X.A)^* &= \forall X.A^* \\ \left(\dagger^{(f,g)} \langle A_i \rangle_{i \in \mathcal{I}} \right)^* &= \forall X. \langle \langle A_{f(j)}^* \rangle_{j \in g^{-1}(k)} \supset X \rangle_{k \in \mathcal{K}} \supset X & \text{(for } X \text{ not free in any } A_i) \end{aligned}$$

for any diagram $\mathcal{I} \xleftarrow{f} \mathcal{J} \xrightarrow{g} \mathcal{K}$

REMARK 2.8. For readability, we will abbreviate $(A \supset X) \supset ((B \supset X) \supset X)$ as $A \vee B$ (and thus $(A \vee B)^*$ as $\forall X.(A^* \vee B^*)$), and similarly $(A \supset (B \supset X)) \supset (C \supset X) \supset X$ as $\circ(A, B, C)$ (and thus $\bullet(A, B, C)^*$ as $\forall X.\circ(A^*, B^*, C^*)$).

The RP-translation scales well from formulas to derivations, yielding an embedding $*$: $\text{NI}^{2p} \mapsto \text{NI}^2$. The embedding for usual connectives is recalled in [33] Def. 2.2. Its extension to NI^p is defined in detail in Appendix C in λ -calculus notation.

REMARK 2.9. If $\mathcal{D}$ is an $\text{NI}^{2\bullet}$ -derivation of A from undischarged assumptions $A_1, \dots, A_n$, then $\mathcal{D}^*$ is an NI^2 -derivation of A^* from undischarged assumptions $A_1^*, \dots, A_n^*$. (For a proof for the whole NI^{2p} , see Appendix C).

If we restrict the language of NI^2 by requiring every universal formula to be universal polynomial, then we obtain a fragment of NI^2 that we will call the *Russell–Prawitz fragment*.

DEFINITION 2.10. (*Russell–Prawitz fragment*) We let $\mathcal{L}_{RP}^2$ be the subset of $\mathcal{L}^2$ in which all universal formulas are universal polynomial, and we let NI_{RP}^2 , called the *Russell–Prawitz fragment of NI^2* , be the fragment of NI^2 obtained by restricting to $\mathcal{L}_{RP}^2$ -formulas. The system NI_o^2 is the subsystem of NI_{RP}^2 in which universal formulas $\forall X.A$ are all of the form $\forall X.\circ(A_1, A_2, A_3)$.

REMARK 2.11. The RP-fragment of $\mathcal{L}^2$ can be equivalently defined inductively as follows:

$$\mathcal{L}_{RP}^2 ::= X \mid A \supset B \mid \forall X. \langle \langle A_{f(j)} \rangle_{j \in g^{-1}(k)} \supset X \rangle_{k \in \mathcal{K}} \supset X$$

where in the last clause we assume $\langle A_i \rangle$ to be an $\mathcal{I}$ -indexed family of $\mathcal{L}_{RP}^2$ -formulas with $X \notin FV(A_i)$ for all $i \in \mathcal{I}$ and $\mathcal{I} \xleftarrow{f} \mathcal{J} \xrightarrow{g} \mathcal{K}$ to be the diagram determining a polynomial connective. A quantified formula of $\mathcal{L}_{RP}^2$ may therefore have quantified formulas as proper subformulas, provided they are in turn universal polynomial formulas.

REMARK 2.12. It is clear that the restriction of the RP-translation to $\mathcal{L}^p$ is into $\mathcal{L}_{RP}^2$, and that if $\mathcal{D}$ is an $\text{NI}^\bullet$ -derivation then $\mathcal{D}^*$ is an NI_{RP}^2 -derivation.

REMARK 2.13. It is easy to check that the equational theory of NI_{RP}^2 is well-defined, since $\mathcal{L}_{RP}^2$ is closed under substitution.

REMARK 2.14. The system NI_{RP}^2 is the smallest fragment of NI^2 closed under $\supset$ and containing the RP-translation of all polynomial connectives.

Observe that not only the translation of a formula $\dagger^{f,g}\langle A_i \rangle$ is a universal polynomial formula, but for every universal polynomial formula A there is at least (but in general more than) one configuration $\mathcal{I} \xleftarrow{f} \mathcal{J} \xrightarrow{g} \mathcal{K}$ such that $A = (\dagger^{f,g}\langle A_i \rangle)^*$. To wit, let $A = \forall X. A_1 \supset \cdots \supset A_n \supset X$, where $A_i = A_{i1} \supset \cdots \supset A_{ik_i} \supset X$. It is enough to take $\mathcal{I} = \mathcal{J} = \langle \langle 1, 1 \rangle, \dots, \langle 1, k_1 \rangle, \dots, \langle n, 1 \rangle, \dots, \langle n, k_n \rangle \rangle$ (so that $A_{\langle i, j \rangle} = A_{ij}$), $\mathcal{K} = \langle 1, \dots, n \rangle$, f to be the identity function and $g = \pi_1^2 : \langle i, j \rangle \mapsto i$.

Given this it is easily seen by induction that for any $A \in \mathcal{L}_{RP}^2$ there is at least (but in general more than) one $B \in \mathcal{L}^p$ such that $B^* = A$. Note however that not every NI_{RP}^2 -derivation is the image of some NI^p -derivation under the RP-translation.

REMARK 2.15. In [33] we alluded to the fact that the translation of polynomial connectives can be described through a class of formulas called *nested sp-X*. There we used “sp-X” as short for “strictly positive” rather than “strongly positive”, where strictly positive formulas are a slight generalization of strongly positive formulas in which X may not occur at all, and nested sp-X formulas roughly stand to universal polynomial as strictly positive formulas stand to strongly positive formulas.

The class of polynomial formulas we consider here is thus a proper subset of the class of *nested sp-X* formulas and it is already sufficient for the goal of RP-translating polynomial connectives. However, most of the results we prove for NI_{RP}^2 can be extended to a similar (and slightly larger) system defined using nested sp-X (more generally these systems are fragments of a more general system $\Lambda 2^{\kappa \leq 2}$ studied in [20]).

3. The ε -Conversions

In this section we introduce some notational conventions and shortly recall some notions and results from the previous paper. These are based on the introduction of a notation for functors in natural deduction and of a class of conversions, called the ε -conversions, that express a naturality condition for natural deduction derivations.

3.1. Weak Expansion

DEFINITION 3.1. Given a formula $A \in \mathcal{L}^2$ (whose structure can be described as in $(*)$, see end of Section 2.1 above), the *weak expansion* of A is the derivation that one obtains by repeatedly applying η -expansions to the derivation consisting only of the assumption of A until the minimal formula of the main branch is atomic, i.e.:

$$\begin{array}{c}
 \frac{\frac{A}{F_1 \supset \forall \langle Y_2 \rangle (F_2 \supset \dots \supset \forall \langle Y_n \rangle (F_n \supset \forall \langle Y_{n+1} \rangle (X)) \dots)} \forall E \quad \frac{m_1}{F_1}}{\forall \langle Y_2 \rangle (F_2 \supset \dots \supset \forall \langle Y_n \rangle (F_n \supset \forall \langle Y_{n+1} \rangle (X)) \dots)} \supset E \\
 \vdots \\
 \frac{\frac{\forall \langle Y_n \rangle (F_n \supset \forall \langle Y_{n+1} \rangle (X))}{F_n \supset \forall \langle Y_{n+1} \rangle (X)} \forall E \quad \frac{m_n}{F_n}}{\frac{\forall \langle Y_{n+1} \rangle (X)}{X} \forall I} \supset E \\
 \frac{\frac{\frac{\frac{\frac{\frac{\forall \langle Y_{n+1} \rangle (X)}{F_n \supset \forall \langle Y_{n+1} \rangle (X)} \supset I (m_n)}{\forall \langle Y_n \rangle (F_n \supset \forall \langle Y_{n+1} \rangle (X))} \forall I}{\vdots}}{\frac{F_1 \supset \forall \langle Y_2 \rangle (F_2 \supset \dots \supset \forall \langle Y_n \rangle (F_n \supset \forall \langle Y_{n+1} \rangle (X)) \dots)} \supset I (m_1)}{\frac{A}{\vdots}} \forall I
 \end{array}$$

where $\frac{\forall \langle Y_i \rangle G}{G} \forall E$ and $\frac{G}{\forall \langle Y_i \rangle G} \forall I$ indicate (possibly empty) sequences of applications of $\forall E$ and $\forall I$.

We will indicate by $\text{El}_A^{\vec{m}}$ the “first half” of the derivation above, consisting of a chain of elimination rules depending on indices $\vec{m} = m_1, \dots, m_n$ for the undischarged assumptions $F_1^{m_1}, \dots, F_n^{m_n}$. Similarly, we will indicate by $\text{In}(\vec{m})$ the “second half” of the derivation above, consisting of a chain of introduction rules discharging the assumptions with indices $\vec{m}$ (a more rigorous description of this construction is provided in Appendix E through the notion of *expansion pair*).

REMARK 3.2. The weak expansion of A might differ from the expanded normal form (also known as η -long normal form) of the derivation consisting only of the assumption of A , since in the latter not only the minimal formula in the main branch, but the minimal formulas of all branches are atomic, see [23, §II.3.2.2.].

3.2. C -Expansion

DEFINITION 3.3. If $C = C_1 \supset \dots \supset C_n \supset X$ is $\text{sp-}X$ and $\mathcal{D}$ is a derivation of B from undischarged assumptions A, Δ (where A is freely chosen among the undischarged assumptions of $\mathcal{D}$), the C -expansion of $\mathcal{D}$ relative to X on

main assumption A , notation $C^X \begin{array}{c} A \\ \mathcal{D} \\ B \end{array}$ is the derivation of $C[B/X]$ from $C[A/X], \Delta$ defined by induction on n as follows:

- If $n = 0$, then $C = X$ and $C^X \begin{array}{c} A \\ \mathcal{D} \\ B \end{array}$ is just $\mathcal{D}$.
- If $n \geq 1$ then $C = C_1 \supset D$ where $X \notin FV(C_1)$ and D is $\text{sp-}X$. We define:

$$C^X \begin{array}{c} A \\ \mathcal{D} \\ B \end{array} = \frac{\frac{C_1 \supset D[A/X] \quad \tilde{C}_1^n}{\supset E} \quad \begin{array}{c} D^X \begin{array}{c} A \\ \mathcal{D} \\ B \end{array}}{\frac{\supset I (n)}{C_1 \supset D[B/X]}} \supset I (n)$$

REMARK 3.4. In ([33], sec. 3.2) we defined the notion of C -expansion for a broader class of formulas called $\text{pn-}X$ formulas. The restriction of the definition to the class of $\text{sp-}X$ formulas allows a straightforward reformulation of the definition using the notion of weak expansion that we give in Appendix E.

REMARK 3.5. In the functorial semantics of NI^2 , what we call the C -expansion of a derivation $\mathcal{D}$ is just the result of applying the functor interpreting C to the morphism interpreting $\mathcal{D}$. As any $\mathcal{L}^2$ -formula can be interpreted as a functor, the notion of C -expansion can be extended to any $C \in \mathcal{L}^2$.

REMARK 3.6. Whenever X is clear from the context, with the notation for substitution introduced in ([33], sec. 3.2), we write $C[B/X]$ and $C[A/X]$ as $C \begin{array}{c} B \\ A \end{array}$ and $C \begin{array}{c} A \\ A \end{array}$, and leaving the main assumption A implicit we indicate the C -expansion of $\mathcal{D}$ as $C \begin{array}{c} \mathcal{D} \end{array}$.

REMARK 3.7. In the present paper, we will only be concerned with the C -expansion of derivations of the form $\text{El}_A^{\vec{m}}$. Observe that, by the definition of C -expansion, the undischarged assumptions of $C \begin{array}{c} \text{El}_A^{\vec{m}} \end{array}$ besides $C \begin{array}{c} A \end{array}$ are

the same as the undischarged assumptions of $\text{El}_A^{\vec{m}}$ besides A , namely (see Definition 3.1 above) $F_1^{m_1}, \dots, F_n^{m_n}$.

3.3. The ε -Conversions

As recalled in the introduction, it is common to characterize identity of proofs using an equivalence induced by (the symmetric closure of) some reduction relation over derivations. The equivalence $\simeq_{\beta\eta}$ generated by the β - and η -conversions for NI^2 is standard (and recalled in Appendix D). Propositional connectives like $\vee$ require, in addition to β - and η -conversion rules, the so-called permutative conversions (which allow to permute an elimination rule upwards a $\vee$ -elimination rule), that we call here γ -conversions, generating an equivalence relation $\simeq_{\beta\eta\gamma}$. Analogous conversions can be defined for all polynomial connectives, see Appendix D.

REMARK 3.8. The equivalence relation $\simeq_{\beta\eta\gamma}$ can be further extended in two equivalent ways: either by replacing γ by a stronger permutation γ^+ which allows the permutation of an arbitrary derivation upwards across an application of $\dagger\text{E}$, or by replacing η by a stronger η^+ which expresses in categorical terms the universality of the connective (η^+ in fact subsumes both γ and γ^+), see Appendix D.

The equivalence-preservation problem of the RP-translation can be formulated as the failure of the implication below (see [33]):

$$\mathcal{D}_1 \simeq_{\beta\eta\gamma} \mathcal{D}_2 \quad \Rightarrow \quad \mathcal{D}_1^* \simeq_{\beta\eta} \mathcal{D}_2^*$$

In [33] (see Section 4.1) we showed that the implication above does hold when the equivalence we consider for NI^2 is the one induced by adding to β - and η -conversions a new class of conversions, called ε -conversions (for the case of $\circ$ these are shown in Table 2, for arbitrary universal polynomial formulas see Appendix F).

PROPOSITION 1. ([33]) *For all $\text{NI}^\vee$ -derivations $\mathcal{D}_1, \mathcal{D}_2$: if $\mathcal{D}_1 \rightsquigarrow_{\beta\eta\gamma} \mathcal{D}_2$, then $\mathcal{D}_1^* \simeq_{\beta\eta\varepsilon} \mathcal{D}_2^*$. More precisely:*

1. *if $\mathcal{D}_1 \rightsquigarrow_{\beta} \mathcal{D}_2$, then $\mathcal{D}_1^* \rightsquigarrow_{\beta} \mathcal{D}_2^*$;*
2. *if $\mathcal{D}_1 \rightsquigarrow_{\eta} \mathcal{D}_2$, then $\mathcal{D}_1^* \simeq_{\beta\eta\varepsilon} \mathcal{D}_2^*$;*
3. *if $\mathcal{D}_1 \rightsquigarrow_{\gamma} \mathcal{D}_2$, then $\mathcal{D}_1^* \rightsquigarrow_{\beta\varepsilon} \mathcal{D}_2^*$.*

Semantically, the ε -conversions express a naturality condition for NI^2 -derivations. Thus, the results of our previous paper were a syntactic reformulation of some well-known properties which hold in parametric models of NI^2 (see [1] and [14]).

Table 2. ε -conversion

$ \begin{array}{c} \mathcal{D}_1 \\ \forall X. \circ(F_1, F_2, F_3) \\ \circ(F_1, F_2, F_3) \ A \\ \hline (F_3 \supset X) \supset X \ A \end{array} $	$\supset E$	$ \begin{array}{c} F_1 \supset F_2 \supset X \ A \\ \hline F_3 \supset X \ A \end{array} $	$\supset E$	
$ \begin{array}{c} A \\ \mathcal{D}_2 \\ B \end{array} $				
$\rightsquigarrow_\varepsilon$				
$ \begin{array}{c} \mathcal{D}_1 \\ \forall X. \circ(F_1, F_2, F_3) \\ \circ(F_1, F_2, F_3) \ B \end{array} $	$\forall E$	$ \begin{array}{c} F_1 \supset F_2 \supset X \ A \\ \hline F_3 \supset X \ A \end{array} $	$\supset E$	
		$ \begin{array}{c} A \\ \mathcal{D}_2 \\ B \end{array} $		
		$ \begin{array}{c} (F_3 \supset X) \supset X \ B \\ \hline B \end{array} $		
			$\supset E$	

 (ε)

REMARK 3.9. The proof of Proposition 1 scales straightforwardly to the whole of NI^P (see Appendix G), and it actually holds if one replaces η (resp. γ) with the more general η^+ (resp. γ^+) (see Remark 3.8 above, Remark D.1 in Appendix D and Section 5.1 below). The class of ε -conversions needed to establish the analog of Proposition 1 for $\text{NI}^\bullet$ are depicted in Table 2.

4. Predicative Translations via Atomization

In this section we show that the RP-translation can be related to the predicative translations by embedding the NI_{RP}^2 fragment of NI^2 into the atomic fragment NI_{at}^2 .

4.1. The FF- and ESF-Translations

As we recalled in the introduction, Ferreira and Ferreira [7–9, 11], proposed an alternative translation of NI -derivations into NI^2 that we call the FF-translation. The FF-translation agrees with the RP-translation on how to translate formulas, but not on how to translate derivations. In particular, the FF-translation does not use the full power of the $\forall E$ rule of NI^2 , but rather it maps derivations in NI into derivations in NI_{at}^2 .

The FF-translation exploits the property of *instantiation overflow*:

DEFINITION 4.1. A formula $\forall X.A \in \mathcal{L}^2$ enjoys the instantiation overflow property iff for all formulas $B \in \mathcal{L}^2$

$$\forall X.A \vdash_{\text{NI}_{at}^2} A[B/X]$$

While Ferreira and Ferreira only address standard intuitionistic connectives, it is easily seen that the instantiation overflow property holds for all universal polynomial formulas $\forall X.A \in \mathcal{L}_{RP}^2$ (and actually for many more, see [19]). The results of Ferreira and Ferreira thus scale to the whole system NI^p . For simplicity we will focus here on the connective $\bullet$ and on the fragments $\text{NI}^\bullet$ and NI_o^2 of NI^p and NI_{RP}^2 respectively. However, all definitions and results here presented scale to all polynomial connectives, as shown in appendix.

By reformulating Ferreira and Ferreira's insight, we define an embedding of NI_o^2 into NI_{at}^2 that we call FF-atomization, and using it we define the FF-translation from $\text{NI}^\bullet$ into NI_{at}^2 as the composition of the RP-translation and FF-atomization.

DEFINITION 4.2. (*FF-atomization, FF-translation*) If $\mathcal{D}$ is an NI_o^2 -derivation, the *FF-atomization* of $\mathcal{D}$, which we indicate as $\mathcal{D}^\downarrow$, is the NI_{at}^2 -derivation defined by induction on $\mathcal{D}$ as follows. We only consider the case in which the last rule $\mathcal{D}$ is $\forall E$ with a non-atomic witness, since all other rules are translated in a trivial way. In this case observe that

$$\mathcal{D} = \frac{\mathcal{D}' \quad \forall X.o(A, B, C)}{o(A, B, C)[F/X]} \forall E$$

We define $\mathcal{D}^\downarrow$ by a sub-induction on F .

- If $F = F_1 \supset F_2$ then

$$\mathcal{D}^\downarrow = \frac{\left(\frac{\mathcal{D}' \quad \forall X.o(A, B, C)}{o(A, B, C)[F_2/X]} \forall E \right)^\downarrow \quad \frac{\frac{A \supset B \supset (F_1 \supset F_2) \quad \overset{\circ}{A}}{B \supset F_1 \supset F_2} \supset E \quad \frac{\overset{\circ'}{B}}{F_1 \supset F_2} \supset E \quad \frac{\overset{m}{F_1}}{C \supset (F_1 \supset F_2)} \supset E}{\frac{\frac{F_2}{B \supset F_2} \supset I (o') \quad \frac{C \supset (F_1 \supset F_2) \quad \overset{\circ}{C}}{F_1 \supset F_2} \supset E \quad \frac{\overset{m}{F_1}}{C \supset F_2} \supset E}{\frac{A \supset B \supset F_2}{(C \supset F_2) \supset F_2} \supset E} \supset E$$

$$\frac{\frac{F_2}{F_1 \supset F_2} \supset I (m) \quad \frac{(C \supset (F_1 \supset F_2)) \supset (F_1 \supset F_2)}{o(A, B, C)[F_1 \supset F_2/X]} \supset I (k_2) \quad \frac{}{o(A, B, C)[F_1 \supset F_2/X]} \supset I (k_1)}{\frac{}{o(A, B, C)[F_1 \supset F_2/X]} \supset I (k_1)} \supset E$$

- If $F = \forall Z.F'$, the clause is analogous to the previous one (for a fully detailed definition, see Definition H.1 in Appendix H)

If $\mathcal{D}$ is an $\text{NI}^\bullet$ -derivation, we call the NI_{at}^2 -derivation $\mathcal{D}^{*\downarrow}$ the FF-translation of $\mathcal{D}$.

REMARK 4.3. Ferreira and Ferreira present their result in a different way, by using the inductive clauses of Definition 4.2 to give a direct proof of instantiation overflow for the universal formulas of the form $(A \vee B)^*$, and

they refer to what we here called the FF-translation of $\mathcal{D}$ as to “the canonical translation of $\mathcal{D}$ in F_{at} provided by instantiation overflow” [11]. This difference in presentation will allow a more straightforward formulation of our results.

Whereas the FF- translation is defined by combining the RP-translation from $\text{NI}^\bullet$ into NI_{RP}^2 with the FF-atomization embedding NI_{RP}^2 into NI_{at}^2 , more recently Espírito Santo and Ferreira [5] introduced an alternative translation by “directly” defining an embedding $\sharp$ from NI into NI_{at}^2 . We will refer to this translation as the *ESF-translation*. Using the notation we introduced, the definition of $\sharp$ can be adapted to $\text{NI}^\bullet$ (the definition actually scales to the whole of NI^p , see Def. H.5 in Appendix H), and the crucial case of the definition, that of an $\text{NI}^\bullet$ -derivation ending with an application of $\bullet\text{E}$ runs as follows (assuming $\text{at}(C^*) = Z$):

$$\left(\frac{\begin{array}{ccc} & [A_1][A_2] & [A_3] \\ \mathcal{D}_1 & \mathcal{D}_2 & \mathcal{D}_3 \\ \bullet(A_1, A_2, A_3) & C & C \end{array}}{C} \bullet\text{E} \right)^\sharp = \frac{\frac{\mathcal{D}_1^\sharp}{\forall X. \circ(A_1, A_2, A_3)} \quad \frac{\frac{\frac{\mathcal{D}_2^\sharp}{C^*} \quad \text{El}_{C^*}^{\vec{m}}}{Z} \supset \text{I} (n_2)}{\frac{\forall X. \circ(A_1, A_2, A_3)[Z/X] \quad \forall \text{E}_{at}}{A_1^* \supset A_2^* \supset Z^*} \supset \text{I} (n_1)} \quad \frac{\frac{\mathcal{D}_3^\sharp}{[A_3^*]} \quad C^*}{\text{El}_{C^*}^{\vec{m}}} \supset \text{I} (m)}{\frac{\frac{\frac{Z^*}{\text{In}_{C^*}(\vec{m})}}{(A_3^* \supset X) \supset X \quad Z} \quad \frac{Z^*}{A_3^* \supset Z^*} \supset \text{I} (m)}}{C^*}}$$

These two translation not only allow one to translate $\text{NI}^\bullet$ into a very weak predicative fragment of NI^2 , but they also preserve slightly *more* equations than the RP-translation. In particular, the equivalences induced by γ - and η -conversion are preserved without appealing to additional conversions for NI^2 beside β - and η -conversions. In fact, both translations map not only β -, but also η -reduction steps in NI onto chains of $\beta\eta$ -reduction steps in NI^2 and γ -reduction steps in NI onto chains of $\beta\eta$ -reduction and expansion steps in NI^2 :

PROPOSITION 2. *Let $\natural \in \{*\downarrow, \sharp\}$. For all $\text{NI}^\bullet$ -derivations $\mathcal{D}_1, \mathcal{D}_2$: if $\mathcal{D}_1 \rightsquigarrow_{\beta\eta\gamma} \mathcal{D}_2$, then $\mathcal{D}_1^\natural \simeq_{\beta\eta} \mathcal{D}_2^\natural$.*

More precisely:

1. if $\mathcal{D}_1 \rightsquigarrow_\beta \mathcal{D}_2$, then $\mathcal{D}_1^\natural \rightsquigarrow_{\beta\eta} \mathcal{D}_2^\natural$;
2. if $\mathcal{D}_1 \rightsquigarrow_\eta \mathcal{D}_2$, then $\mathcal{D}_1^\natural \rightsquigarrow_{\beta\eta} \mathcal{D}_2^\natural$;
3. if $\mathcal{D}_1 \rightsquigarrow_\gamma \mathcal{D}_2$, then $\mathcal{D}_1^\natural \simeq_\beta \mathcal{D}_2^\natural$.

PROOF. For the FF-translation $*\downarrow$, point 1 and 3 of the proposition were proven for the whole of NI by Ferreira and Ferreira [9] and [8] respectively,

and point 2 by Ferreira [11] (although in this case by adding a primitive conjunction to NI^2 , see below Section 5.1), and we claim that, at the cost of tedious computations, those proofs can be scaled to the whole of NI^P . In Appendix I we give a proof of the proposition for the ESF-translation $\sharp$ scaled to the whole of NI^P , that generalizes the analogous result for NI of Espírito Santo and Ferreira in [5]. ■

4.2. The ε -Translation

We now show that the ε -rule can be used to clarify the relationship between the RP-translation and the FF- and ESF-translations. To see how, we define an alternative atomization procedure yielding yet another translation of $\text{NI}^\bullet$ into NI_{at}^2 (also this atomization scales to the whole of NI^P , see Definition H.2 in Appendix H):

DEFINITION 4.4. (ε -atomization, ε -translation) The definition differs from Definition 4.2 in the following respect: In case

$$\mathcal{D}' = \frac{\forall X. \circ(A, B, C)}{\circ(A, B, C)[F/X]} \forall E$$

where F is not atomic and Z is the variable in rightmost position in F , using the notation introduced in Sections 3.1–3.2 (see in particular Remark 3.7), we define :

$$\mathcal{D}^{\downarrow \varepsilon} = \frac{\frac{\mathcal{D}'^{\downarrow \varepsilon}}{\frac{\forall X. \circ(A, B, C)}{\circ(A, B, C)[Z/X]} \forall E_{at}} \quad \frac{\begin{array}{c} k_1 \\ \boxed{A \supset B \supset X \quad \begin{array}{c} F \\ \text{El}_{\vec{m}} \\ Z \end{array}} \end{array}}{(C \supset Z) \supset Z} \supset E \quad \frac{\begin{array}{c} k_2 \\ \boxed{C \supset X \quad \begin{array}{c} F \\ \text{El}_{\vec{m}} \\ Z \end{array}} \end{array}}{\supset E} \quad \frac{\frac{\text{In}_F(\vec{m})}{F} \supset I(k_2)}{\frac{(C \supset F) \supset F}{\circ(A, B, C)[F/X]} \supset I(k_1)} \supset E$$

If $\mathcal{D}$ is an $\text{NI}^\bullet$ -derivation, we call the NI_{at}^2 -derivation $\mathcal{D}^{*\downarrow \varepsilon}$ the ε -translation of $\mathcal{D}$.

The name “ ε -atomization” is justified by the fact that, as can be easily verified in the case of $\circ$, the ε -atomization $\mathcal{D}^{\downarrow \varepsilon}$ of an NI_o^2 -derivation $\mathcal{D}$ is the result of applying η -expansions and ε -conversions to $\mathcal{D}$:

PROPOSITION 3. If $\mathcal{D}$ is an NI_o^2 -derivation, then $\mathcal{D} \simeq_{\eta\varepsilon} \mathcal{D}^{\downarrow \varepsilon}$.

PROOF. See Appendix J. ■

The relationship between the three predicative translations is very close. In fact, they yield β -equivalent derivations, with the ε -translation lying between the FF-translation and the ESF-translation, in the sense below:

PROPOSITION 4. *For all $\text{NI}^\bullet$ -derivations $\mathcal{D}$, $\mathcal{D}^{*\downarrow} \rightsquigarrow_\beta \mathcal{D}^{*\downarrow^\varepsilon} \rightsquigarrow_\beta \mathcal{D}^\sharp$.*

PROOF. See Appendix K. ■

REMARK 4.5. That $\mathcal{D}^{*\downarrow} \rightsquigarrow_\beta \mathcal{D}^\sharp$ was already known from [5].

By putting together Proposition 2 and 4 we can deduce that also the ε -translation preserves the equivalence induced by permutative conversions and η -conversions:

COROLLARY 4.6. *For all $\text{NI}^\bullet$ -derivations $\mathcal{D}_1, \mathcal{D}_2$, if $\mathcal{D}_1 \simeq_{\beta\eta\gamma} \mathcal{D}_2$, then $\mathcal{D}_1^{*\downarrow^\varepsilon} \simeq_{\beta\eta} \mathcal{D}_2^{*\downarrow^\varepsilon}$.*

REMARK 4.7. The statement of Corollary 4.6 cannot be expressed in terms of reduction, but only in terms of equivalence, due to the fact that neither $\mathcal{D}_1 \rightsquigarrow_\eta \mathcal{D}_2$ nor $\mathcal{D}_1 \rightsquigarrow_\gamma \mathcal{D}_2$ imply $\mathcal{D}_1^{*\downarrow^\varepsilon} \rightsquigarrow_{\beta\eta} \mathcal{D}_2^{*\downarrow^\varepsilon}$, but only $\mathcal{D}_1^{*\downarrow^\varepsilon} \simeq_{\beta\eta} \mathcal{D}_2^{*\downarrow^\varepsilon}$.

By combining Proposition 3 with Proposition 4 we deduce that the FF- and ESF-translations are $\beta\eta\varepsilon$ -equivalent to the RP-translation:

COROLLARY 4.8. *For all $\text{NI}^\bullet$ -derivations $\mathcal{D}$, $\mathcal{D}^* \simeq_{\beta\eta\varepsilon} \mathcal{D}^{*\downarrow}$ and $\mathcal{D}^* \simeq_{\beta\eta\varepsilon} \mathcal{D}^\sharp$.*

The relationship between the four different translations is illustrated in Table 3.

Summing up, all three predicative translations are equivalent, modulo $\beta\eta\varepsilon$ -equivalence to the RP-translation and thus, semantically, they all interpret an $\text{NI}^\bullet$ -derivation using the same second-order proof (provided one understands second-order proofs as satisfying the naturality conditions expressed by ε -conversions).

Moreover, our results offer a new insight for the fact that the predicative translations do preserve the η - and γ -equivalences of $\text{NI}^\bullet$ resorting only to the $\beta\eta$ -equivalences of NI^2 : the predicative translations encode those bits of ε -equivalences (or, semantically speaking, of naturality) needed to translate the η - and γ -equivalences of $\text{NI}^\bullet$ into NI^2 .

5. Limitations of the Predicative Translations

By comparing Propositions 1 and 2 of the previous sections, it might be tempting to say that the approach based on atomic polymorphism provides

Table 3. Relationship among the RP-, FF-, ESF- and ε -translations

For all NI [•] derivations $\mathcal{D}_1$ and $\mathcal{D}_2$ such that $\mathcal{D}_1 \rightsquigarrow_{\beta\eta\gamma} \mathcal{D}_2$:				
		$\mathcal{D}_1^*$		
		$\simeq_{\eta\varepsilon}$		
$\mathcal{D}_1^{*\downarrow}$	$\rightsquigarrow_\beta$	$\mathcal{D}_1^{*\downarrow^\varepsilon}$	$\rightsquigarrow_\beta$	$\mathcal{D}_1^\sharp$
$\simeq_{\beta\eta}$		$\simeq_{\beta\eta}$		$\simeq_{\beta\eta}$
$\mathcal{D}_2^{*\downarrow}$	$\rightsquigarrow_\beta$	$\mathcal{D}_2^{*\downarrow^\varepsilon}$	$\rightsquigarrow_\beta$	$\mathcal{D}_2^\sharp$
		$\simeq_{\eta\varepsilon}$		
		$\mathcal{D}_2^*$		

a fully syntactic alternative to categorical semantics and related techniques for the study of identity of proofs for propositional connectives.

In this section we argue that this is not entirely the case, by stressing two important limitations of the predicative translations. First, they do not preserve the whole equational theory of propositional connectives, and in particular the predicative translations of $\wedge, \vee$ do not yield products and co-products in NI_{at}^2 . Hence, *strictu sensu* the predicative translations do not fully solve the equivalence-preservation problem. Second, the $\forall E$ rule restricted to atomic witnesses is too weak to prove the logical equivalence between a connective and its second-order translation. In Prawitz' terminology ([22]), we show that the connective $\vee$ is not *strongly* definable in NI_{at}^2 but only *weakly* definable.

5.1. Predicative Translations and Generalized Permutations

As is well known, in order to obtain a perfect match between the syntax of NI and the free bi-cartesian closed category, it is necessary to consider the following stronger permutative conversions in which every chunk of derivation (and not just those consisting of applications of elimination rules) can be permuted-up across an application of disjunction elimination:^{6,7}

⁶These stronger permutations were first formulated in natural deduction format by [29] and have been recently discussed in the context of proof-theoretic semantics by [31, 32].

⁷The general form of the γ^+ -conversions for an arbitrary polynomial connective (of which the one given here is a simple instance) is given in Appendix D.

$$\begin{array}{c}
\begin{array}{ccc}
& \overset{n}{[A]} & \overset{m}{[B]} \\
\mathscr{D} & \mathscr{D}_1 & \mathscr{D}_2 \\
\hline
A \vee B & C & C \\
\hline
& [C] & \\
& \mathscr{D}_3 & \\
& D &
\end{array}
\vee E(n, m) \rightsquigarrow_{\gamma^+}
\begin{array}{ccc}
& \overset{n}{[A]} & \overset{m}{[B]} \\
\mathscr{D} & \mathscr{D}_1 & \mathscr{D}_2 \\
& [C] & [C] \\
& \mathscr{D}_3 & \mathscr{D}_3 \\
\hline
A \vee B & D & D \\
\hline
& D &
\end{array}
\vee E(n, m)
\end{array}$$

Point 3 of Proposition 1 can be strengthened by replacing γ with γ^+ , so that for any instance of γ^+ , the RP-translation of its left-hand side $\beta\varepsilon$ -reduces to the RP-translation of the right-hand side (see Appendix G).

However, Point 3 of Proposition 2 ceases to hold as soon as one replaces γ with γ^+ . That is, although the predicative translations preserve the equivalences induced by η - and γ -conversions, they do not preserve all equivalences needed to interpret NI as a bi-cartesian closed category.

To see this, it is enough to consider the instance of γ^+ in which A, B, C and D are atoms, $\mathscr{D}$ (respectively $\mathscr{D}_1, \mathscr{D}_2$) consists only of the assumptions of $A \vee B$ (resp. C), and $\mathscr{D}_3$ is the derivation $\frac{C \supset D}{D} C \supset E$. As the reader can easily check, the predicative translations of the left and right-hand side of this instance of $\rightsquigarrow_{\gamma^+}$ are not $\beta\eta$ -equivalent, just like their RP-translations.

That the predicative translations fail to preserve the equivalence induced by the stronger permutations provides an explanation for another puzzling aspect of the approach of Ferreira and co-authors. As observed in [11], Point 2 of Proposition 2 fails if $\text{NI}^\vee$ is extended to include the standard rules of conjunction. In particular, the predicative translations fail to preserve all instances of the η -equation for the standard conjunction rules. Nonetheless, when the two standard elimination rules $\wedge E_1$ and $\wedge E_2$ are replaced by the general elimination rule $\wedge E_p$ as in NI^p (see above Remarks 2.2 and 2.4), Proposition 2 applies to the equivalence relation induced by the β -, η - and γ -conversions for “generalized” conjunction (see Appendix D for a formulation of these conversions in λ -calculus notation and Appendix I for a proof of Proposition 2 scaled to NI^p).

The standard elimination rules can be defined using the general elimination rule as follows:

$$\frac{A \wedge B \quad [A]}{A} \wedge E_p \qquad \frac{A \wedge B \quad [B]}{B} \wedge E_p$$

and given these definitions, the η -equations for the standard conjunction rules become the following:

$$\frac{\mathcal{D}}{A \wedge B} \simeq \frac{\frac{\mathcal{D}}{A \wedge B} \quad \frac{m}{A} \wedge E_p(m)}{A \wedge B} \quad \frac{\frac{\mathcal{D}}{A \wedge B} \quad \frac{n}{B} \wedge E_p(n)}{A \wedge B} \wedge I$$

The different behavior of the standard and the general elimination rules is explained by the fact that, as in the case of disjunction, the predicative translations of two γ^+ -equivalent NI^P -derivations need not be $\beta\eta$ -equivalent NI_{at}^2 -derivations, and that the γ^+ -conversion for conjunction:⁸

$$\frac{\frac{\mathcal{D}}{A \wedge B} \quad \frac{\mathcal{D}_1}{C} \wedge E(n, m)}{[C]} \rightsquigarrow_{\gamma^+} \frac{\frac{\mathcal{D}}{A \wedge B} \quad \frac{\mathcal{D}_2}{D} \wedge E(n, m)}{D}$$

is essential to define the η -conversion for the standard elimination rules using the η -conversion for the general elimination rule (see Remark D.3 in Appendix D for a proof).

5.2. Definability of Disjunction in NI_{at}^2

A second disadvantage of the predicative translations arises when one considers the extension $\text{NI}^{2\vee}$ of NI^2 with a primitive disjunction (for an early semantic investigation of this system see [12,30]), in which the following holds:

PROPOSITION 5. For all $A \in \mathcal{L}^{2\vee}$:

$$A \dashv\vdash_{\text{NI}^{2\vee}} A^*$$

PROOF. See Appendix L. ■

That is, as soon as one extends NI^2 with a primitive disjunction, one can show in the extended system $\text{NI}^{2\vee}$ that every formula is interderivable with its RP-translation.⁹

In contrast to what happens in $\text{NI}^{2\vee}$, a propositional formula A and its RP-translation A^* may fail to be interderivable in the extension of NI_{at}^2

⁸The same remarks of footnote 7 apply.

⁹Actually, A and its RP-translation are not just interderivable, but can be shown to be isomorphic modulo $\simeq_{\beta\eta\gamma^+\varepsilon}$ (see for details [20]).

with a primitive disjunction. By inspecting one direction of the proof of Proposition 5 it is clear that $A \vdash_{\text{NI}_{at}^{2\vee}} A^*$. However, the inspection of the other direction clearly suggests that, at least in some cases, in order to establish that $A^* \vdash_{\text{NI}_{at}^{2\vee}} A$ it is essential to apply $\forall E$ with a non-atomic witness. In particular,

PROPOSITION 6. *For any two distinct propositional variables Y, Z , $(Y \vee Z)^* \not\vdash_{\text{NI}_{at}^{2\vee}} Y \vee Z$*

PROOF. See Appendix L. ■

In $\text{NI}_{at}^{2\vee}$ the RP-translation of $Y \vee Z$ is thus strictly weaker than $Y \vee Z$.

REMARK 5.1. An immediate consequence of Proposition 6 concerns the faithfulness of the RP-translation. We recall that a translation $(\cdot)^\natural$ from $\mathcal{N}$ to $\mathcal{N}'$ is faithful iff for all Γ, A , if $\Gamma^\natural \vdash_{\mathcal{N}'} A^\natural$ then $\Gamma \vdash_{\mathcal{N}} A$. Although the FF-translation is a faithful translation of $\text{NI}^\vee$ into NI_{at}^2 ([10]), it is *not* a faithful translation of $\text{NI}_{at}^{2\vee}$ into NI_{at}^2 . In fact, keeping in mind that the FF- and the RP-translations of formulas coincide, if we let $\Gamma = (Y \vee Z)^*$ and $A = Y \vee Z$ we have that $((Y \vee Z)^*)^* = (Y \vee Z)^*$ and hence obviously $\Gamma^* \vdash_{\text{NI}_{at}^2} A^*$, but $\Gamma \vdash_{\text{NI}_{at}^{2\vee}} A$ does not hold by the above proposition.

REMARK 5.2. In addition to the failure of faithfulness, also the disjunction property fails for the translation of disjunction in NI_{at}^2 in the following sense: there exist formulas A, B of $\text{NI}_{at}^{2\vee}$ such that $\vdash_{\text{NI}_{at}^2} (A \vee B)^*$ holds but neither of $\vdash_{\text{NI}_{at}^2} A^*$ and $\vdash_{\text{NI}_{at}^2} B^*$ holds. The following example establishes this fact.

For readability, we write $\exists Z.C$ as shorthand for $\forall W.(\forall Z.C \supset W) \supset W$ (with W a fresh propositional variable) and $\frac{C[Y/Z]}{\exists Z.C} \exists I$ as shorthand for

$$\frac{\frac{\frac{\forall Z.C \supset W}{C[Y/Z] \supset W} \forall E}{W} \supset I (p)}{\exists Z.C} \forall I$$

Let $A = \forall X \forall Y. X \supset Y$; neither A nor $\exists X.A \supset X$ are provable in NI_{at}^2 (see Proposition 22 in Appendix L). Notice that $A^* = A$ and $(\exists X.A \supset X)^* = \exists X.A \supset X$. On the other hand, one can show that $\vdash_{\text{NI}_{at}^2} (A \vee (\exists X.A \supset X))^*$ as follows:

$$\frac{\frac{\frac{(\exists X.A \supset X) \supset X}{X} \supset I (m)}{(A \supset X) \supset ((\exists X.A \supset X) \supset X) \supset X} \supset I (m, n)}{(A \vee (\exists X.A \supset X))^*} \forall I$$

Another way of highlighting the difference between $(A \vee B)^*$ and $(A \vee B)$ in $\text{NI}_{at}^{2\vee}$ is by observing that, whereas the disjunction elimination rule warrants that $A \vee B \vdash_{\text{NI}_{at}^{2\vee}} (A \supset C) \supset (B \supset C) \supset C$ for all A, B, C in $\mathcal{L}^{2\vee}$, the same does not hold if one replaces $A \vee B$ with $\forall X.(A \vee B)$:

PROPOSITION 7. *There exist $\mathcal{L}^{2\vee}$ -formulas A, B, C such that*

$$\forall X.(A \vee B) \not\vdash_{\text{NI}_{at}^{2\vee}} (\forall X.(A \vee B)) \llbracket C/X \rrbracket$$

PROOF. See Appendix L ■

As a consequence, in contrast to what happens in NI_{at}^2 , the instantiation overflow property fails for $\forall X.(A \vee B)$ in $\text{NI}_{at}^{2\vee}$.

Given Proposition 6, the analog of Proposition 5 cannot hold for $\text{NI}_{at}^{2\vee}$. One may therefore wonder in which sense, if at all, can one say that disjunction is definable in NI_{at}^2 . Following Prawitz ([22], p. 58), we can distinguish between *strong and weak definability* of a connective. In Prawitz' terminology, disjunction is weakly definable in NI_{at}^2 if and only if there is a faithful translation of $\text{NI}^\vee$ into NI_{at}^2 , which is the case, since the FF-translation is faithful (see Remark 5.1 above); on the other hand, disjunction is strongly definable in NI_{at}^2 iff for all $A, B \in \mathcal{L}^\vee$ there is an $\mathcal{L}^2$ -formula C , such that $C \dashv\vdash_{\text{NI}_{at}^{2\vee}} A \vee B$.¹⁰ Then, not only Proposition 6 shows that disjunction is not strongly defined by its RP-translation, but there is in fact no other formula strongly defining it in NI_{at}^2 :

PROPOSITION 8. *$\vee$ is not strongly definable in NI_{at}^2 .*

PROOF. See Appendix L ■

6. Concluding Remarks

6.1. Summary of the Results

In this paper we have shown how the category-theory-inspired ε -conversions introduced in our previous paper can be used to clarify the relationship between the alternative translations proposed by Ferreira, Ferreira and Espírito Santo and the original Russell-Prawitz translation, and to provide semantic insights on the proof-theoretic properties of the former translations.

Our approach consisted in focusing on an atomizing translation from a suitable fragment of NI^2 (the fragment that we called NI_{RP}^2 , in which

¹⁰Note that strong definability implies the existence of a faithful translation from $\text{NI}_{at}^{2\vee}$ to NI_{at}^2 .

universal formulas correspond to the translation of polynomial connectives) into NI_{at}^2 , that we defined using the ε -conversions.

This made it possible to show that the predicative translations produce derivations that are equivalent to the RP-translation modulo the ε -conversions (hence, in semantic terms, modulo parametricity), and that their proof-theoretic properties (the preservation of the equivalence induced by η - and permutative conversions) result from the fact that ε -conversions express a naturality condition for the proofs denoted by NI^2 -derivations.

We thus used the connections of ε -conversions with category theory to provide an explanation of how and why the predicative translations work, by embedding the syntactic results about them into a wider picture.

Moreover, we showed that the predicative translations do not offer a full solution to the equivalence preservation problem, due to the fact that the predicative translations of the right-hand and left-hand side of generalized permutations in NI are NI_{at}^2 -derivations which in general may fail to be $\beta\eta$ -equivalent. This suggests that the system NI_{at}^2 coupled with $\beta\eta$ -equivalence is not the best setting to study the notion of identity of proof resulting from the full equational theory on NI -derivations. We provided further hints for this conclusion by showing that disjunction is only weakly and not strongly definable in NI_{at}^2 .

6.2. Related and Further Work

The $\beta\eta\gamma^+$ -equational theory has been recently shown to be the maximum consistent equational theory in NI [28]. As shown by the first and second author in [20], the $\beta\eta\varepsilon$ -equational theory in the fragment NI_{RP}^2 is equivalent (in the sense of category theory) to the $\beta\eta\gamma^+$ -equational theory of NI . From this they inferred both the maximality of the $\beta\eta\varepsilon$ -equational theory in NI_{RP}^2 and some decidability results for the fragment NI_{RP}^2 (decidability of type inhabitation, i.e. of provability, and of contextual equivalence). In contrast to what happens in NI_{RP}^2 , provability and contextual equivalence are both undecidable in NI_{at}^2 (although type checking and typability are decidable in NI_{at}^2 , see [21] for an overview of these results).

Although a generalization of our ε -conversions (expressing a dinaturality condition, rather than just naturality) can be formulated for any formula $\forall X.A$ of $\mathcal{L}^2$, the tight connection between the ε -conversions and the phenomenon of instantiation overflow underlying the present paper seems to be limited to the class of universal polynomial formulas. In fact, in [33] the ε -conversions were defined for a broader class of formulas, that of nested p - X formulas, all enjoying the instantiation overflow property. However, it

does not seem possible to define a uniform atomization procedure in the style of our ε -atomization for the NI^2 -derivations in which the premises of all applications of $\forall\text{E}$ are nested $p\text{-}X$ (rather than universal polynomial) formulas. For a characterization of the class of $\mathcal{L}^2$ -formulas enjoying instantiation overflow, see [3] and [19].

As to further directions of investigations, we observe that, in the extension of System F with primitive propositional connectives, not only every proposition and its Russell-Prawitz translation are interderivable, but they are actually *isomorphic* (modulo the equivalence relation induced by β -, η -, γ^+ -, and ε -equations). From a categorical perspective, such isomorphisms belong to a more general class of isomorphisms induced by a proof-theoretic formulation of the *Yoneda lemma*. This and further categorical aspects underlying the present work have been investigated by the authors in [20].

As observed, the ε -conversions allow one to translate not only the γ - but also γ^+ -permutations. As a consequence, the study of the reduction (as opposed to equivalence) induced by ε -conversions can be expected to be rather involved. In particular, the disadvantages of the fully extensional reduction theory for disjunction, such as non-confluence and non-termination (see e.g. [17] for a discussion) would be imported inside NI^2 .

As in the case of disjunction, these problems do not exclude the possibility of considering restricted forms of ε -conversion (essentially, those in which only elimination rules are permuted across an application of $\forall\text{E}$) corresponding to the restricted permutations of NI , and it is not implausible to thereby obtain a well-behaving reduction theory.

In a recent paper [6], Espírito Santo and Ferreira developed an analogous approach by devising further conversions in NI^2 corresponding to those steps of ε -conversions and of η -expansions needed to atomize applications of $\forall\text{E}$ in the RP -translation of NI -derivations. Using the reduction relation induced by the β -, η - and the newly introduced conversions in place of ε -conversions they could obtain a result analogous to our Proposition 1 but establishing the preservation of reduction rather than just equivalence also in the case of η -conversions.

Acknowledgements. Many thanks to the referees for the many suggestions that substantially contributed to improve the paper. Many thanks also to Gilda Ferreira, Fernando Ferreira and José Espírito Santo for several discussions on atomic polymorphism.

Funding Open Access funding enabled and organized by Projekt DEAL. Mattia Petrolo gratefully acknowledges the support of the São Paulo

Research Foundation (FAPESP) through the Project Auxlio Pesquisa Jovem Pesquisador no.2016/25891-3 and the support of the French National Research Agency (ANR) through the Project ANR-20-CE27-0004. Luca Tranchini gratefully acknowledges the support of the Deutsche Forschungsgemeinschaft (DFG) through the project “Falsity and Refutation. On the negative side of logic” (DFG TR1112/4-1).

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A. The System NI^{2p}

We indicate by $\mathcal{TV}$ a countably infinite set of *term variables* $x, y, z, \dots$. The terms of the system NI^{2p} (we adopt a domain-free formulation, see [2]) are given by the following grammar (where, given a list $\mathcal{A} = \langle a_1, \dots, a_n \rangle$, we indicate an $\mathcal{A}$ -indexed list of terms $t_{a_1} \dots t_{a_n}$ as $\langle t_a \rangle_{a \in \mathcal{A}}$, or simply with $\langle t_a \rangle$ whenever $\mathcal{A}$ is clear from the context, see Remark 2.3):

$$t, u := x \in \mathcal{TV} \mid \lambda x.t \mid tu \mid \Lambda X.t \mid tB \mid \iota_{\dagger}^k \langle t_j \rangle_{j \in g^{-1}(k)} \\ \mid \delta_{\dagger} \left(t, \left\langle \langle x_j \rangle_{j \in g^{-1}(k)} \cdot s_k \right\rangle_{k \in \mathcal{K}} \right)$$

Observe that, for every connective $\dagger$ generated by $\mathcal{I} \xleftarrow{f} \mathcal{J} \xrightarrow{g} \mathcal{K}$, we have a family of generalized injections $\iota_{\dagger}^k$ indexed by elements of $\mathcal{K}$, and a generalized “case” constructor $\delta_{\dagger}$ with $|\mathcal{K}| + 1$ arguments.

Let $\mathcal{A} = \langle a_1, \dots, a_k \rangle$ be a finite list. We will use the following abbreviated notation to indicate $\mathcal{A}$ -indexed sequences of λ -abstractions and applications: if $\langle x_a \rangle_{a \in \mathcal{A}}$ is an $\mathcal{A}$ -indexed list of variables, for any term t , we let $\lambda \langle x_a \rangle_{a \in \mathcal{A}}.t$ (often abbreviated as $\lambda \langle x_a \rangle.t$) indicate the term $\lambda x_{a_1} \dots \lambda x_{a_k}.t$. Similarly,

if $\langle t_a \rangle$ is an $\mathcal{A}$ -indexed list of terms, for any term u , we let $u\langle t_a \rangle_{a \in \mathcal{A}}$ (often abbreviated as $u\langle t_a \rangle$) indicate the term $ut_{a_1} \dots t_{a_k}$.

As usual, by a *typing context* we indicate a finite set of type declarations $x : A$ where all declared variables are distinct. We indicate typing contexts with Γ, Δ, Σ .

The typing rules for NI^{2p} are the following:

$$\begin{array}{c}
\frac{}{\Gamma, x : A \vdash x : A} \text{Ax} \quad \frac{\Gamma, x : A \vdash t : B}{\Gamma \vdash \lambda x. t : A \supset B} \supset \text{I} \quad \frac{\Gamma \vdash t : A \supset B \quad \Gamma \vdash u : A}{\Gamma \vdash tu : B} \supset \text{E} \\
\frac{\Gamma \vdash t : A}{\Gamma \vdash \Lambda X. t : \forall X. A} \forall \text{I} \text{ (Proviso: } X \notin \text{FV}(\Gamma)) \quad \frac{\Gamma \vdash t : \forall X. A}{\Gamma \vdash tB : A[B/X]} \forall \text{E} \\
\left\langle \frac{\Gamma \vdash t_j : A_{f(j)} \rangle_{j \in g^{-1}(k)}}{\Gamma \vdash \iota_{\dagger}^k \langle t_j \rangle : \dagger \langle A_i \rangle} \dagger \text{I}_k \right\rangle_{k \in \mathcal{K}} \quad \frac{\Gamma \vdash t : \dagger \langle A_i \rangle \quad \left\langle \Gamma, \langle y_j : A_{f(j)} \rangle_{j \in g^{-1}(k)} \vdash s_k : C \right\rangle_{k \in \mathcal{K}}}{\Gamma \vdash \delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle) : C} \dagger \text{E}
\end{array}$$

We write $\Gamma \vdash_{\text{NI}^{2p}} t : A$ iff there is a derivation of $\Gamma \vdash t : A$ using the above rules. Similarly for $\Gamma \vdash_{\text{NI}^2} t : A$ and all other systems introduced in Sections 2.1 and 2.2.

B. Contexts

Let $\mathcal{A}$ be a finite (possibly empty) list. By an $\mathcal{A}$ -*multicontext* T (henceforth simply referred to as context, or term context to avoid confusion with typing contexts) we indicate a term containing exactly one occurrence of $|\mathcal{A}|$ distinct special variables noted as $[\]_a$ (and referred to as “holes”), for each $a \in \mathcal{A}$.

Given an $\mathcal{A}$ -multicontext T , we let $\mathsf{T}[t]$ (resp. $\mathsf{T}[\mathsf{U}]$) indicate usual—i.e. non variable-capturing—substitution (short n.v.c.-substitution) of the term t (resp. context U) for each $[\]_a$ in T , and $\mathsf{T}\{t\}$ (respectively $\mathsf{T}\{\mathsf{U}\}$) indicate *variable-capturing* substitution (short v.c.-substitution) of the term t (resp. context U) for $[\]_a$ in T .

We will use the following special families of contexts:

- the *principal contexts* (indicated as $\mathsf{C}, \mathsf{D}, \dots$) are defined by the grammar below:

$$\mathsf{C} := [\] \mid \mathsf{Cu} \mid \mathsf{CB} \mid \delta_{\dagger}(\mathsf{C}, \langle \langle y_j \rangle . s_k \rangle) \mid \lambda x. \mathsf{C} \mid \Lambda Y. \mathsf{C} \mid \iota_{\dagger}^k \langle t_1, \dots, t_{l-1}, \mathsf{C}, t_{l+1}, \dots, t_{|g^{-1}(k)|} \rangle$$

- the *introduction contexts* are defined by dropping the cases Cu , CB and $\delta_{\dagger}(\mathsf{E} \langle \langle y_j \rangle . s_k \rangle)$ from the grammar above;
- the *elimination contexts* are defined by dropping the cases $\lambda x. \mathsf{C}$, $\Lambda Y. \mathsf{C}$ and $\iota_{\dagger}^k \langle t_1, \dots, t_{l-1}, \mathsf{C}, t_{l+1}, \dots, t_{|g^{-1}(k)|} \rangle$ from the grammar above;
- the *atomic elimination contexts* are defined by restricting B to be a type variable in the grammar for elimination contexts.

Observe that all principal contexts are contexts in the standard sense (i.e. they all contain exactly one occurrence of the hole). Note that if $\mathcal{C}, \mathcal{D}$ are principal contexts, then $\mathcal{C}\{\mathcal{D}\}$ and $\mathcal{C}[\mathcal{D}]$ are principal contexts as well.

It is easily checked that if $\mathcal{E}$ is an elimination context, then for all terms t , $\mathcal{E}\{t\}$ cannot capture variables of t , whence $\mathcal{E}\{t\} = \mathcal{E}[t]$.

We let $\mathsf{T} : A \vdash^\Gamma B$ be a shorthand for $\Gamma, x : A \vdash \mathsf{T}[x] : B$ (where x is a fresh term variable). Note that if $\mathsf{T} : A \vdash^\Gamma B$ and $\Gamma \vdash t : A$, $\Gamma \vdash \mathsf{T}[t] : B$ holds, but $\Gamma \vdash \mathsf{T}\{t\} : B$ in general does not hold. At crucial points below we will however rely on an analogous (but weaker) property for v.c.-substitution (namely Fact B.2 below), which requires a more fine-grained control over v.c.-substitution in order to be formulated and proven. The remaining part of this appendix is devoted to making this formally precise.

Given an $\mathcal{A}$ -multicontext T and an $\mathcal{A}$ -indexed finite list of terms $\langle u_a \rangle_{a \in \mathcal{A}}$, we let $\mathsf{T}\{\langle u_a \rangle_{a \in \mathcal{A}}\}$ be the term obtained by simultaneous variable-capturing substitution of $[\]_a$ by u_a . (Note that $\mathsf{T}\{t\}$ is just an abbreviation of $\mathsf{T}\{\langle u_a \rangle_{a \in \mathcal{A}}\}$ when $u_a = t$ for all $a \in \mathcal{A}$.)

We now introduce a refined typing for contexts, where a refined typing $\mathsf{T} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma B$ indicates, informally speaking, that in $\mathsf{T}\{\langle u_a \rangle_{a \in \mathcal{A}}\}$ the variables declared in Δ_a are bound in the term u_a inserted in the a -indexed hole in T . Refined typing judgments are defined inductively by appropriate rules. Being rather obvious, we omit some of them. In the refined typing rules, when $\mathcal{A}$ and $\mathcal{B}$ are lists we indicate with $\mathcal{A} * \mathcal{B}$ the concatenation of $\mathcal{A}$ and $\mathcal{B}$.

REMARK B.1. The refined typing judgements could be generalized to the following: $\mathsf{T} : \langle A_a \rangle_{a \in \mathcal{A}} \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma B$, i.e. the holes could have different types, and an opportune generalization of Fact B.2 below would hold for this generalized refined typing. As we make no use of the more general case, to avoid additional notational overhead we restrict our presentation to the relevant special case.

$$\begin{array}{c}
\frac{}{[\] : A \vdash_{\{\emptyset\}}^\Gamma A} \text{Ax}_1 \qquad \frac{}{y : A \vdash_{\langle y \rangle}^{y:A, \Gamma} A} \text{Ax}_2 \\
\\
\frac{\mathsf{T} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^{\Gamma, y:C} B}{\lambda y. \mathsf{T} : A \vdash_{\langle \Delta_a \cup \{y:C\} \rangle_{a \in \mathcal{A}}}^\Gamma C \supset B} \supset \text{I} \qquad \frac{\mathsf{T} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma C \supset B \quad \mathsf{U} : A \vdash_{\langle \Delta_b \rangle_{b \in \mathcal{B}}}^\Gamma C}{\mathsf{TU} : A \vdash_{\langle \Delta_c \rangle_{c \in \mathcal{A} * \mathcal{B}}}^\Gamma B} \supset \text{E} \\
\\
\frac{\mathsf{T} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma B}{\lambda Y. \mathsf{T} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma \forall Y. B} \forall \text{I (Proviso: } Y \notin FV(\Gamma)) \qquad \frac{\mathsf{T} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma \forall X. B}{\mathsf{TC} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma B[C/X]} \forall \text{E}
\end{array}$$

Observe that if $\mathsf{T} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma B$ holds, then $\Gamma \cup \Delta_a$ is a well-defined context, for any $a \in \mathcal{A}$.

The following fact can be easily established by induction on the derivation of $\mathsf{T} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma B$ (for some explanatory remarks, see below):

FACT B.2. *If $\mathsf{T} : A \vdash_{\langle \Delta_a \rangle_{a \in \mathcal{A}}}^\Gamma B$ and $\langle u_a \rangle_{a \in \mathcal{A}}$ is an $\mathcal{A}$ -indexed family of terms such that $\Gamma \cup \Delta_a \vdash u_a : A$ for all $a \in \mathcal{A}$, then $\Gamma \vdash \mathsf{T}\{\langle u_a \rangle_{a \in \mathcal{A}}\} : B$.*

REMARK B.3. We observe that:

- the conditions for applying the induction hypothesis in the case of rules that discharge hypothesis are guaranteed by the fact that the variable(s) to be bound swaps between Γ and the Δ s in passing from the premise to the conclusion of the refined typing rules;
- it is crucial that the refined typings keep apart the variables captured by the distinct occurrences of the hole: Had one merged all the Δ_a s in each refined typing into a single Δ , the inductive steps for rules with more than one premise would not go through, since the fact that e.g., in the case of $\supset \mathsf{E}$, $\Gamma \cup \Delta_1 \cup \Delta_2 \vdash u_a : A$ would guarantee neither $\Gamma \cup \Delta_1 \vdash u_a : A$ nor $\Gamma \cup \Delta_2 \vdash u_a : A$;
- in the refined typing judgments there is no need of keeping track also of bound propositional variables, as they play no role in the proof;
- in the case in which the same term t is inserted in all holes of a context, the premise of the lemma is to be understood as meaning that $FV(t) \subseteq \bigcap_{a \in \mathcal{A}} (FV(\Gamma) \cup FV(\Delta_a))$ (where $FV(t)$ indicates the set of free variables of t).

C. The RP-Translation

DEFINITION C.1. (*RP-translation*) Given the RP-translation of formulas (Definition 2.7 in Section 2.2), for every term u such that $\Gamma \vdash_{\mathsf{NI}^{2p}} u : A$ we define a term u^* as follows:

$$\left. \begin{array}{l} x^* = x \\ (\lambda x.t)^* = \lambda x.t^* \\ (ts)^* = t^* s^* \\ (\Lambda X.t)^* = \Lambda X.t^* \\ (tB)^* = t^* B^* \end{array} \right| \begin{array}{l} (\iota_{\dagger}^{k_0} \langle t_j \rangle)^* = \Lambda X. \lambda \langle x_k \rangle. x_{k_0} \langle t_j^* \rangle \\ (\delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle))^* = t^* A^* \langle \lambda \langle y_j \rangle . s_k^* \rangle \end{array}$$

The following fact is easily checked by induction on a derivation of $\Gamma \vdash_{\mathsf{NI}^{2p}} u : A$:

FACT C.2. *If $\Gamma \vdash_{\mathsf{NI}^{2p}} u : A$, then $\Gamma^* \vdash_{\mathsf{NI}^2} u^* : A^*$ and if $\Gamma \vdash_{\mathsf{NI}^p} u : A$, then $\Gamma^* \vdash_{\mathsf{NI}_{RP}^2} u^* : A^*$.*

REMARK C.3. The clauses for ι - and δ -terms generalize the standard translation of disjunction and conjunction constructors in System F (where $\iota_\wedge(t_1, t_2)$ is prefix notation for the standard pairing).

$$\begin{aligned} (\iota_\vee^u t)^* &= \Lambda X. \lambda x_1 x_2. x_u t^* \quad (u = 1, 2) \quad (\delta_\vee(t, y.s_1, y.s_2))^* = t^* A^* (\lambda y.s_1^*) (\lambda y.s_2^*) \\ (\iota_\wedge(t_1, t_2))^* &= \Lambda X. \lambda x. x t_1^* t_2^* \quad (\delta_\wedge(t, y_1 y_2. s))^* = t^* A^* \lambda y_1 y_2. s^* \end{aligned}$$

REMARK C.4. It can be checked that if $\mathbb{C} : A \vdash^\Gamma B$ is a context in the language of NI^{2p} , then $\mathbb{C}^* : A^* \vdash^{\Gamma^*} B^*$ is a context in the language of NI^2 . Moreover, if $\mathbb{E}$ is an elimination context then $\mathbb{E}^*$ is also an elimination context.

D. The Standard Equivalence on Derivations

The rules of equivalence are the following:

$$\begin{aligned} & \frac{\Gamma, x : A \vdash t : B \quad \Gamma \vdash s : A}{\Gamma \vdash (\lambda x. t) s \simeq t[s/x] : B} \supset \beta \quad \frac{\Gamma \vdash t : A \supset B}{\Gamma \vdash t \simeq \lambda x. tx : A \supset B} \supset \eta \quad (x \notin FV(t)) \\ & \frac{\Gamma \vdash t : A}{\Gamma \vdash (\Lambda X. t) B \simeq t[B/X] : A[B/X]} \forall \beta \quad (X \notin FV(\Gamma)) \quad \frac{\Gamma \vdash t : \forall X. A}{\Gamma \vdash t \simeq \Lambda X. tX : \forall X. A} \forall \eta \quad (X \notin FV(t)) \\ & \left\langle \frac{\langle \Gamma \vdash t_j : A_{f(j)} \rangle_{j \in g^{-1}(h)} \quad \langle \Gamma, \langle y_j : A_{f(j)} \rangle_{j \in g^{-1}(k)} \vdash s_k : C \rangle_{k \in \mathcal{K}}}{\Gamma \vdash \delta_\dagger(\iota_\dagger^h \langle t_j \rangle_{j \in g^{-1}(h)}, \langle \langle y_j \rangle_{j \in g^{-1}(k)}, s_k \rangle_{k \in \mathcal{K}}) \simeq s_h[\langle t_j \rangle / \langle y_j \rangle] : C} \dagger \beta_h \right\rangle_{h \in \mathcal{K}} \\ & \frac{\Gamma \vdash t : \dagger \langle A_i \rangle \quad \mathbb{U} : \dagger \langle A_i \rangle \vdash^\Gamma C}{\Gamma \vdash \delta_\dagger(t, \langle \langle y_j \rangle, \mathbb{U}[\iota_\dagger^k \langle y_j \rangle] \rangle_{k \in \mathcal{K}}) \simeq \mathbb{U}[t] : C} \dagger \gamma^+ \end{aligned}$$

together with reflexivity, transitivity, symmetry and congruence rules.

REMARK D.1. As in the case of disjunction (see, e.g. [29] and [17]) the “generalized” rule η^+ for $\dagger$ -connectives can be decomposed into a simple form of η -rule and a “generalized” permutation rule:

$$\begin{aligned} & \frac{\Gamma \vdash t : \dagger \langle A_i \rangle}{\Gamma \vdash \delta_\dagger(t, \langle \langle y_j \rangle, \iota_\dagger^k \langle y_j \rangle \rangle_{k \in \mathcal{K}}) \simeq t : \dagger \langle A_i \rangle} \dagger \eta \\ & \frac{\Gamma \vdash t : \dagger \langle A_i \rangle \quad \langle \Gamma, \langle y_j : A_{f(j)} \rangle \vdash s_k : C \rangle_{k \in \mathcal{K}} \quad \mathbb{U} : C \vdash^\Gamma D}{\Gamma \vdash \mathbb{U}[\delta_\dagger(t, \langle \langle y_j \rangle, s_k \rangle_{k \in \mathcal{K}})] \simeq \delta_\dagger(t, \langle \langle y_j \rangle, \mathbb{U}[s_k] \rangle_{k \in \mathcal{K}}) : D} \dagger \gamma^+ \end{aligned}$$

The rule $\dagger \gamma$ expressing the standard permuting conversions used in establishing the subformula of normal derivations in NI is the special case of $\dagger \gamma^+$ in which the context $\mathbb{U}$ consists in a single elimination rule (with the hole $[]$ corresponding to the major premise).

We write

$$\begin{array}{l} \Gamma \vdash_{\mathbf{NI}^{2p}} t \simeq_{\beta} s : A \\ \Gamma \vdash_{\mathbf{NI}^{2p}} t \simeq_{\eta^+} s : A \\ \Gamma \vdash_{\mathbf{NI}^{2p}} t \simeq_{\eta} s : A \\ \Gamma \vdash_{\mathbf{NI}^{2p}} t \simeq_{\gamma^+} s : A \\ \Gamma \vdash_{\mathbf{NI}^{2p}} t \simeq_{\gamma} s : A \end{array} \left| \begin{array}{l} \text{iff there is a derivation of} \\ \Gamma \vdash t \simeq s \text{ using all the typing} \\ \text{rules of } \mathbf{NI}^{2p} \text{ together with} \end{array} \right. \begin{array}{l} \supset \beta, \forall \beta, \dagger \beta_h \\ \supset \eta, \forall \eta, \dagger \eta^+ \\ \supset \eta, \forall \eta, \dagger \eta \\ \dagger \gamma^+ \\ \dagger \gamma \end{array}$$

as well as reflexivity, transitivity, symmetry and the congruence rules. With e.g. $\Gamma \vdash_{\mathbf{NI}^{2p}} t \simeq_{\beta\eta} s : A$ we indicate that there is a derivation using the rules underlying both $\Gamma \vdash_{\mathbf{NI}^{2p}} t \simeq_{\beta} s : A$ and $\Gamma \vdash_{\mathbf{NI}^{2p}} t \simeq_{\eta} s : A$, and similarly for other “combined” labels.

Similar notation will be used for sub-systems of $\mathbf{NI}^{2p}$. If the derivations do not use symmetry we write $\rightsquigarrow$ in place of $\simeq$.

REMARK D.2. The rules $\dagger \eta^+$ and $\dagger \gamma^+$ above are formulated using n.v.c.-substitution of a term in the context $\mathbf{U}$. One can also consider variants $\dagger \eta_{\text{v.c.}}^+$ and $\dagger \gamma_{\text{v.c.}}^+$ of these rules with v.c.-substitution in place of n.v.c.-substitution (Fact B.2 crucially warrants that both terms in the conclusion of the two rules can be typed by C and D respectively). However, these new rules are derivable from $\dagger \eta^+$ and $\dagger \gamma^+$ (*together with the other conversion rules and symmetry*). The proof of the derivability of the v.c.-variants is by induction on the context $\mathbf{U}$, and the only critical cases are those in which $\mathbf{U} = \lambda x.U'$, $\mathbf{U} = \Lambda X.U'$ or $\mathbf{U}$ is a $\delta_{\dagger}$ -term in which one or more of its immediate subterms in $\langle s_k \rangle$ are a context $\mathbf{U}'$. We give the details of the inductive case for $\dagger \gamma_{\text{v.c.}}^+$ in which $\mathbf{U} = \lambda x.U'$ by informally showing how to construct a derivation of $\Gamma \vdash \mathbf{U}\{\delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle)\} \simeq_{\dagger} \delta_{\dagger}(t, \langle \langle y_j \rangle . \mathbf{U}\{s_k\}\rangle) : G$ as follows:

$$\begin{aligned} \lambda x.U'\{\delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle)\} &\stackrel{\text{I.H.}}{\simeq} \lambda x.\delta_{\dagger}(t, \langle \langle y_j \rangle . \mathbf{U}'\{s_k\}\rangle) \simeq_{\supset \beta} \\ &\simeq_{\supset \beta} \lambda x.\delta_{\dagger}(t, \langle \langle y_j \rangle . (\lambda x.U'\{s_k\})x \rangle) \simeq_{\dagger \gamma} \\ &\simeq_{\dagger \gamma} \lambda x.\delta_{\dagger}(t, \langle \langle y_j \rangle . \lambda x.U'\{s_k\}\rangle)x \rightsquigarrow_{\supset \eta} \delta_{\dagger}(t, \langle \langle y_j \rangle . \lambda x.U'\{s_k\}\rangle) \end{aligned}$$

The other critical cases for $\dagger \gamma^+$ and those for $\dagger \eta^+$ are similar.

REMARK D.3. Assuming that the standard elimination rules for $\wedge$ are defined as $\pi_1(t) := \delta_{\wedge}(t, y_1 y_2 . y_1)$ and $\pi_2(t) := \delta_{\wedge}(t, y_1 y_2 . y_2)$ (and taking $\iota_{\wedge}(t_1, t_2)$ as prefix notation for the standard pairing), we can derive the standard η -rule for conjunction $\iota_{\wedge}(\pi_1(t), \pi_2(t)) \simeq t$ as follows:

$$\begin{array}{llll}
\iota_\wedge(\delta_\wedge(t, y_1 y_2 . y_1), \delta_\wedge(t, y_3 y_4 . y_4)) & \rightsquigarrow_{\dagger\gamma^+} & & \\
\rightsquigarrow_{\dagger\gamma^+} & \delta_\wedge(t, y_1 y_2 . \iota_\wedge(y_1, \delta_\wedge(t, y_3 y_4 . y_4))) & \rightsquigarrow_{\dagger\gamma^+} & \\
\rightsquigarrow_{\dagger\gamma^+} & \delta_\wedge(t, y_1 y_2 . \delta_\wedge(t, y_3 y_4 . \iota_\wedge(y_1, y_4))) & \simeq_{\dagger\eta} & \\
\simeq_{\dagger\eta} & \delta_\wedge\left(\left(\delta_\wedge(t, z_1 z_2 . \iota_\wedge(z_1, z_2))\right), y_1 y_2 . \delta_\wedge\left(\left(\delta_\wedge(t, z_1 z_2 . \iota_\wedge(z_1, z_2))\right), y_3 y_4 . \iota_\wedge(y_1, y_4)\right)\right) & \rightsquigarrow_{\dagger\gamma^+} & \\
\rightsquigarrow_{\dagger\gamma^+} & \delta_\wedge(t, z_1 z_2 . \delta_\wedge\left(\left(\iota_\wedge(z_1, z_2)\right), y_1 y_2 . \delta_\wedge\left(\left(\iota_\wedge(z_1, z_2)\right), y_3 y_4 . \iota_\wedge(y_1, y_4)\right)\right)) & \rightsquigarrow_{\dagger\beta} & \\
\rightsquigarrow_{\dagger\beta} & \delta_\wedge(t, z_1 z_2 . \delta_\wedge\left(\left(\iota_\wedge(z_1, z_2)\right), y_1 y_2 . \iota_\wedge(y_1, z_2)\right)) & \rightsquigarrow_{\dagger\beta} & \\
\rightsquigarrow_{\dagger\beta} & \delta_\wedge(t, z_1 z_2 . \iota_\wedge(z_1, z_2)) & \rightsquigarrow_{\dagger\eta} & \\
& & \rightsquigarrow_{\dagger\eta} t &
\end{array}$$

where the three applications of $\dagger\gamma^+$ can be described as:

- $\mathsf{U}[\delta_\wedge(t, y_1 y_2 . y_1)] \rightsquigarrow_{\gamma^+} \delta_\wedge(t, y_1 y_2 . \mathsf{U}[y_1])$ with $\mathsf{U} := \iota_\wedge([], \delta_\wedge(t, y_3 y_4 . y_4))$;
- $\mathsf{U}[\delta_\wedge(t, y_3 y_4 . y_4)] \rightsquigarrow_{\gamma^+} \delta_\wedge(t, y_3 y_4 . \mathsf{U}[y_4])$ with $\mathsf{U} := \iota_\wedge(y_1, [])$;
- $\mathsf{U}[\delta_\wedge(t, z_1 z_2 . \iota_\wedge(z_1, z_2))] \rightsquigarrow_{\gamma^+} \delta_\wedge(t, z_1 z_2 . \mathsf{U}[\iota_\wedge(z_1, z_2)])$ with $\mathsf{U} := \delta_\wedge([], y_1 y_2 . \delta_\wedge([], y_3 y_4 . \iota_\wedge(y_1, y_4)))$.

E. Weak Expansion and A-Expansion

DEFINITION E.1. For any $A \in \mathcal{L}^2$ we define a family of atomic elimination contexts $\text{Elim}(A)$ and a family of introduction contexts $\text{Intro}(A)$ by induction as follows:

- If $A = Z$, then $\text{Elim}(A) = \text{Intro}(A) = \{[]\}$;
- If $A = B \rightarrow C$, then $\text{Elim}(A) = \{\mathsf{E}[[]x] \mid x \text{ not free in } \mathsf{E} \text{ and } \mathsf{E} \in \text{Elim}(C)\}$ and $\text{Intro}(A) = \{\lambda x. \mathsf{I} \mid \mathsf{I} \in \text{Intro}(C) \text{ and } x \text{ does not occur in } \mathsf{I}\}$
- If $A = \forall Z. C$, then $\text{Elim}(A) = \{\mathsf{E}[[]X] \mid X \text{ not free in } \mathsf{E} \text{ and } \mathsf{E} \in \text{Elim}(C)\}$ and $\text{Intro}(A) = \{\Lambda X. \mathsf{I} \mid \mathsf{I} \in \text{Intro}(C) \text{ and } X \text{ does not occur in } \mathsf{I}\}$

An *expansion pair* for A is any pair of contexts $(\mathsf{El}, \mathsf{In})$ such that $\mathsf{El} \in \text{Elim}(A)$, $\mathsf{In} \in \text{Intro}(A)$ and $\mathsf{In}\{\mathsf{El}\} \rightsquigarrow_\eta []$.

We list some useful and easily established facts about introduction and elimination contexts (recalling that for all elimination contexts E and terms t , $\mathsf{E}\{t\} = \mathsf{E}[t]$):

FACT E.2. If $\Gamma \vdash_{NT^2} t : A$ then for all expansion pairs $(\mathsf{El}, \mathsf{In})$ for A , $\Gamma \vdash t \simeq_\eta \mathsf{In}\{\mathsf{El}[t]\} : A$

FACT E.3. For all expansion pairs $(\mathsf{El}, \mathsf{In})$ for $A \supset B$ there exists $x \in TV$ such that $\mathsf{El} = \mathsf{E}[[]x]$ and $\mathsf{In} = \lambda x. \mathsf{I}$, where (E, I) is an expansion pair for

B. Similarly, for all expansion pairs $(E\mathbf{l}, I\mathbf{n})$ for $\forall X.A$ there exists $X \in \mathcal{V}$ such that $E\mathbf{l} = E[\]X$ and $I\mathbf{n} = \Lambda X.I$, where (E, I) is an expansion pair for A .

FACT E.4. *If (E_1, I_1) and (E_2, I_2) are two expansion pairs for A , then for any $U : \mathbf{at}(A) \vdash^\Gamma \mathbf{at}(A)$ and term $\Gamma \vdash t : A$, if no variable free in U and t is bound in either I_1 or I_2 , then $I_1\{U[E_1[t]]\} = I_2\{U[E_2[t]]\}$.*

In the following we will suppose fixed for any formula $A \in \mathcal{L}^2$ an expansion pair $(E\mathbf{l}_A, I\mathbf{n}_A)$.

DEFINITION E.5. (*Weak expansion and A-expansion*) For all $A \in \mathcal{L}^2$, the *weak expansion* of A is the context $I\mathbf{n}_A\{E\mathbf{l}_A\}$. Moreover, if A is $\text{sp-}X$, for all $U : B \vdash^\Gamma C$, we let the *A-expansion* of U be $A^x U = I\mathbf{n}_A\{U[E\mathbf{l}_A]\}$.

Observe that by Fact E.4, the context $A^x U$ does not depend on the chosen expansion pair for A . Moreover, it is easily checked that for all $\text{sp-}X$ formulas $A \in \mathcal{L}^2$, if $U : B \vdash^\Gamma C$ then $A^x U : A[B/X] \vdash^\Gamma A[C/X]$.

REMARK E.6. In [33] we defined the notion of A -expansion in a more direct manner for the broader class of strictly positive formulas (and not, as done here, strongly positive, see Remark 2.15). The interested reader can easily check that the two definitions coincide in the case of strongly positive formulas.

F. The ε -Equation

We write $\Gamma \vdash t \simeq_\varepsilon u : A$ iff there is a derivation of $\Gamma \vdash t \simeq u$ using the rules of congruence, symmetry, transitivity, reflexivity and any instance of the rule below:

$$\frac{\Gamma \vdash t : (\dagger(A_i))^* \quad \left\langle \Gamma \vdash u_k : ((A_{f(j)})^* \supset X) \llbracket C/X \rrbracket \right\rangle_{k \in \mathcal{K}} \quad U : C \vdash^\Gamma D}{\Gamma \vdash U[tC\langle u_k \rangle_{k \in \mathcal{K}}] \simeq tD \left\langle ((A_{f(j)})^* \supset X)^x U[u_k] \right\rangle_{k \in \mathcal{K}} : D} \varepsilon$$

For the use of e.g. $\simeq_{\eta\varepsilon}$ and $\rightsquigarrow_\varepsilon$ we adopt the same conventions introduced in Appendix D.

REMARK F.1. As for the rules $\dagger\eta^+$ and $\dagger\gamma^+$, we formulated the rule ε using n.v.c.-substitution of a term for the hole of U . Similarly to the other cases (see Remark D.2), a variant of ε with v.c.-substitution in place of n.v.c.-substitution can be derived from the rule ε along with β - and η -conversions and symmetry.

G. Proof of Proposition 1

We will establish the following strengthening of the generalization of Proposition 1 to NI^{2p} :

PROPOSITION 9. *For all u, v if $\Gamma \vdash_{\text{NI}^{2p}} u \rightsquigarrow_{\beta\eta\gamma+} v : A$ then $\Gamma^* \vdash_{\text{NI}^2} u^* \simeq_{\beta\eta\epsilon} v^* : A^*$. More precisely:*

1. *if $\Gamma \vdash_{\text{NI}^{2p}} u \rightsquigarrow_{\beta} v : A$ then $\Gamma^* \vdash_{\text{NI}^2} u^* \rightsquigarrow_{\beta} v^* : A^*$;*
2. *if $\Gamma \vdash_{\text{NI}^{2p}} u \rightsquigarrow_{\eta} v : A$ then $\Gamma^* \vdash_{\text{NI}^2} u^* \simeq_{\beta\eta\epsilon} v^* : A^*$;*
3. *if $\Gamma \vdash_{\text{NI}^{2p}} u \rightsquigarrow_{\gamma+} v : A$ then $\Gamma^* \vdash_{\text{NI}^2} u^* \rightsquigarrow_{\beta\epsilon} v^* : A^*$.*

We need the following two lemmas:

LEMMA G.1. *For all $U : A \vdash_{\text{NI}^{2p}}^{\Gamma} B$ and t such that $\Gamma \vdash_{\text{NI}^{2p}} t : A$, $(\mathcal{U}[t])^* = \mathcal{U}^*[t^*]$.*

PROOF. The lemma is easily established by induction on U . ■

LEMMA G.2. *For all $sp\text{-}X$ formulas $A \in \mathcal{L}^2$ such that $A = \langle A_a \rangle_{a \in \mathcal{A}} \supset X$, if $U : B \vdash^{\Gamma} C$ and $\Gamma, \langle x_a : A_a \rangle_{a \in \mathcal{A}} \vdash_{\text{NI}^2} t : B$, the following hold:*

$$\Gamma \vdash (\langle A_a \rangle_{a \in \mathcal{A}} \supset X)^X \mathcal{U}[\lambda \langle x_a \rangle_{a \in \mathcal{A}} . t] \rightsquigarrow_{\beta} \lambda \langle x_a \rangle_{a \in \mathcal{A}} . \mathcal{U}[t] : \langle A_a \rangle_{a \in \mathcal{A}} \supset C$$

Moreover, if no free variable of $\text{El}_{\langle A_a \rangle_{a \in \mathcal{A}} \supset X}$ is bound in U , then

$$\Gamma \vdash (\langle A_a \rangle_{a \in \mathcal{A}} \supset X)^X \mathcal{U}\{\lambda \langle x_a \rangle_{a \in \mathcal{A}} . t\} \rightsquigarrow_{\beta} \lambda \langle x_a \rangle_{a \in \mathcal{A}} . \mathcal{U}\{t\} : \langle A_a \rangle_{a \in \mathcal{A}} \supset C$$

PROOF. The lemma is easily established by induction on the length of the list $\mathcal{A}$. ■

PROPOSITION 10. *For all u, v if $\Gamma \vdash_{\text{NI}^{2p}} u \simeq_{\eta} v : A$ then $\Gamma^* \vdash_{\text{NI}^2} u^* \simeq_{\beta\eta\epsilon} v^* : A^*$.*

PROOF. We can argue by induction on the derivation $\mathcal{D}$ of $\Gamma \vdash_{\text{NI}^{2p}} u \simeq_{\eta} v : A$. The only non-trivial case is when $\mathcal{D}$ ends with an application of $\dagger\eta$. So suppose $u = \delta_{\dagger}(v, \langle \langle y_j \rangle . \iota_{\dagger}^k \langle y_j \rangle \rangle_k)$. Hence $\Gamma \vdash_{\text{NI}^{2p}} v : \dagger \langle A_i \rangle$ for some $\dagger$, and thus $\Gamma^* \vdash_{\text{NI}^2} v^* : \forall X. \langle \langle A_{f(j)}^* \rangle \supset X \rangle \supset X$, and we can construct a derivation $\mathcal{D}'$ of $\Gamma^* \vdash_{\text{NI}^2} u^* \simeq_{\beta\eta\epsilon} v^* : A^*$ as shown below:

$$v^* \simeq_{\eta} \Lambda X. \lambda \langle f_k \rangle_{k \in \mathcal{K}} . v^* X \langle \lambda \langle y_j \rangle . f_k \langle y_j \rangle \rangle_{k \in \mathcal{K}} = \mathbf{C} \{v^* X \langle \lambda \langle y_j \rangle . f_k \langle y_j \rangle \rangle_{k \in \mathcal{K}}\} \overset{\text{Rem. F.1}}{\simeq_{\beta\eta\epsilon}}$$

$$(v^* A^*) \left\langle (\langle A_{f(j)}^* \rangle \supset X)^X \mathbf{C} \{ \lambda \langle y_j \rangle . f_k \langle y_j \rangle \} \right\rangle_{k \in \mathcal{K}} \overset{\text{Lemma G.2}}{\rightsquigarrow_{\beta}}$$

$$(v^* A^*) \langle \lambda \langle y_j \rangle . \mathbf{C} \{ f_k \langle y_j \rangle \} \rangle \overset{\text{Lemma. G.1}^+ \text{ Def. C.1}}{=} u^*$$

where $\mathbf{C} = \Lambda X. \lambda \langle f_k \rangle_{k \in \mathcal{K}} . [\]$ and the variables X, f_k are chosen so as not to occur free in the contexts $\text{El}_{\langle A_{f(j)}^* \rangle \supset X}$. ■

PROPOSITION 11. *For all u, v if $\Gamma \vdash_{\text{NI}^{2p}} u \rightsquigarrow_{\gamma^+} v : A$ then $\Gamma^* \vdash_{\text{NI}^2} u^* \rightsquigarrow_{\beta\varepsilon} v^* : A^*$.*

PROOF. By induction on the typing derivation $\mathcal{D}$ of $\Gamma \vdash_{\text{NI}^{2p}} u \rightsquigarrow_{\gamma^+} v : A$. If $\mathcal{D}$ ends with an application of either reflexivity, transitivity or of one of the congruence rules, it is enough to apply the induction hypothesis to the derivations of the premises, and then applying the reflexivity, transitivity or the congruence rules.

If $\mathcal{D}$ ends with an application of $\dagger\gamma^+$, then for some context $\mathbf{U} : B \vdash^\Gamma A$, $u = \mathbf{U}[\delta_\dagger(t, \langle \langle y_j \rangle . s_k \rangle)]$ and $v = \delta_\dagger(t, \langle \langle y_j \rangle . \mathbf{U}[s_k] \rangle)$, where $\Gamma \vdash_{\text{NI}^{2p}} \delta_\dagger(t, \langle \langle y_j \rangle . s_k \rangle) : B$, $\Gamma \vdash_{\text{NI}^{2p}} t : \dagger \langle A_i \rangle$ for some $\dagger$, and $\Gamma, \langle y_j : A_{f(j)} \rangle \vdash_{\text{NI}^{2p}} s_k : B$ for every $k \in \mathcal{K}$. Thus $\Gamma^* \vdash_{\text{NI}^2} t^* : \forall X. \langle \langle A_{f(j)}^* \rangle \supset X \rangle \supset X$, $\Gamma^*, \langle y_j : A_{f(j)}^* \rangle \vdash_{\text{NI}^2} s_k^* : B^*$ for every $k \in \mathcal{K}$ and we can construct a derivation $\mathcal{D}'$ of $\Gamma^* \vdash_{\text{NI}^2} u^* \rightsquigarrow_{\beta\varepsilon} v^* : A^*$ as shown below:

$$\begin{aligned} u^* &\stackrel{\text{Def. C.1} + \text{Lemma G.1}}{=} \mathbf{U}^* [(t^* B^*) \langle \lambda \langle y_j \rangle . s_k^* \rangle] \rightsquigarrow_\varepsilon \\ (t^* A^*) \left\langle \left(\langle \langle A_{f(j)}^* \rangle \supset X \rangle^X \right) \mathbf{U}^* [\lambda \langle y_j \rangle . s_k^*] \right\rangle &\stackrel{\text{Lemma G.2}}{\rightsquigarrow_\beta} \\ (t^* A^*) \langle \lambda \langle y_j \rangle . \mathbf{U}^* [s_k^*] \rangle &\stackrel{\text{Lemma G.1} + \text{Def. C.1}}{=} v^* \end{aligned}$$

■

PROOF OF PROPOSITION 9. It suffices to check rule by rule that points 1–3 of the proposition hold. For point 1., the only critical case is $\dagger\beta$ and its verification is a straightforward generalization of the preservation of the β -reduction for disjunction by the RP-translation [15]. For points 2. and 3., the critical cases are $\dagger\eta$ and $\dagger\gamma^+$ which follow from Proposition 10 and Proposition 11 respectively. ■

H. The Three Translations into NI_{at}^2

DEFINITION H.1. (*Atomization*) For every term u such that $\Gamma \vdash_{\text{NI}_{RP}^2} u : A$ we define a term $u^\downarrow$ as follows (observe that when $u = tB$ then $\Gamma \vdash_{\text{NI}_{RP}^2} t : \forall W.F$ and since F is polynomial in W , by what we observed in Subsection 2.2, F can be written as $\langle \langle A_{f(j)}^* \rangle \supset W \rangle \supset W$, with the A_i having no occurrence of W):

$$\begin{array}{l|l} x^\downarrow = x & (tY)^\downarrow = t^\downarrow Y \\ (\lambda x.t)^\downarrow = \lambda x.t^\downarrow & (t(C \supset D))^\downarrow = \lambda \langle y_k \rangle . \lambda x. (tD)^\downarrow \langle \lambda \langle z_j \rangle . y_k \langle z_j \rangle x \rangle \\ (ts)^\downarrow = t^\downarrow s^\downarrow & (t(\forall X.D))^\downarrow = \lambda \langle y_k \rangle . \Lambda X. (tD)^\downarrow \langle \lambda \langle z_j \rangle . y_k \langle z_j \rangle X \rangle \\ (\Lambda X.t)^\downarrow = \Lambda X.t^\downarrow & \end{array}$$

DEFINITION H.2. (ε -atomization) For every term u such that $\Gamma \vdash_{\mathbf{NI}_{RP}^2} u : A$ we define a term $u^{\downarrow^\varepsilon}$ by replacing in Definition H.1 the last two clauses for terms of the form tB (i.e. with B non-atomic) with the following single clause (assuming $Z = \mathbf{at}(B)$):

$$(tB)^{\downarrow^\varepsilon} = \lambda \langle y_k \rangle . \mathbf{In}_B \left\{ t^{\downarrow^\varepsilon} Z \left\langle \left(\langle A_{f(j)}^* \rangle \supset W \right)^W \mathbf{El}_B [y_k] \right\rangle \right\}$$

It is easy to check by induction on the derivation of $\Gamma \vdash_{\mathbf{NI}_{RP}^2} u : A$ that the following holds:

FACT H.3. For $\natural \in \{\downarrow, \downarrow^\varepsilon\}$, if $\Gamma \vdash_{\mathbf{NI}_{RP}^2} u : A$, then $\Gamma \vdash_{\mathbf{NI}_{at}^2} u^\natural : A$.

From this and Fact C.2 it follows moreover the following:

COROLLARY H.4. For $\natural \in \{\downarrow, \downarrow^\varepsilon\}$, if $\Gamma \vdash_{\mathbf{NI}^p} u : A$, then $\Gamma^* \vdash_{\mathbf{NI}_{at}^2} u^{*\natural} : A^*$.

DEFINITION H.5. (ESF-translation) The ESF-translation is obtained by replacing in Definition C.1 $\mathbf{NI}^{2p}$ with $\mathbf{NI}^p$ and the last clause with the following (we assume here $\Gamma \vdash_{\mathbf{NI}^p} u : A$ and $Z = \mathbf{at}(A)$):

$$(\delta_\dagger(t, \langle \langle y_j \rangle . s_k \rangle))^\sharp = \mathbf{In}_{A^*} \left\{ (t^\sharp Z) \left\langle \lambda \langle y_j \rangle . \mathbf{El}_{A^*} [s_k^\sharp] \right\rangle \right\}$$

Also in this case, it is easily checked by induction on the derivation of $\Gamma \vdash_{\mathbf{NI}^p} u : A$ that the following holds:

FACT H.6. If $\Gamma \vdash_{\mathbf{NI}^p} u : A$, then $\Gamma^* \vdash_{\mathbf{NI}_{at}^2} u^\sharp : A^*$.

The interested reader can check that the refined interpretation proposed by Espírito Santo and Ferreira in [5] coincides essentially with ours in the case of $\mathbf{NI}^\vee$.

REMARK H.7. All three translations $*\downarrow, *\downarrow^\varepsilon, \sharp : \mathbf{NI}^p \rightarrow \mathbf{NI}_{at}^2$ extend straightforwardly to contexts, similarly to what happens for the RP-translation (see Remark C.4). Moreover, for each of the three translations $^\top$ it can be checked by induction that for all elimination contexts $\mathbf{E}$ of $\mathbf{NI}^p$, $\mathbf{E}^\top$ is a principal context and that $\mathbf{E}^\top\{t^\top\} = \mathbf{E}^\top[t^\top] = (\mathbf{E}[t])^\top$. Also note that for all terms t and s , $(t[s/x])^\top = t^\top[s^\top/x]$.

I. Proof of Proposition 2

We prove the following generalization to $\mathbf{NI}^p$ of Proposition 2 for the case $\natural = \sharp$:

PROPOSITION 12. For all u, v , if $\Gamma \vdash_{\mathbf{NI}^p} u \rightsquigarrow_{\beta\eta\gamma} v : A$ then $\Gamma^* \vdash_{\mathbf{NI}_{at}^2} u^\sharp \simeq_{\beta\eta} v^\sharp : A^*$. More precisely:

1. if $\Gamma \vdash_{NI^p} u \rightsquigarrow_{\beta} v : A$ then $\Gamma^* \vdash_{NI^2} u^{\sharp} \rightsquigarrow_{\beta\eta} v^{\sharp} : A^*$;
2. if $\Gamma \vdash_{NI^p} u \rightsquigarrow_{\eta} v : A$ then $\Gamma^* \vdash_{NI^2} u^{\sharp} \rightsquigarrow_{\beta\eta} v^{\sharp} : A^*$;
3. if $\Gamma \vdash_{NI^p} u \rightsquigarrow_{\gamma} v : A$ then $\Gamma^* \vdash_{NI^2} u^{\sharp} \simeq_{\beta} v^{\sharp} : A^*$.

To prove the proposition we will need the following lemma, whose formulation relies on the refined typing for contexts introduced in Appendix B.

LEMMA I.1. *Let*

- A, B be formulas of NI^2 ;
- $\langle u_a \rangle_{a \in \mathcal{A}}$ be a finite list of terms such that $\Gamma, \Delta_a \vdash_{NI_{at}^2} u_a : A$;
- $C : A \vdash^{\Gamma} B$ be a principal context of NI_{at}^2 ;
- $T : X \vdash^{\Gamma}_{\langle \Delta_a \rangle_{a \in \mathcal{A}}} X$ be an $\mathcal{A}$ -multicontext in NI^2 such that X does not occur free neither in Γ nor in the Δ_a .

Moreover, suppose that T binds no term or type variable of El_A and that In_A binds no term or type variable of T , C and of u_a for all $a \in \mathcal{A}$. Then

$$C \left[In_A \{ T_{at(A)} \{ \langle El_A[u_a] \rangle_{a \in \mathcal{A}} \} \} \right] \simeq_{\beta} In_B \left\{ T_{at(B)} \{ \langle El_B[C[u_a]] \rangle_{a \in \mathcal{A}} \} \right\}$$

where T_Y is shorthand for $T[Y/X]$.

PROOF. We argue by induction on the principal context C :

- if $C = []$ the claim is immediate;
- if $C = C'v$, then $C' : A \vdash^{\Gamma}_{NI_{at}^2} C \supset B$ and $\Gamma \vdash_{NI_{at}^2} v : C$, and noticing that $at(B) = at(C \supset B)$ we then have

$$\begin{aligned} C \left[In_A \{ T_{at(A)} \{ \langle El_A[u_a] \rangle_{a \in \mathcal{A}} \} \} \right] &= C' \left[In_A \{ T_{at(A)} \{ \langle El_A[u_a] \rangle_{a \in \mathcal{A}} \} \} \right] v \\ &\stackrel{I.H.}{\simeq_{\beta}} \left(In_{C \supset B} \{ T_{at(B)} \{ \langle El_{C \supset B}[C'[u_a]] \rangle_{a \in \mathcal{A}} \} \} \right) v = \left(\lambda x. In_B \{ T_{at(B)} \{ \langle El_B[C'[u_a]x] \rangle_{a \in \mathcal{A}} \} \} \right) v \\ &\rightsquigarrow_{\beta} In_B \{ T_{at(B)} \{ \langle El_B[C'[u_a]v] \rangle_{a \in \mathcal{A}} \} \} = In_B \{ T_{at(B)} \{ \langle El_B[C[u_a]] \rangle_{a \in \mathcal{A}} \} \} \end{aligned}$$

- if $C = C'W$, then $C' : A \vdash^{\Gamma} \forall Y. B'$, $B = B'[W/Y]$; by noticing that $(at(\forall Y. B'))[W/Y] = at(B)$ we then have

$$\begin{aligned} C \left[In_A \{ T_{at(A)} \{ \langle El_A[u_a] \rangle_{a \in \mathcal{A}} \} \} \right] &= C' \left[In_A \{ T_{at(A)} \{ \langle El_A[u_a] \rangle_{a \in \mathcal{A}} \} \} \right] W \\ &\stackrel{I.H.}{\simeq_{\beta}} \left(In_{\forall Y. B'} \{ T_{at(\forall Y. B')} \{ \langle El_{\forall Y. B'}[C'[u_a]] \rangle_{a \in \mathcal{A}} \} \} \right) W \\ &= \left(\lambda Y. In_{B'} \{ T_{at(\forall Y. B')} \{ \langle El_{B'}[C'[u_a]Y] \rangle_{a \in \mathcal{A}} \} \} \right) W \rightsquigarrow_{\beta} In_B \{ T_{at(B)} \{ \langle El_B[C'[u_a]W] \rangle_{a \in \mathcal{A}} \} \} \\ &= In_B \{ T_{at(B)} \{ \langle El_B[C[u_a]] \rangle_{a \in \mathcal{A}} \} \} \end{aligned}$$

- if $C = \lambda w. C'$, then $B = B_1 \supset B_2$ and $C' : A \vdash^{\Gamma, w:B_1} B_2$, so by noticing that $at(B) = at(B_2)$ we have

$$\begin{aligned}
\text{In}_B \left\{ \text{T}_{\text{at}(B)} \{ \langle \text{El}_B [\mathbf{C}[u_a]] \rangle_{a \in \mathcal{A}} \} \right\} &= \lambda w'. \text{In}_{B_2} \left\{ \text{T}_{\text{at}(B)} \{ \langle \text{El}_{B_2} [(\lambda w. \mathbf{C}'[u_a])w'] \rangle_{a \in \mathcal{A}} \} \right\} \\
&\rightsquigarrow_{\beta} \lambda w'. \text{In}_{B_2} \left\{ \text{T}_{\text{at}(B)} \{ \langle \text{El}_{B_2} [\mathbf{C}'[w'/w][u_a]] \rangle_{a \in \mathcal{A}} \} \right\} \stackrel{\text{I.H.}}{\simeq_{\beta}} \lambda w'. \mathbf{C}'[w'/w] \\
&\quad \left[\text{In}_A \{ \text{T}_{\text{at}(A)} \{ \langle \text{El}_A [u_a] \rangle_{a \in \mathcal{A}} \} \} \right] \\
&= \mathbf{C} \left[\text{In}_A \{ \text{T}_{\text{at}(A)} \{ \langle \text{El}_A [u_a] \rangle_{a \in \mathcal{A}} \} \} \right]
\end{aligned}$$

- if $\mathbf{C} = \Lambda W. \mathbf{C}'$, then $B = \forall W. B'$ and $\mathbf{C}' : A \vdash^{\Gamma} B'$, so by noticing that $\text{at}(B) = \text{at}(B')$ we have

$$\begin{aligned}
\text{In}_B \left\{ \text{T}_{\text{at}(B)} \{ \langle \text{El}_B [\mathbf{C}[u_a]] \rangle_{a \in \mathcal{A}} \} \right\} &= \Lambda W'. \text{In}_{B'} \left\{ \text{T}_{\text{at}(B)} \{ \langle \text{El}_{B'} [(\Lambda W. \mathbf{C}'[u_a])W'] \rangle_{a \in \mathcal{A}} \} \right\} \\
&\rightsquigarrow_{\beta} \Lambda W'. \text{In}_{B'} \left\{ \text{T}_{\text{at}(B)} \{ \langle \text{El}_{B'} [\mathbf{C}'[W'/W][u_a]] \rangle_{a \in \mathcal{A}} \} \right\} \stackrel{\text{I.H.}}{\simeq_{\beta}} \Lambda W'. \mathbf{C}'[W'/W] \\
&\quad \left[\text{In}_A \{ \text{T}_{\text{at}(A)} \{ \langle \text{El}_A [u_a] \rangle_{a \in \mathcal{A}} \} \} \right] \\
&= \mathbf{C} \left[\text{In}_A \{ \text{T}_{\text{at}(A)} \{ \langle \text{El}_A [u_a] \rangle_{a \in \mathcal{A}} \} \} \right]
\end{aligned}$$

■

PROPOSITION 13. For all u, v , if $\Gamma \vdash_{\text{NIP}} u \rightsquigarrow_{\gamma} v : A$, then $\Gamma^* \vdash_{\text{NI}_{\text{at}}^2} u^{\sharp} \simeq_{\beta} v^{\sharp} : A^*$.

PROOF. By induction on the typing derivation $\mathcal{D}$ of $\Gamma \vdash_{\text{NIP}} u \rightsquigarrow_{\gamma} v : A$. If $\mathcal{D}$ ends with an application of either reflexivity, transitivity or of one of the congruence rules, it is enough to apply the induction hypothesis to the derivations of the premises, and then applying the reflexivity, transitivity or the congruence rules.

If $\mathcal{D}$ ends with an application of $\dagger\gamma$, $u = \mathbf{E}[\delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle)]$ and $v = \delta_{\dagger}(t, \langle \langle y_j \rangle . \mathbf{E}[s_k] \rangle)$ for some elimination context $\mathbf{E} : B \vdash^{\Gamma} A$ (hence $\Gamma \vdash_{\text{NIP}} \delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle) : B$ and $\Gamma \vdash_{\text{NIP}} t : \dagger \langle A_i \rangle$ for some $\dagger$). We show how to construct a derivation $\mathcal{D}'$ of $\Gamma^* \vdash_{\text{NI}_{\text{at}}^2} u^{\sharp} \simeq_{\beta} v^{\sharp} : A^*$. Let $Z = \text{at}(A^*)$ and $Z' = \text{at}(B^*)$. By Remark H.7 we know that $\mathbf{E}^{\sharp}$ is a principal context and that $u^{\sharp} = (\mathbf{E}[\delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle)])^{\sharp} = (\mathbf{E}^{\sharp}[(\delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle))^{\sharp}]) = (\mathbf{E}^{\sharp}\{(\delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle))^{\sharp}\})$. Moreover, consider the $\mathcal{K}$ -multicontext $\mathbf{T} : X \vdash_{\langle \Delta_k \rangle_{k \in \mathcal{K}}}^{\Gamma} X$ given by $\mathbf{T} = (t^{\sharp} X) \langle \lambda \langle y_j \rangle_{j \in g^{-1}(k)} . [\]_k \rangle_{k \in \mathcal{K}}$, where $\Delta_k = \langle y_j : A_{f(j)} \rangle_{j \in g^{-1}(k)}$ and X is a fresh variable. The reader can check that all conditions of Lemma I.1 do hold with $\mathbf{E}^{\sharp}$ in place of $\mathbf{C}$, B^* , A^* in place of A, B and the $s_k^{\sharp}$ in place of the u_a . Then we have:

$$\begin{aligned}
u^{\sharp} &\stackrel{\text{Def. H.5}}{=} \mathbf{E}^{\sharp} \left[\text{In}_{B^*} \left\{ (t^{\sharp} Z') \left\langle \lambda \langle y_j \rangle . \text{El}_{B^*} [s_k^{\sharp}] \right\rangle \right\} \right] \\
&\stackrel{\text{Lemma I.1}}{\simeq_{\beta}} \text{In}_{A^*} \left\{ (t^{\sharp} Z) \left\langle \lambda \langle y_j \rangle . \text{El}_{A^*} \left[\mathbf{E}^{\sharp} [s_k^{\sharp}] \right] \right\rangle \right\} \stackrel{\text{Def. H.5}}{=} v^{\sharp}
\end{aligned}$$

■

PROPOSITION 14. For all u, v if $\Gamma \vdash_{\text{NIP}} u \rightsquigarrow_{\eta} v : A$ then $\Gamma^* \vdash_{\text{NI}_{\text{at}}^2} u^{\sharp} \rightsquigarrow_{\beta\eta} v^{\sharp} : A^*$.

PROOF. By induction on the typing derivation $\mathcal{D}$ of $\Gamma \vdash_{\text{NIP}} u \rightsquigarrow_{\eta} v : A$. If $\mathcal{D}$ ends with an application of either reflexivity, transitivity, one of the congruence rules or $\supset\eta$ it is enough to apply the induction hypothesis to the derivations of the premises, and then applying reflexivity, transitivity, the congruence rules and $\supset\eta$ or $\forall\eta$.

If $\mathcal{D}$ ends with an application of $\dagger\eta$, $u = \delta_{\dagger}(v, \langle \langle y_j \rangle . \iota_{\dagger}^k \langle y_j \rangle \rangle)$ (where $\Gamma \vdash_{\text{NIP}} v : \dagger \langle A_i \rangle$ for some $\dagger$), and we informally show how to construct a derivation $\mathcal{D}'$ of $\Gamma^* \vdash_{\text{NI}_{at}^2} u^{\#} \rightsquigarrow_{\beta\eta} v^{\#} : A^*$. By letting $\mathbf{C} = \Lambda X. \lambda \langle x_h \rangle_{h \in \mathcal{K}}. []$ and In_{A^*} be $\Lambda Z. \lambda \langle z_h \rangle_{h \in \mathcal{K}}. []$, we have that:

$$u^{\#} \stackrel{\text{Def. H.5}}{=} \text{In}_{A^*} \left\{ (v^{\#} Z) \langle \lambda \langle y_j \rangle . \text{El}_{A^*} [\mathbf{C} \{x_k \langle y_j \rangle\}] \rangle_{k \in \mathcal{K}} \right\} \rightsquigarrow_{\beta}$$

$$\rightsquigarrow_{\beta} \text{In}_{A^*} \left\{ (v^{\#} Z) \langle \lambda \langle y_j \rangle . z_h \langle y_j \rangle \rangle_{h \in \mathcal{K}} \right\} \rightsquigarrow_{\eta} \text{In}_{A^*} \left\{ (v^{\#} Z) \langle z_h \rangle_{h \in \mathcal{K}} \right\} \rightsquigarrow_{\eta} v^{\#} \quad \blacksquare$$

PROPOSITION 15. *For all u, v if $\Gamma \vdash_{\text{NIP}} u \rightsquigarrow_{\beta} v : A$ then $\Gamma^* \vdash_{\text{NI}_{at}^2} u^{\#} \rightsquigarrow_{\beta\eta} v^{\#} : A^*$.*

PROOF. By induction on the typing derivation $\mathcal{D}$ of $\Gamma \vdash_{\text{NIP}} u \rightsquigarrow_{\beta} v : A$. If $\mathcal{D}$ ends with an application of either reflexivity, transitivity, one of the congruence rules or $\supset\beta$ it is enough to apply the induction hypothesis to the derivations of the premises, and then applying reflexivity, transitivity, the congruence rules and $\supset\beta$ or $\forall\beta$.

If $\mathcal{D}$ ends with an application of $\dagger\beta_h$, $u = \delta_{\dagger}(\iota_{\dagger}^h \langle t_j \rangle_{j \in g^{-1}(h)}, \langle \langle y_j \rangle_{j \in g^{-1}(k)} . s_k \rangle_{k \in \mathcal{K}})$, and we informally show how to construct a derivation $\mathcal{D}'$ of $\Gamma^* \vdash_{\text{NI}_{at}^2} u^{\#} \rightsquigarrow_{\beta\eta} v^{\#} : A^*$ as follows:

$$\begin{aligned} u^{\#} &\stackrel{\text{Def. H.5}}{=} \text{In}_{A^*} \left\{ ((\Lambda X. \lambda \langle x_k \rangle_{k \in \mathcal{K}}. x_h \langle t_j^{\#} \rangle_{j \in g^{-1}(h)}) Z) \langle \lambda \langle y_j \rangle_{j \in g^{-1}(k)} . \text{El}_{A^*} [s_k^{\#}] \rangle_{k \in \mathcal{K}} \right\} \rightsquigarrow_{\beta} \\ &\rightsquigarrow_{\beta} \text{In}_{A^*} \left\{ \text{El}_{A^*} \left[s_h^{\#} \llbracket \langle t_j^{\#} \rangle / \langle y_j \rangle \rrbracket \right] \right\} \stackrel{\text{Def. E.1}}{\rightsquigarrow_{\eta}} s_h^{\#} \llbracket \langle t_j^{\#} \rangle / \langle y_j \rangle \rrbracket \stackrel{\text{Remark H.7}}{=} \\ &\quad (s_h \llbracket \langle t_j \rangle / \langle y_j \rangle \rrbracket)^{\#} = v^{\#} \quad \blacksquare \end{aligned}$$

PROOF OF PROPOSITION 12. It suffices to check rule by rule that points 1-3 of the proposition hold. For points 1, 2 and 3, the critical cases are $\dagger\beta_h$, $\dagger\eta$ and $\dagger\gamma$ (respectively), which follow from Propositions 15, 14 and 13 respectively. $\blacksquare$

J. Proof of Proposition 3

We prove the following generalization of Proposition 3 to NI_{RP}^2 :

PROPOSITION 16. *For all u such that $\Gamma \vdash_{NI_{RP}^2} u : A$ we have that $\Gamma \vdash_{NI_{RP}^2} u \simeq_{\eta\varepsilon} u^{\downarrow^\varepsilon} : A$.*

PROOF. By induction on the typing derivation $\mathcal{D}$ of $\Gamma \vdash_{NI_{RP}^2} u : A$. If $\mathcal{D}$ ends with an application of either Ax , $\supset\text{I}$, $\supset\text{E}$, $\forall\text{I}$ or one of $\forall\text{E}$ with an atomic witness, (i.e. u is either a variable or of the form $\lambda x.t$, ts , $\Lambda X.t$ or tY), then it is enough to use reflexivity or to apply the induction hypothesis to the immediate sub-derivations of $\mathcal{D}$ and use the congruence rules.

Otherwise $u = tB$ with B non atomic, where the same remarks as in Definition H.1 apply to t (in particular $\Gamma \vdash_{NI_{RP}^2} t : \forall W.F$, where F is of the form $F_1 \supset \dots \supset F_{|\mathcal{K}|} \supset W$), and we informally show how to construct a derivation $\mathcal{D}'$ of $\Gamma \vdash_{NI_{at}^2} u \simeq_{\eta\varepsilon} u^{\downarrow^\varepsilon} : A$. Assuming $Z = \text{at}(B)$ (so that $\text{El}_B : B \vdash^\Gamma Z$) we have that:

$$\begin{aligned} tB &\stackrel{\text{Fact E.2}}{\simeq_\eta} \text{In}_{F[B/W]} \{ \text{El}_{F[B/W]}[tB] \} \stackrel{\text{Fact E.3}}{=} \lambda \langle y_k \rangle . \text{In}_B \{ \text{El}_B[(tB) \langle y_k \rangle] \} \\ &\simeq_\varepsilon \lambda \langle y_k \rangle . \text{In}_B \left\{ (tZ) \left\langle \left((A_{f(j)})^* \supset Z \right)^Z \text{El}_B[y_k] \right\rangle \right\} \stackrel{\text{Definition H.2+I.H.}}{\simeq_{\eta\varepsilon}} (tB)^{\downarrow^\varepsilon} \end{aligned}$$

where in the penultimate step the rule ε is applied to the term in braces by taking El_B for U . ■

K. Proof of Proposition 4

We prove the following generalization of Proposition 4 to NI^P :

PROPOSITION 17. *For all u such that $\Gamma \vdash_{NI^P} u : A$: (1) $\Gamma^* \vdash_{NI_{at}^2} u^{*\downarrow} \rightsquigarrow_\beta u^{*\downarrow^\varepsilon} : A^*$; (2) $\Gamma^* \vdash_{NI_{at}^2} u^{*\downarrow^\varepsilon} \rightsquigarrow_\beta u^\sharp : A^*$.*

The proposition follows immediately from the following two:

PROPOSITION 18. *For all u such that $\Gamma \vdash_{NI_{RP}^2} u : A$ we have that: $\Gamma \vdash_{NI_{at}^2} u^{\downarrow} \rightsquigarrow_\beta u^{\downarrow^\varepsilon} : A$.*

PROOF. By induction on the typing derivation $\mathcal{D}$ of $\Gamma \vdash_{NI_{RP}^2} u : A$. If $\mathcal{D}$ ends with an application of either Ax , $\supset\text{I}$, $\supset\text{E}$, $\forall\text{I}$ or one of $\forall\text{E}$ with an atomic witness, (i.e. u is either a variable or of the form $\lambda x.t$, ts , $\Lambda X.t$ or tY), then it is enough to use reflexivity or to apply the induction hypothesis to the immediate sub-derivations of $\mathcal{D}$ and use the congruence rules.

Otherwise $u = tB$ with B non atomic, where the same remarks as in Definition H.1 apply to t , and we informally show how to construct a derivation $\mathcal{D}'$ of $\Gamma \vdash_{NI_{at}^2} u^{\downarrow} \rightsquigarrow_\beta u^{\downarrow^\varepsilon} : A$ by induction on B .

If $B = C \supset D$, assuming $Z = \text{at}(B)$, we have:

$$\begin{aligned}
& (tB)^{\downarrow} \stackrel{\text{Definition H.1+I.H}}{=} \lambda \langle v_k \rangle \lambda u. ((tD)^{\downarrow}) \langle \lambda \langle x_j \rangle . v_k \langle x_j \rangle u \rangle \\
& \stackrel{\text{Definition H.2}}{=} \lambda \langle v_k \rangle \lambda u. (\lambda \langle y_k \rangle . \text{In}_D \left\{ (t^{\downarrow}) Z \left\langle \left(\langle A_{f(j)}^* \rangle \supset Z \right)^Z \text{El}_D [y_k] \right\rangle \right\}) \langle \lambda \langle x_j \rangle . v_k \langle x_j \rangle u \rangle \\
& \rightsquigarrow_{\beta} \lambda \langle v_k \rangle \lambda u. \text{In}_D \left\{ (t^{\downarrow}) Z \left\langle \left(\langle A_{f(j)}^* \rangle \supset Z \right)^Z \text{El}_D [\lambda \langle x_j \rangle . v_k \langle x_j \rangle u] \right\rangle \right\} \\
& \stackrel{\text{Definition E.5}}{=} \lambda \langle v_k \rangle \lambda u. \text{In}_D \left\{ (t^{\downarrow}) Z \langle \lambda \langle w_j \rangle . \text{El}_D [(\lambda \langle x_j \rangle . v_k \langle x_j \rangle u) \langle w_j \rangle] \rangle \right\} \\
& \rightsquigarrow_{\beta} \lambda \langle v_k \rangle \lambda u. \text{In}_D \left\{ (t^{\downarrow}) Z \langle \lambda \langle w_j \rangle . \text{El}_D [v_k \langle w_j \rangle u] \rangle \right\} \\
& \stackrel{\text{Fact E.4}}{=} \lambda \langle v_k \rangle . \text{In}_{C \supset D} \left\{ (t^{\downarrow}) Z \langle \lambda \langle w_j \rangle . \text{El}_{C \supset D} [v_k \langle w_j \rangle] \rangle \right\} \\
& \stackrel{\text{Definition E.5}}{=} \lambda \langle v_k \rangle . \text{In}_{C \supset D} \left\{ (t^{\downarrow}) Z \left\langle \left(\langle A_{f(j)}^* \rangle \supset Z \right)^Z \text{El}_{C \supset D} [v_k] \right\rangle \right\} \stackrel{\text{Definition H.2}}{=} (tB)^{\downarrow}
\end{aligned}$$

One can argue in a similar way if $B = \forall U.D$ (just replace $C \supset D$ with $\forall U.D$ and u with U). ■

PROPOSITION 19. *For all u such that $\Gamma \vdash_{\text{NIP}} u : A : \Gamma^* \vdash_{\text{NI}_{at}^2} u^{*\downarrow} \rightsquigarrow_{\beta} u^{\sharp} : A^*$.*

PROOF. By induction on the typing derivation $\mathcal{D}$ of $\Gamma \vdash_{\text{NIP}} u : A$. If $\mathcal{D}$ ends with an application of either Ax , $\supset\text{I}$, $\supset\text{E}$ or $\dagger\text{I}_k$, (i.e. u is either a variable or of the form $\lambda x.t$, ts or $i_k^{\dagger} \langle t_j \rangle$), then it is enough to use reflexivity or to apply the induction hypothesis to the immediate sub-derivations of $\mathcal{D}$ and use the congruence rules.

Otherwise $u = \delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle)$ and hence (by Definition C.1) $u^* = (t^* A^*) \langle \lambda \langle y_j \rangle . s_k^* \rangle$, where the same remarks as in Definition H.1 apply to t^* , and we informally show how to construct a derivation $\mathcal{D}'$ of $\Gamma \vdash_{\text{NI}_{at}^2} u^{*\downarrow} \rightsquigarrow_{\beta} u^{\sharp} : A$.

If A is atomic, i.e. $A = Y$ we have that:

$$\begin{aligned}
& (\delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle))^{*\downarrow} \stackrel{\text{Definition C.1}}{=} ((t^* Y) \langle \lambda \langle y_j \rangle . s_k^* \rangle)^{\downarrow} \stackrel{\text{Definition H.2}}{=} (t^{*\downarrow} Y) \left\langle \lambda \langle y_j \rangle . s_k^{*\downarrow} \right\rangle \stackrel{\text{I.H.}}{\rightsquigarrow}_{\beta} \\
& (t^{\sharp} Y) \left\langle \lambda \langle y_j \rangle . s_k^{\sharp} \right\rangle \stackrel{\text{Definition E.1}}{=} \text{In}_Y \left\{ (t^{\sharp} Y) \left\langle \lambda \langle y_j \rangle . \text{El}_Y [s_k^{\sharp}] \right\rangle \right\} \stackrel{\text{Definition H.5}}{=} u^{\sharp}
\end{aligned}$$

Otherwise, assuming $Z = \text{at}(A^*)$ we have that:

$$\begin{aligned}
& (\delta_{\dagger}(t, \langle \langle y_j \rangle . s_k \rangle))^{*\downarrow} \stackrel{\text{Definition C.1}}{=} ((t^* A^*) \langle \lambda \langle y_j \rangle . s_k^* \rangle)^{\downarrow} \stackrel{\text{Definition H.2}}{=} \\
& = \left(\lambda \langle z_k \rangle . \text{In}_{A^*} \left\{ (t^{*\downarrow}) Z \left\langle \left(\langle A_{f(j)}^* \rangle \supset Z \right)^Z \text{El}_{A^*} [z_k] \right\rangle \right\} \right) \left\langle \lambda \langle y_j \rangle . s_k^{*\downarrow} \right\rangle \rightsquigarrow_{\beta} \\
& \rightsquigarrow_{\beta} \text{In}_{A^*} \left\{ (t^{*\downarrow}) Z \left\langle \left(\langle A_{f(j)}^* \rangle \supset Z \right)^Z \text{El}_{A^*} [\lambda \langle y_j \rangle . s_k^{*\downarrow}] \right\rangle \right\} \stackrel{\text{Definition E.5}}{=} \\
& = \text{In}_{A^*} \left\{ (t^{*\downarrow}) Z \left\langle \lambda \langle v_j \rangle . \text{El}_{A^*} [(\lambda \langle y_j \rangle . s_k^{*\downarrow}) \langle v_j \rangle] \right\rangle \right\} \rightsquigarrow_{\beta} \\
& \rightsquigarrow_{\beta} \text{In}_{A^*} \left\{ (t^{*\downarrow}) Z \left\langle \lambda \langle y_j \rangle . \text{El}_{A^*} [s_k^{*\downarrow}] \right\rangle \right\} \stackrel{\text{I.H.}}{\rightsquigarrow}_{\beta} u^{\sharp}
\end{aligned}$$

L. Proofs of Proposition 5, 6, 7 and 8

We establish the following generalization of Proposition 5 to the whole of NI^{2p} .

PROPOSITION 20. For all $A \in \mathcal{L}^{2p}$, $A \dashv\vdash_{NI^{2p}} A^*$.

PROOF. We can construct by induction on A term contexts $\mathbf{C} : A^* \vdash^\emptyset A$ and $\mathbf{D} : A \vdash^\emptyset A^*$. The only non trivial case is $A = \dagger\langle A_i \rangle$. In this case we have that:

$$\begin{aligned}\mathbf{C} &= [\] \dagger\langle A_i \rangle \langle \lambda \langle y_j \rangle . \iota_{\dagger}^k \langle \mathbf{C}_j[y_j] \rangle \rangle_{k \in \mathcal{K}} \\ \mathbf{D} &= \Lambda X. \delta_{\dagger}([\], \langle \langle y_j \rangle . \lambda \langle x_h \rangle_{h \in \mathcal{K}} . x_k \langle \mathbf{D}_j[y_j] \rangle \rangle)\end{aligned}$$

where by induction hypothesis $\mathbf{C}_j : A_{f(j)}^* \vdash^\emptyset A_{f(j)}$ and $\mathbf{D}_j : A_{f(j)} \vdash^\emptyset A_{f(j)}^*$ ■

To establish Proposition 6 we will exploit a sound and complete semantics for NI_{at}^{2v} from [30], that we briefly recall.

For any partially ordered set $\langle W, \leq \rangle$, let $W^\uparrow = \{a \subseteq W \mid \forall \alpha, \beta \in W (\alpha \in a \wedge \alpha \leq \beta \Rightarrow \beta \in a)\}$ the set of upward closed subsets of W .

A NI_{at}^{2v} -model is a tuple $\mathcal{M} = \langle W, \leq, \perp, D, g \rangle$ such that $\langle W, \leq, \perp \rangle$ is a partially ordered set with a bottom element $\perp$, D is a monotone function from W to $\wp(W^\uparrow)$ (i.e. $\alpha \leq \beta \Rightarrow D(\alpha) \subseteq D(\beta)$) and $g : \mathcal{V} \rightarrow \bigcup_{\alpha \in W} D(\alpha)$.

For any formula $A \in \mathcal{L}^{2v}$ and model $\mathcal{M} = \langle W, \leq, \perp, D, g \rangle$, for all $\alpha \in W$ such that $g(\mathcal{V}) \subseteq D(\alpha)$ (i.e. for all $X \in \mathcal{V}$, $g(X) \in D(\alpha)$), we let the relation $\alpha \Vdash_{\mathcal{M}} A$ be defined by:

- $\alpha \Vdash_{\mathcal{M}} X$ iff $\alpha \in g(X)$
- $\alpha \Vdash_{\mathcal{M}} A \supset B$ iff for all $\beta \geq \alpha$, if $\beta \Vdash_{\mathcal{M}} A$ then $\beta \Vdash_{\mathcal{M}} B$
- $\alpha \Vdash_{\mathcal{M}} A \vee B$ iff $\alpha \Vdash_{\mathcal{M}} A$ or $\alpha \Vdash_{\mathcal{M}} B$
- $\alpha \Vdash_{\mathcal{M}} \forall X. A$ iff for all $\beta \geq \alpha$ and $a \in D(\beta)$, $\beta \Vdash_{\mathcal{M}[X \mapsto a]} A$

where $\mathcal{M}[X \mapsto a]$ is $\langle W, \leq, \perp, D, g[X \mapsto a] \rangle$, with $g[X \mapsto a]$ being like g except that it sends X to a .

The following fact (easily proved by induction on A) will be used below:

FACT L.1. For any model $\mathcal{M} = \langle W, \leq, \perp, D, g \rangle$, $\alpha, \beta \in W$ and formula A , if $g(\mathcal{V}) \subseteq D(\alpha)$, $\alpha \leq \beta$ and $\alpha \Vdash_{\mathcal{M}} A$, then $\beta \Vdash_{\mathcal{M}} A$.

A model $\mathcal{M} = \langle W, \leq, \perp, D, g \rangle$ is called *regular* if $g(\mathcal{V}) \subseteq D(\perp)$. We let $\Gamma \models A$ if for any regular model $\mathcal{M}$, $\perp \Vdash_{\mathcal{M}} \Gamma$ implies $\perp \Vdash_{\mathcal{M}} A$.

PROPOSITION 21. (Soundness and completeness [30]) $\Gamma \vdash_{NI_{at}^{2v}} A$ iff $\Gamma \models A$.

We will exploit this sound and complete semantics for NI_{at}^{2v} to show that a few formulas (which are all provable in NI^{2v}) are not provable in NI_{at}^{2v} .

PROOF OF PROPOSITION 6. By Proposition 21, it suffices to show that $(\forall X. Y \vee Z) \not\models Y \vee Z$, and for this we will construct a model $\mathcal{M}$ such that

$\perp \models_{\mathcal{M}} \forall X.Y \vee Z$ and $\perp \not\models_{\mathcal{M}} Y \vee Z$. Let $\mathcal{M} = \langle W, \leq, \perp, D, g \rangle$ be the regular model given by:

- $W = \{\perp, \alpha, \beta\}$, $\leq$ is reflexive, $\perp \leq \alpha, \beta$;
- $D(\perp) = D(\alpha) = D(\beta) = \{\{\alpha\}, \{\beta\}\}$;
- $g(Y) = \{\alpha\}$, $g(Z) = \{\beta\}$ and $g(W)$ is chosen arbitrarily in $D(\perp)$ for $W \neq Y, Z$.

We have $\alpha, \beta \models_{\mathcal{M}} Y \vee Z$ but $\perp \not\models_{\mathcal{M}} Y \vee Z$. Given $a \in D(\perp) = D(\alpha) = D(\beta)$, we have the following facts (where $\alpha \models_{\mathcal{M}} A_1, \dots, A_n$ is short for $\alpha \models_{\mathcal{M}} A_1$ and, ..., and $\alpha \models_{\mathcal{M}} A_n$, and similarly for $\not\models$):

- if $a = \{\alpha\}$, then
 - $-\alpha \models_{\mathcal{M}[X \mapsto a]} Y \supset X, Z \supset X, X$
 - $-\beta \models_{\mathcal{M}[X \mapsto a]} Y \supset X$
 - $-\beta \not\models_{\mathcal{M}[X \mapsto a]} Z \supset X, X$
- if $a = \{\beta\}$, then
 - $-\alpha \models_{\mathcal{M}[X \mapsto a]} Z \supset X,$
 - $-\alpha \not\models_{\mathcal{M}[X \mapsto a]} Y \supset X, X$
 - $-\beta \models_{\mathcal{M}[X \mapsto a]} Y \supset X, Z \supset X, X$

From these facts we deduce in turn:

- if $a = \{\alpha\}$, then
 - $-\perp \models_{\mathcal{M}[X \mapsto a]} Y \supset X$ (since $\perp \not\models_{\mathcal{M}[X \mapsto a]} Y$, $\beta \not\models_{\mathcal{M}[X \mapsto a]} Y$ and $\alpha \models_{\mathcal{M}[X \mapsto a]} Y, X$)
 - $-\perp \not\models_{\mathcal{M}[X \mapsto a]} Z \supset X, X$ (since $\beta \models_{\mathcal{M}[X \mapsto a]} Z$ but $\beta \not\models_{\mathcal{M}[X \mapsto a]} X$)
- if $a = \{\beta\}$, then
 - $-\perp \models_{\mathcal{M}[X \mapsto a]} Z \supset X$ (since $\perp \not\models_{\mathcal{M}[X \mapsto a]} Z$, $\alpha \not\models_{\mathcal{M}[X \mapsto a]} Z$ and $\beta \models_{\mathcal{M}[X \mapsto a]} Z, X$)
 - $-\perp \not\models_{\mathcal{M}[X \mapsto a]} Y \supset X, X$ (since $\alpha \models_{\mathcal{M}[X \mapsto a]} Y$ but $\alpha \not\models_{\mathcal{M}[X \mapsto a]} X$)

We now deduce that

- if $a = \{\alpha\}$, then $\perp \models_{\mathcal{M}[X \mapsto a]} Y \vee Z$, since the only $\gamma \geq \perp$ such that $\gamma \models_{\mathcal{M}[X \mapsto a]} Y \supset X$ and for all $\gamma' \geq \gamma$, $\gamma' \models_{\mathcal{M}[X \mapsto a]} Z \supset X$ is α and $\alpha \models_{\mathcal{M}[X \mapsto a]} X$;
- if $a = \{\beta\}$, then $\perp \models_{\mathcal{M}[X \mapsto a]} Y \vee Z$, since the only $\gamma \geq \perp$ such that $\gamma \models_{\mathcal{M}[X \mapsto a]} Y \supset X$ and for all $\gamma' \geq \gamma$, $\gamma' \models_{\mathcal{M}[X \mapsto a]} Z \supset X$ is β and $\beta \models_{\mathcal{M}[X \mapsto a]} X$.

We deduce then (using Fact L.1) that $\perp \models_{\mathcal{M}} \forall X.Y \vee Z$, and since $\perp \not\models_{\mathcal{M}} Y \vee Z$, we conclude that $(\forall X.Y \vee Z) \not\models Y \vee Z$. ■

PROPOSITION 22. Let $A = \forall X. \forall Y. X \supset Y$. The formulas A , $A \supset Z$ and $\exists X. A \supset X$ (that is, $\forall Y. (\forall X. (A \supset X) \supset Y) \supset Y$) are not provable in $\mathbf{NI}_{at}^2$.

PROOF. We will construct counter-models for each of the formulas above:

- let $\mathcal{M}$ be the model defined in the proof of Proposition 6 above. If $\perp \Vdash_{\mathcal{M}} A$ holds, then by definition we deduce that $\alpha \Vdash_{\mathcal{M}[X \mapsto \{\alpha\}, Y \mapsto \{\beta\}]} X \supset Y$, but this is impossible since $\alpha \Vdash_{\mathcal{M}[X \mapsto \{\alpha\}, Y \mapsto \{\beta\}]} X$ but $\alpha \not\Vdash_{\mathcal{M}[X \mapsto \{\alpha\}, Y \mapsto \{\beta\}]} Y$.
- Let $W = \{\perp, \alpha\}$, $D(\perp) = D(\alpha) = \{\{\alpha\}\}$ and let g constantly return $\{\alpha\}$. Observe that for all $\beta \in W$ and $a, b \in D(\beta)$, $\beta \Vdash_{\mathcal{M}[X \mapsto a, Y \mapsto b]} X$ iff $\beta = \alpha$ iff $\beta \Vdash_{\mathcal{M}[X \mapsto a, Y \mapsto b]} Y$, since the only possibility is that $a = b = \{\alpha\}$. From this we deduce that for all $\beta \in W$, $\beta \Vdash_{\mathcal{M}} A$, and, in particular, that $\perp \Vdash_{\mathcal{M}} A$. On the other hand, $\perp \not\Vdash_{\mathcal{M}} Z$, since $\perp \notin g(Z) = \{\alpha\}$. This is enough to conclude that $\perp \not\Vdash_{\mathcal{M}} A \supset Z$.
- Let $\mathcal{M}$ be the same model as in the case above. Suppose by contradiction that $\perp \Vdash_{\mathcal{M}} \exists X. A \supset X$; we deduce that $\perp \Vdash_{\mathcal{M}[Y \mapsto \{\alpha\}]} (\forall X. (A \supset X) \supset Y) \supset Y$. This implies that either $\perp \not\Vdash_{\mathcal{M}[Y \mapsto \{\alpha\}]} \forall X. (A \supset X) \supset Y$ or $\perp \Vdash_{\mathcal{M}[Y \mapsto \{\alpha\}]} Y$; since the latter is impossible, $\perp \not\Vdash_{\mathcal{M}[Y \mapsto \{\alpha\}]} \forall X. (A \supset X) \supset Y$ must hold, and this implies that either $\perp \not\Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} (A \supset X) \supset Y$ or $\alpha \not\Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} (A \supset X) \supset Y$:
 - if $\perp \not\Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} (A \supset X) \supset Y$, then for some $\beta \geq \perp$, $\beta \Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} A \supset X$ and $\beta \not\Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} Y$; if $\beta = \perp$, then we reach a contradiction since $\perp \Vdash_{\mathcal{M}[Y, X \mapsto \{\alpha\}]} A$ (one can argue similarly to the case above) and $\perp \not\Vdash_{\mathcal{M}[Y, X \mapsto \{\alpha\}]} X$; if $\beta = \alpha$, then we reach a contradiction since $\beta \Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} Y$.
 - if $\alpha \not\Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} (A \supset X) \supset Y$, then we deduce $\alpha \Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} A \supset X$ and $\alpha \not\Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} Y$, and we reach a contradiction since $\alpha \Vdash_{\mathcal{M}[X, Y \mapsto \{\alpha\}]} Y$.

Hence in any possible case we reach a contradiction, and we conclude that $\perp \not\Vdash_{\mathcal{M}} \exists X. A \supset X$. ■

PROOF OF PROPOSITION 7. Take $A = Y$, $B = Z$ and $C = Y \vee Z$. If there were an $\mathbf{NI}_{at}^{2\vee}$ -context $\mathbf{I0} : \forall X. Y \vee Z \vdash^{\emptyset} (Y \vee Z) \llbracket Y \vee Z / X \rrbracket$, then one could derive $Y \vee Z$ from $(Y \vee Z)^*$ by replacing $[\] \dagger \langle A_i \rangle$ with $\mathbf{I0}$ in the context $\mathbf{C}$ of the proof of proposition 5, thereby contradicting Proposition 6.

PROOF OF PROPOSITION 8. For all $\Gamma \subset \mathcal{L}^2$ such that $\Gamma \vdash_{\mathbf{NI}_{at}^{2\vee}} A \vee B$, either $\Gamma \vdash_{\mathbf{NI}_{at}^{2\vee}} A$ or $\Gamma \vdash_{\mathbf{NI}_{at}^{2\vee}} B$. This can be established following Prawitz' proof

of the generalized disjunction property of NI (see Corollary 6 ([22], p. 56)). We suppose (given the normalization theorem for $\text{NI}_{at}^{2\vee}$) the existence of a normal derivation of $\Gamma \vdash_{\text{NI}_{at}^{2\vee}} A \vee B$ and we reason as Prawitz. The only additional case to consider is that in which $A \vee B$ is the conclusion of a (sequence of) applications of $\forall E$, but due to the atomic restriction this case can be treated as the other cases of NI.

Now suppose $\vee$ is strongly definable in $\text{NI}_{at}^{2\vee}$ and let $C \Vdash_{\text{NI}_{at}^{2\vee}} (Y \vee Z)$. Then $Y \vdash_{\text{NI}_{at}^{2\vee}} C$, $Z \vdash_{\text{NI}_{at}^{2\vee}} C$ and $C \vdash_{\text{NI}_{at}^{2\vee}} Y \vee Z$ hold; by the above we deduce then that either $C \vdash_{\text{NI}_{at}^{2\vee}} Y$ or $C \vdash_{\text{NI}_{at}^{2\vee}} Z$ holds, and so we conclude that either $Z \vdash_{\text{NI}_{at}^{2\vee}} Y$ or $Y \vdash_{\text{NI}_{at}^{2\vee}} Z$, which is impossible. ■

References

- [1] BAINBRIDGE, E.S., P. J. FREYD, A. SCEDROV, and P. J. SCOTT, Functorial polymorphism, *Theoretical Computer Science* 70:35–64, 1990.
- [2] BARTHE, G., and M. H. SØRENSEN, Domain-free pure type systems, *Journal of Functional Programming* 10:417–452, 2000.
- [3] DINIS, B., and G. FERREIRA, Instantiation overflow, *Reports on Mathematical Logic* 51:15–33, 2016.
- [4] DOŠEN, K., Identity of proofs based on normalization and generality, *Bulletin of Symbolic Logic* 9(4):477–503, 2003.
- [5] ESPÍRITO SANTO, J., and G. FERREIRA, A refined interpretation of intuitionistic logic by means of atomic polymorphism, *Studia Logica* 108:477–507, 2020.
- [6] ESPÍRITO SANTO, J., and G. FERREIRA, The Russell-Prawitz embedding and the atomization of universal instantiation, *Logic Journal of the IGPL* 29(5):823–858, 2021.
- [7] FERREIRA, F., Comments on predicative logic, *Journal of Philosophical Logic* 35:1–8, 2006.
- [8] FERREIRA, F., and G. FERREIRA, Commuting conversions vs. the standard conversions of the “good” connectives, *Studia Logica* 92(1):63–84, 2009.
- [9] FERREIRA, F., and G. FERREIRA, Atomic polymorphism, *Journal of Symbolic Logic* 78(1):260–274, 2013.
- [10] FERREIRA, F., and G. FERREIRA, The faithfulness of atomic polymorphism, in *Proceedings of Trends in Logic XIII*, Łódź University Press, 2014, pp. 55–65.
- [11] FERREIRA, G., η -conversions of IPC implemented in atomic F, *Logic Journal of the IGPL* 25(2):115–130, 2017.
- [12] GABBAY, D., *Semantical Investigations in Heyting’s Intuitionistic Logic*, Springer Science + Business, 1981.
- [13] GAMBINO, N., and J. KOCK, Polynomial functors and polynomial monads, *Mathematical Proceedings of the Cambridge Philosophical Society* 154(1):153–192, 2013.
- [14] GIRARD, J.-Y., A. SCEDROV, and P. J. SCOTT, Normal forms and cut-free proofs as natural transformations, in Y. Moschovakis, (ed.), *Logic from Computer Science, vol. 21 of Mathematical Sciences Research Institute Publications*, Springer-Verlag, 1992, pp. 217–241.

- [15] GIRARD J.-Y., Y. LAFONT, and P. TAYLOR, *Proofs and Types*, Cambridge University Press, Cambridge, 1989.
- [16] LEIVANT, D., Finitely stratified polymorphism, *Information and Computation* 93(1):93–113, 1991.
- [17] LINDLEY, S., Extensional rewriting with sums, in S. Ronchi Della Rocca, (ed.), *Typed Lambda Calculi and Applications, TLCA 2007*, vol. 4583 of *Lecture Notes in Computer Science*, Springer Verlag, 2007, pp. 255–271.
- [18] OLKHOVIKOV, G., and P. SCHROEDER-HEISTER, On flattening elimination rules, *The Review of Symbolic Logic* 7:60–72, 2014.
- [19] PISTONE, P., *Proof nets and the instantiation overflow property*, available at [arXiv:1803.09297](https://arxiv.org/abs/1803.09297), 2018.
- [20] PISTONE, P., and L. TRANCHINI, The Yoneda Reduction of Polymorphic Types, in Ch. Baier, and J. Goubault-Larrecq, (eds.), 29th *EACSL Annual Conference on Computer Science Logic (CSL 2021)*, vol.183 of *Leibniz International Proceedings in Informatics (LIPIcs)*, Dagstuhl, Germany, 2021, pp. 35:1–35:22
- [21] PISTONE, P., and L. TRANCHINI, What’s Decidable about (Atomic) Polymorphism?, in N. Kobayashi, (ed.), *Proceedings of Formal Structures of Computation and Deduction (FSCD 2021)*, vol. 195 of *Leibniz International Proceedings in Informatics (LIPIcs)*, Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2021, pp. 27:1–27:23.
- [22] PRAWITZ, D., *Natural deduction, a proof-theoretical study*, Almqvist & Wiskell, 1965.
- [23] PRAWITZ, D., Ideas and results in proof theory, in J.E. Fenstad, (ed.), *Proceedings of the 2nd Scandinavian Logic Symposium (Oslo)*, vol. 63 of *Studies in logic and foundations of mathematics*, North-Holland, 1971, pp. 235–307.
- [24] PRAWITZ, D., Proofs and the meaning and completeness of the logical constants, in J. Hintikka, I. Niiniluoto, and E. Saarinen, (eds.), *Essays on Mathematical and Philosophical Logic: Proceedings of the Fourth Scandinavian Logic Symposium and the First Soviet-Finnish Logic Conference, Jyväskylä, Finland, June 29-July 6, 1976*, Kluwer, Dordrecht, 1979, pp. 25–40.
- [25] REYNOLDS, J. C., Types, abstraction and parametric polymorphism, in R.E.A. Mason, (ed.), *Information Processing '83*, North-Holland, 1983, pp. 513–523.
- [26] SCHROEDER-HEISTER, P., A natural extension of natural deduction, *Journal of Symbolic Logic* 49(4):1284–1300, 1984.
- [27] SCHWICHTENBERG, H., and A. S. TROELSTRA, *Basic proof theory*, Cambridge University Press, 2000.
- [28] SCHERER, G., Deciding equivalence with sums and the empty type, in *POPL 2017: Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages*, 2017, pp. 374–386.
- [29] SEELY, R.A.G., Weak adjointness in proof theory, in *Proceedings of the Durham Conference on Applications of Sheaves*, vol. 753 of *Springer Lecture Notes in Mathematics*, Springer, Berlin, 1979, pp. 697–701.
- [30] SOBOLEV, S.K., The intuitionistic propositional calculus with quantifiers, *Mathematical Notes* 22(1):528–532, 1977.

- [31] TRANCHINI, L., Proof-theoretic harmony: towards an intensional account. *Synthese* 198:1145–1176, 2021.
- [32] TRANCHINI, L., Stabilizing quantum disjunction, *Journal of Philosophical Logic* 47(6):1029–1047, 2018.
- [33] TRANCHINI, L., P. PISTONE, and M. PETROLO, The naturality of natural deduction, *Studia Logica* 107(1):195–231, 2019.
- [34] WADLER, P., Theorems for free!, in *Proceedings of the Fourth International Conference on Functional Programming Languages and Computer Architecture, FPCA '89*, ACM, New York, 1989, pp. 347–359.

P. PISTONE
Dipartimento di Informatica-Scienza e Ingegneria
Università di Bologna
Bologna
Italy

L. TRANCHINI
Wilhelm-Schickard-Institut
Universität Tübingen
Tübingen
Germany
luca.tranchini@gmail.com

M. PETROLO
Centro de Ciências Naturais e Humanas
Universidade Federal do ABC
São Bernardo do Campo
Brazil