Permissive nominal terms and their unification

Gilles Dowek[a], Murdoch J. Gabbay[b], Dominic P. Mulligan[c]

[a]http://www.lix.polytechnique.fr/~dowek
[b]http://www.gabbay.org.uk
[c]http://www.macs.hw.ac.uk/~dpm8

Abstract

We introduce *permissive nominal terms*, and their unification.

Nominal terms are one way to extend first-order terms with binding. However, they lack some useful properties of first- and higher-order terms: Terms must be reasoned about in a context of 'freshness assumptions'; it is not always possible to 'choose a fresh variable symbol' for a nominal term; and it is not always possible to 'α-convert a bound variable symbol'.

Permissive nominal terms closely resemble nominal terms, but they recover these useful 'always fresh' and 'always alpha-rename' properties, familiar from first- and higher-order syntax. In the permissive world, freshness contexts are elided, and the notion of unifier is based on substitution alone, rather than on nominal terms' notion of substitution plus freshness conditions.

We prove that expressivity is not lost moving to the permissive case. We provide a translation from nominal terms into permissive nominal terms and we prove that a nominal unification problem is solvable if and only if its translation into permissive nominal terms is.

Finally, we investigate the precise relation between nominal unification and Miller's higher-order pattern unification. We translate nominal terms into the λ-calculus and show that the translation may also be applied to unification problems; the result is pattern unification. This cements an existing intuition that higher-order patterns are what is needed to unify encodings of nominal terms. This builds on a translation by Levy and Villaret, and refines it; both translations are parameterised by sets of atoms, but we identify a smaller parameter set and prove that it is as small as possible. We also translate solutions of nominal unification problems to solutions of higher-order pattern unification problems. We exhibit a general class of higher-order pattern solutions and show that every pattern solution in that class is the translation of a nominal unification solution up to a permutative renaming.

Key words: Nominal unification, higher-order pattern unification, nominal techniques

February 25, 2009

Contents

1 Introduction **3**

2 Permissive nominal terms **4**

3 Substitutions **10**

4 Relation to nominal terms **12**

5 Support inclusion problems **17**
5.1 Simplification reduction and normal forms 17
5.2 Building solutions 19

6 Permissive nominal unification problems **21**
6.1 Problems, solutions, the unification algorithm 21
6.2 Simplification rewrites calculate principal solutions 24

7 The λ-calculus **27**

8 Translating nominal terms into the λ-calculus **32**
8.1 The translation $[\![\text{-}]\!]^D$, and its soundness 32
8.2 Capturable atoms; injectivity and minimality 34

9 Translating substitutions; relating solutions of nominal and pattern unification problems **37**
9.1 Translating substitutions 37
9.2 Reducing permissive nominal unification to pattern unification; soundness, weak completeness 40
9.3 Strong Completeness 42

10 Conclusions **45**

1. Introduction

Many formal languages feature *variable binding*: examples include quantification, λ-abstraction, sets comprehension $\{x \mid \phi(x)\}$, and process-calculi name-hiding. Binding is ubiquitous, because variables are there to be bound or substituted.

In contrast, variables cannot be bound in first-order terms: In first-order logic, variables are bound in propositions by quantifiers and not at all in terms; first-order rewriting does not allow binding as it is based on first-order terms; and, many programming languages and proof systems allow datatypes of terms, but only of first-order terms. This motivates logics where variables can be bound by any function or predicate symbol [8], extensions of rewriting on terms with binders [18, 20, 9], and programming languages and proof systems allowing datatypes with binders [25, 21, 26, 17, 24] and more generally, definitions of a notion of term where variables may be bound: *nominal terms* [27].

Introducing binding opens new possibilities: when a variable occurs in the scope of a binder, for instance in the term $f([a]X)$ we may decide that the substitution of some term for X may capture the bound variable a or not. In the original formulation of nominal terms, we could decide to exclude this capture by imposing a condition $a\#X$, forbidding a to appear in X. This example explains some of the features of nominal terms: two levels of variable (*atoms*, such as a, and *unknowns*, such as X), *freshness* conditions, such as $a\#X$, and atom *permutations*.

Nominal terms preserved much of the flavour of first-order terms, while extending them, so that we could represent informal statements like "If $y \notin fv(t)$ then $\lambda x.t$ is α-equivalent with $\lambda y.[y/x]t$" and "How can we choose t and u to make $\lambda x.\lambda y.(y\ t)$ equal to $\lambda x.\lambda x.(x\ u)$?". For instance, the first statement above may be rendered as a nominal term as the equality judgement $b\#X \vdash [a]X = [b](b\ a) \cdot X$ where a and b denote atoms, which represent the 'x' and 'y'; X denotes an unknown, it represents the 't'; $b\#X$ is a freshness side-condition, it represents the '$y \notin fv(t)$'; $(b\ a)$ is a *permutation* meaning 'map a to b and b to a', it represents the '$[y/x]$' (we assumed $y \notin fv(t)$, so this is possible). Yet original nominal terms possess some less attractive properties too:

- Freshness contexts are not fixed so we must often prove properties of *terms-in-freshness context*. This is harder than reasoning just about terms.
- We expect that we can *always* pick a fresh variable symbol and α-rename a bound variable. Not so in nominal terms; for X in the empty freshness context, there is — by definition of the empty freshness context — no a such that we know $a\#X$; further, we cannot α-rename abstracted a to a 'fresh b' to obtain $[b](b\ a)\cdot X$, because there is no fresh b (this is useful e.g. for moving syntax under a binder).

'Freshness contexts' sound like 'typing contexts' for the λ-calculus, but freshness contexts' effects are more complex and harder to control. Extending a typing context may make more terms typable, but will not typically make more terms *equal*; α-equivalence is independent of typing, but it depends on freshness. This then complicates, for example, normal forms and theories of reduction (lack of α-convertibility may block a reduction; adding a fresh atom may change the normal form of a term) as is explicit in [13], and implicit e.g. in [12, 11].

In this paper, we propose an alternative way to handle these conditions by associating a freshness context once and for all to each unknown. This leads to a new definition of nominal terms: *permissive nominal terms*. An unknown takes the form X^S where S is a single fixed permission sort (Definition 3); thus, we can reason about terms, rather than about terms-in-freshness-context. In a further departure from the usual nominal style, permission sorts are sets of atoms that are both infinite and co-infinite (see Definition 2).

Thus, we can always choose a fresh atom for a term, always α-convert, and α-equivalence is inherent rather than depending on a freshness context (Definition 11, Theorem 13, and Corollary 14).

Permissive nominal terms allow to simplify several results and algorithms of the theory of nominal terms, in particular their unification. The solutions of a problem, such as $(a\ b) \cdot X^S = X^S$, is the set of all substitutions mapping the variable X to a term containing no occurrences of a and b. The unifier of such a problem is simply the substitution $X^S := Y^T$ where $T = S \setminus \{a, b\}$.

Thus, after laying down definitions and basic properties of permissive nominal terms (Sections 2 and 3) and studying their relation to original nominal terms (Section 4), we focus on the unification of permissive nominal terms (Sections 5 and 6).

The rest of the paper considers translations of permissive nominal terms. It is known that languages containing binders can be encoded as datatypes in some programming languages and proof systems using λ-binding and *higher-order abstract syntax* [25]. We investigate this idea in the general form of a translation of permissive nominal terms to λ-calculus (Section 8). Finally, building up on a recent result of Levy and Villaret, we show (Section 9) that this translation can be applied to unification problems and yields pattern unification problems, as identified by Miller [22, 21], crystallizing the intuition that pattern unification is exactly what is needed to unify encodings of nominal terms.

Levy and Villaret's main result is that a nominal unification problem is solvable if and only if its translation into higher-order patterns, is solvable [19, Corollary 1 and Theorem 2]. We take this further as follows:

- We investigate how the solutions of a problem, and the solutions of its translation, correspond. We translate terms and substitutions from the 'nominal' to the 'higher-order pattern' world (Definitions 117 and 129) and show that the translation of a solution is a solution of the translation (Theorem 141).
- We exhibit a general class of higher-order pattern solutions (Definition 149), a notion of permutative renaming of pattern substitutions (Lemma 147), and show that every pattern substitution in that class is the translation of a nominal unification solution, up to permutative renaming (Theorem 155).
- We refine Levy and Villaret's translation of terms. Their translation is parameterised by a vector of atoms; Levy and Villaret consider a vector containing all the atoms of the problem [19, Definition 2], whereas we identify a smaller vector containing only the *capturable atoms* (Definition 121). We prove that this is minimal, in the sense that if we use any smaller vector, then injectivity is lost (Theorem 128).

This technical report is formed from two conference papers [6, 5], and a journal paper [7]; the technical report includes all material, with full proofs.

2. Permissive nominal terms

Definition 1. Fix a countably infinite set $\mathbb{A}$ of **atoms**. $a, b, c, \ldots$ will range over *distinct* atoms (we call this the **permutative convention**). Fix a set of **term-formers**. $\mathsf{f}, \mathsf{g}, \mathsf{h}$ will range over distinct term-formers.

Definition 2. Call $S \subseteq \mathbb{A}$ co-infinite when $\mathbb{A} \setminus S$ is infinite. Fix an infinite, co-infinite set $\mathit{comb} \subseteq \mathbb{A}$.[1] A **permissions set** has the form $(\mathit{comb} \cup A_1) \setminus A_2$ for finite sets $A_1 \subseteq \mathbb{A}$

[1] The second author helped develop nominal sets [14], which famously disallow sets like *comb* (*comb*

and $A_2 \subseteq \mathbb{A}$.

S, S', T will range over permissions sets, whereas *Permit* is the set of all permissions sets.

$S, T \in \mathit{Permit}$ implies $S \cup T, S \cap T \in \mathit{Permit}$, and S and $\mathbb{A} \setminus S$ are infinite.

Definition 3. For each permission sort S fix a disjoint countably infinite set of **unknowns** of sort S. X^S, Y^S, Z^S, will range over distinct unknowns of sort S. If $S \neq S'$ then there is no particular connection between X^S and $X^{S'}$. $\mathcal{V}$ will range over finite sets of unknowns (we use this from Section 5 onwards).

Definition 4. Define the **domain** of a function from atoms to atoms by:

$$\mathit{dom}(f) = \{a \mid f(a) \neq a\}$$

Definition 5. A **permutation** is a bijection on atoms such that $\mathit{dom}(\pi)$ is finite. π and π' will range over permutations (not necessarily distinct).

Write id for the **identity** permutation such that $\mathit{id}(a) = a$ always. Write $(a\ b) \cdot r$ for the **swapping** permutation that swaps a and b in the term r.

Definition 6. Define **(permissive nominal) terms** by:

$$r, s, t, \ldots ::= a \mid \mathsf{f}(r, \ldots, r) \mid [a]r \mid \pi \cdot X^S$$

We write $\equiv$ for syntactic identity; $r \equiv s$ when r and s denote identical terms.

Atoms represent *variable symbols*; term-formers *functions*; unknowns *meta-variables*; abstraction $[a]r$ *binding*; and $\pi \cdot X^S$ a meta-variable with a suspended substitution, like '$t[y/x]$'. For example, suppose term-formers app and lam:

– $\mathsf{app}(a, b)$ can represent 'xy' (x applied to y).
– $\mathsf{app}(\mathsf{lam}([a]a), b)$ can represent '$(\lambda x.x)y$' (identity applied to y).
– $\mathsf{lam}([a]X^{\mathit{comb}})$ can represent '$\lambda x.t$' if $a \in \mathit{comb}$, and '$\lambda x.t$, where $x \notin \mathit{fv}(t)$' if $a \notin \mathit{comb}$.

See Section 4 for comparison with 'ordinary' nominal terms [27].

Definition 7. Define a **permutation action** by:

$$\pi{\cdot}a \equiv \pi(a) \quad \pi{\cdot}(\mathsf{f}(r_1, \ldots)) \equiv \mathsf{f}(\pi{\cdot}r_1, \ldots) \quad \pi{\cdot}[a]r \equiv [\pi(a)](\pi{\cdot}r) \quad \pi{\cdot}(\pi'{\cdot}X^S) \equiv (\pi \circ \pi'){\cdot}X^S$$

Definition 8. If $S \subseteq \mathbb{A}$, define the **pointwise** action by: $\pi \cdot S = \{\pi(a) \mid a \in S\}$

Definition 9. Define **free atoms** $\mathit{fa}(r)$ by:

$$\mathit{fa}(a) = \{a\} \quad \mathit{fa}(\mathsf{f}(r_1, \ldots, r_n)) = \bigcup_{1 \leq i \leq n} \mathit{fa}(r_i) \quad \mathit{fa}([a]r) = \mathit{fa}(r) \backslash \{a\} \quad \mathit{fa}(\pi{\cdot}X^S) = \pi{\cdot}S$$

Definition 10. Define $\mathit{fV}(r)$ by:

$$\mathit{fV}(a) = \varnothing \quad \mathit{fV}(\mathsf{f}(r_1, \ldots, r_n)) = \bigcup_{1 \leq i \leq n} \mathit{fV}(r_i) \quad \mathit{fV}([a]r) = \mathit{fV}(r) \quad \mathit{fV}(\pi{\cdot}X^S) = \{X^S\}$$

lacks *finite support*) [14]. Here, we are working at the meta-level, where we can talk about any subset or function that we wish.

Definition 11. Define **α-equivalence** $=_\alpha$ inductively by:

$$\frac{}{a =_\alpha a}\,(=_\alpha\mathbf{aa}) \qquad \frac{r_1 =_\alpha s_1 \quad \cdots \quad r_n =_\alpha s_n}{\mathsf{f}(r_1,\ldots,r_n) =_\alpha \mathsf{f}(s_1,\ldots,s_n)}\,(=_\alpha\mathsf{f}) \qquad \frac{r =_\alpha s}{[a]r =_\alpha [a]s}\,(=_\alpha[\mathbf{a}])$$

$$\frac{(b\ a)\cdot r =_\alpha s \qquad (b \notin fa(r))}{[a]r =_\alpha [b]s}\,(=_\alpha[\mathbf{b}]) \qquad \frac{(\pi|_S = \pi'|_S)}{\pi \cdot X^S =_\alpha \pi' \cdot X^S}\,(=_\alpha\mathbf{X})$$

Here, $\pi|_S$ denotes the partial function 'π restricted to S'; similarly for π'.

Remark 12. $r =_\alpha s$ is either true or false. Compare with the corresponding notion for nominal terms, which is subject to a *freshness context* (Definition 37).

Theorem 13 and Corollary 14 are properties that 'ordinary syntax' has, that nominal terms do not have, and that permissive nominal terms recover; we can always choose a fresh variable, and we can always α-rename with it.

Theorem 13. *For any r, there exist infinitely many b such that $b \notin fa(r)$.*

Proof. By induction on r.

- The case a. There are infinitely many atoms not equal to a, as required.
- The case $\mathsf{f}(r_1,\ldots,r_n)$. Suppose $S, S', S'', \ldots$ are the distinct permissions sorts of $fV(r_i)$ for $1 \leq i \leq n$. By Definition 2, each permission sort differs finitely from, *comb*. Therefore, we have for every π and S, that $\pi \cdot S$ differs finitely from *comb*s, too. There exists infinitely many $b \notin \pi \cdot S \cup \pi' \cdot S'$, for every π, S and S'. The result follows from the inductive hypotheses.
- The case $[a]r$. By inductive hypothesis, infinitely many $b \notin fa(r)$ exist. There exists infinitely many $b \notin fa(r) \setminus \{a\}$. The result follows.
- The case $\pi \cdot X^S$. By Definition 9, $fa(\pi \cdot X^S) = \pi \cdot S$. Infinitely many $b \notin S$ exist, by the co-infinite nature of S, therefore infinitely many $b \notin \pi \cdot S$ exist. The result follows.

□

Corollary 14. *For any r and a there exists infinitely many fresh b (so $b \notin fa(r)$) such that for some s, $[a]r =_\alpha [b]s$.*

Proof. Immediate, by Theorem 13 and $(=_\alpha[\mathbf{b}])$. □

Our changes do not affect basic results about nominal terms [27]:

Lemma 15. 1. $id \cdot r \equiv r$
2. $\pi' \cdot (\pi \cdot r) \equiv (\pi' \circ \pi) \cdot r$

Proof. By induction on r.

- The cases a and $\mathsf{f}(r_1,\ldots,r_n)$. These are straightforward.
- The case $[a]r$. We have:

$$\begin{array}{lcll} id \cdot [a]r & \equiv & [a](id \cdot r) & \text{Definition 7} \\ & \equiv & [a]r & \text{Inductive hypothesis} \end{array}$$

$$\begin{array}{lcll} \pi \cdot (\pi' \cdot [a]r) & \equiv & [a](\pi \cdot (\pi' \cdot r)) & \text{Definition 7} \\ & \equiv & [a]((\pi \circ \pi') \cdot r) & \text{Indictive hypothesis} \\ & \equiv & (\pi \circ \pi') \cdot [a]r & \text{Definition 7} \end{array}$$

- The case $\pi'' \cdot X^S$. As $id \circ \pi'' = \pi''$, we have $id \cdot \pi'' \cdot X^S \equiv \pi'' \cdot X^S$. Further:

$$\begin{array}{rcll} \pi \cdot (\pi' \cdot (\pi'' \cdot X^S)) & \equiv & \pi \cdot ((\pi' \circ \pi'') \cdot X^S) & \text{Definition 7} \\ & \equiv & (\pi \circ (\pi' \circ \pi'')) \cdot X^S & \text{Definition 7} \\ & \equiv & ((\pi \circ \pi') \circ \pi'') \cdot X^S & \text{Fact} \\ & \equiv & (\pi \circ \pi') \cdot (\pi'' \cdot X^S) & \text{Definition 7} \end{array}$$

□

Lemma 16. $\pi \cdot fa(r) = fa(\pi \cdot r)$.

Proof. By induction on r.

- The case a and $\mathsf{f}(r_1, \ldots, r_n)$. These are easy.
- The case $[a]r$. We have:

$$\begin{array}{rcll} \pi \cdot fa([a]r) & = & \pi \cdot (fa(r) \setminus \{a\}) & \text{Definition 9} \\ & = & \pi \cdot fa(r) \setminus \{\pi(a)\} & \text{Fact} \\ & = & fa(\pi \cdot r) \setminus \{\pi(a)\} & \text{Inductive hypothesis} \\ & = & fa([\pi(a)](\pi \cdot r)) & \text{Definition 9} \\ & = & fa(\pi \cdot [a]r) & \text{Definition 7} \end{array}$$

- The case $\pi' \cdot X^S$. We have:

$$\begin{array}{rcll} \pi \cdot fa(\pi' \cdot X^S) & = & \pi \cdot (\pi' \cdot S) & \text{Definition 9} \\ & = & (\pi \circ \pi') \cdot S & \text{Fact} \\ & = & fa((\pi \circ \pi') \cdot X^S) & \text{Definition 9} \\ & = & fa(\pi \cdot (\pi' \cdot X^S)) & \text{Definition 7} \end{array}$$

□

Lemma 17. $fV(\pi \cdot r) = fV(r)$

Proof. By induction on r.

- The case a. Since $fV(a) = \varnothing = fV(\pi(a))$, the result follows.
- The case $\mathsf{f}(r_1, \ldots, r_n)$. We have:

$$\begin{array}{rcll} fV(\mathsf{f}(r_1, \ldots, r_n)) & = & \bigcup_{1 \le i \le n} fV(r_i) & \text{Definition 10} \\ & = & \bigcup_{1 \le i \le n} fV(\pi \cdot r_i) & \text{Inductive hypotheses} \\ & = & fV(\mathsf{f}(\pi \cdot r_1, \ldots, \pi \cdot r_n)) & \text{Definition 10} \\ & = & fV(\pi \cdot \mathsf{f}(r_1, \ldots, r_n)) & \text{Definition 7} \end{array}$$

  The result follows.
- The case $[a]r$. We have:

$$\begin{array}{rcll} fV([a]r) & = & fV(r) & \text{Definition 10} \\ & = & fV(\pi \cdot r) & \text{Inductive hypothesis} \\ & = & fV([\pi(a)](\pi \cdot r)) & \text{Definition 10} \\ & = & fV(\pi \cdot [a]r) & \text{Definition 7} \end{array}$$

  The result follows.
- The case $\pi' \cdot X^S$. By Definition 7, $\pi \cdot (\pi' \cdot X^S) \equiv (\pi \circ \pi') \cdot X^S$. Then $fV(\pi' \cdot X^S) = \{X^S\} = fV((\pi \circ \pi') \cdot X^S)$, and the result follows.

□

Lemma 18. *If $r =_\alpha s$ then $\pi \cdot r =_\alpha \pi \cdot s$.*

Proof. By induction on $r =_\alpha s$.

- The case $(=_\alpha \mathbf{a})$. Using $(=_\alpha \mathbf{a})$, $\pi(a) =_\alpha \pi(a)$ always.
- The case $(=_\alpha \mathsf{f})$. By hypothesis, $\pi \cdot r_i =_\alpha \pi \cdot s_i$ for $1 \leq i \leq n$. Extending with $(=_\alpha \mathsf{f})$, we have $\mathsf{f}(\pi \cdot r_1, \ldots, \pi \cdot r_n) =_\alpha \mathsf{f}(\pi \cdot s_1, \ldots, \pi \cdot s_n)$. This implies $\pi \cdot \mathsf{f}(r_1, \ldots, r_n) =_\alpha \pi \cdot \mathsf{f}(s_1, \ldots, s_n)$. The result follows.
- The case $(=_\alpha [\mathbf{a}])$. By hypothesis, $\pi \cdot r =_\alpha \pi \cdot s$. We use $(=_\alpha [\mathbf{a}])$ to conclude $[\pi(a)](\pi \cdot r) =_\alpha [\pi(a)](\pi \cdot s)$. The result follows.
- The case $(=_\alpha [\mathbf{b}])$. Suppose $(b\ a) \cdot r =_\alpha s$ with $b \notin fa(r)$. By Lemma 15 and inductive hypothesis, $(\pi \circ (b\ a)) \cdot r =_\alpha s$. By Lemma 16, $\pi(b) \notin fa(\pi \cdot r)$. Further, $(\pi \circ (b\ a)) = (\pi(b)\ \pi(a)) \circ \pi$. Using $(=_\alpha \mathbf{b})$, $[\pi(b)](\pi \cdot r) =_\alpha [\pi(a)](\pi \cdot s)$. The result follows.
- The case $(=_\alpha \mathbf{X})$. Suppose $\pi'|_S = \pi''|_S$ so that $\pi' \cdot X^S =_\alpha \pi'' \cdot X^S$. Then $(\pi \circ \pi')|_S = (\pi \circ \pi'')|_S$. By $(=_\alpha \mathbf{X})$, $(\pi \circ \pi') \cdot X^S =_\alpha (\pi \circ \pi'') \cdot X^S$. The result follows.

□

Lemma 19. *If $r =_\alpha s$ then $fa(r) = fa(s)$.*

Proof. By induction on $r =_\alpha s$.

- The cases $(=_\alpha \mathbf{a})$, $(=_\alpha \mathsf{f})$ and $(=_\alpha [\mathbf{a}])$. Easy.
- The case $(=_\alpha [\mathbf{b}])$. Suppose $[a]r =_\alpha [b]s$ by $(=_\alpha [\mathbf{b}])$, so $b \notin fa(r)$. We aim to show $fa([a]r) = fa([b]s)$, or $fa(r) \setminus \{a\} = fa(s) \setminus \{b\}$. As $b \notin fa(r)$, $fa(r) \setminus \{a\} = (b\ a) \cdot fa(r) \setminus \{b\}$. By Lemma 16, $(b\ a) \cdot fa(r) \setminus \{b\} = fa((b\ a) \cdot r) \setminus \{b\}$. By hypothesis, $fa((b\ a) \cdot r) = fa(s)$. The result follows.
- The case $(=_\alpha \mathbf{X})$. We have $fa(\pi \cdot X^S) = \pi \cdot S$ and $fa(\pi' \cdot X^S) = \pi' \cdot S$. By assumption, $\pi|_S = \pi'|_S$, therefore $\pi \cdot S = \pi' \cdot S$. The result follows.

□

Lemma 20. *If $\pi|_{fa(r)} = \pi'|_{fa(r)}$ then $\pi \cdot r =_\alpha \pi' \cdot r$.*

Proof. By induction on r.

- The case a. As $fa(a) = \{a\}$, the result follows by $(=_\alpha \mathbf{aa})$.
- The case $\mathsf{f}(r_1, \ldots, r_n)$. Suppose $\pi|_{fa(\mathsf{f}(r_1,\ldots,r_n))} = \pi'|_{fa(\mathsf{f}(r_1,\ldots,r_n))}$. By Definition 9, $\pi|_{fa(r_i)} = \pi'|_{fa(r_i)}$ for $1 \leq i \leq n$. By hypothesis, $\pi \cdot r_i =_\alpha \pi' \cdot r_i$ for $1 \leq i \leq n$. Using $(=_\alpha \mathsf{f})$, the result follows.
- The case $[a]r$. Suppose $\pi|_{fa(r)} = \pi'|_{fa(r)}$ so that $\pi \cdot r =_\alpha \pi' \cdot r$. Then $\pi|_{fa(r)\setminus\{a\}} = \pi'|_{fa(r)\setminus\{a\}}$. By Definition 9, $\pi|_{fa([a]r)} = \pi'|_{fa([a]r)}$. The result follows.
- The case $\pi'' \cdot X^S$. The result follows from $(=_\alpha \mathbf{X})$.

□

To prove Theorem 24, we introduce the following notion:

Definition 21. Define the **size** of a term r by:

$$size(a) = 0 \quad size(\mathsf{f}(r_1, \ldots, r_n)) = \sum_{1 \leq i \leq n} size(r_i) \quad size([a]r) = 1 + size(r) \quad size(\pi \cdot X^S) = 0$$

Lemma 22. $size(r) = size(\pi \cdot r)$

Proof. By induction on r.

- The case a. Since $size(\pi(a)) = 0 = size(a)$.

- The case $\mathsf{f}(r_1, \ldots, r_n)$. We have:

$$\begin{array}{rcll} \mathit{size}(\mathsf{f}(r_1, \ldots, r_n)) & = & \sum_{1 \leq i \leq n} \mathit{size}(r_i) & \text{Definition 21} \\ & = & \sum_{1 \leq i \leq n} \mathit{size}(\pi \cdot r_i) & \text{Inductive hypotheses} \\ & = & \mathit{size}(\mathsf{f}(\pi \cdot r_1, \ldots, \pi \cdot r_n)) & \text{Definition 21} \\ & = & \mathit{size}(\pi \cdot \mathsf{f}(r_1, \ldots, r_n)) & \text{Definition 7} \end{array}$$

- The case $[a]r$. We have:

$$\begin{array}{rcll} \mathit{size}([a]r) & = & 1 + \mathit{size}(r) & \text{Definition 21} \\ & = & 1 + \mathit{size}(\pi \cdot r) & \text{Inductive hypothesis} \\ & = & \mathit{size}([\pi(a)](\pi \cdot r)) & \text{Definition 21} \\ & = & \mathit{size}(\pi \cdot [a]r) & \text{Definition 7} \end{array}$$

- The case $\pi' \cdot X^S$. By Definition 7, $\pi \cdot (\pi' \cdot X^S) \equiv (\pi \circ \pi') \cdot X^S$. Then, $\mathit{size}(\pi' \cdot X^S) = 1 = \mathit{size}((\pi \circ \pi') \cdot X^S)$. The result follows.

□

Lemma 23. *For every term r, the set $\{\mathit{size}(s) \mid s \text{ is a subterm of } r\}$ is well-ordered.*

Proof. As $\{\mathit{size}(s) \mid s \text{ is a subterm of } r\}$ is a subset of the natural numbers. □

Theorem 24. *$=_\alpha$ is transitive, reflexive, and symmetric.*

Proof. We handle the three claims separately:

- The reflexivity case. We prove $r =_\alpha r$ by induction on r.
  - The case a. Straightforward, using $(=_\alpha \mathbf{a})$.
  - The cases $\mathsf{f}(r_1, \ldots, r_n)$ and $[a]r$. These are easy consequences of the inductive hypotheses.
  - The case $\pi \cdot X^S$. Note, $\pi|_S = \pi|_S$. Using $(=_\alpha \mathbf{X})$, $\pi \cdot X^S =_\alpha \pi \cdot X^S$.
- The symmetry case. We prove $s =_\alpha r$ by induction on $r =_\alpha s$.
  - The cases $(=_\alpha \mathbf{a})$ and $(=_\alpha \mathsf{f})$. Easy.
  - The case $(=_\alpha [\mathbf{a}])$. Suppose $r =_\alpha s$ so $s =_\alpha r$ by hypothesis. Using $(=_\alpha [\mathbf{a}])$, $[a]s =_\alpha [a]r$. The result follows.
  - The case $(=_\alpha [\mathbf{b}])$. Suppose $(b\ a) \cdot r =_\alpha s$ with $b \notin \mathit{fa}(r)$. By Lemma 16, $a \notin \mathit{fa}((a\ b) \cdot r)$. By Lemma 15, and as $\pi = \pi^{-1}$, $r =_\alpha (a\ b) \cdot s$. By Lemma 19, $a \notin \mathit{fa}(s)$, and by hypothesis, $(a\ b) \cdot s =_\alpha r$. Using $(=_\alpha [\mathbf{b}])$, $[b]s =_\alpha [a]r$. The result follows.
  - The case $(=_\alpha \mathbf{X})$. Since equality on partial functions is symmetric.
- The transitivity case. By Lemma 23, we may perform induction on the size of a term. We prove, given $r =_\alpha s$ and $s =_\alpha t$, that $r =_\alpha t$ by induction on $\mathit{size}(r)$.
  - The cases a and $\mathsf{f}(r_1, \ldots, r_n)$. Easy.
  - The case $[a]r$. We examine only the most complex case, where all abstracted variables are distinct. Suppose $(b\ a) \cdot r =_\alpha s$ and $(c\ b) \cdot s =_\alpha t$ with $b \notin \mathit{fa}(r)$ and $c \notin \mathit{fa}(s)$. By Lemma 18, $(c\ b) \cdot ((b\ a) \cdot r) =_\alpha (c\ b) \cdot s$. By Lemma 22, $(c\ b) \cdot ((b\ a) \cdot r) =_\alpha t$, equivalent to $(c\ a) \cdot r =_\alpha t$. By Lemma 19, $c \notin \mathit{fa}((b\ a) \cdot r)$. By Lemma 16, $c \notin (b\ a) \cdot \mathit{fa}(r)$. By Lemma 15, $c \notin \mathit{fa}(r)$. Using $(=_\alpha [\mathbf{b}])$, $[a]r =_\alpha [c]t$. The result follows.
  - The case $\pi \cdot X^S$. Since equality on partial functions is transitive.

□

3. Substitutions

Definition 25. A **substitution** θ is a function from unknowns to terms such that $fa(\theta(X^S)) \subseteq S$ always. θ, θ', θ_1, θ_2, will range over substitutions.

Write id for the **identity** substitution mapping X^S to $id \cdot X^S$ always. It will always be clear whether id means the identity substitution or permutation.

Suppose $fa(t) \subseteq S$. Write $[X^S{:=}t]$ for the substitution such that $[X^S{:=}t](X^S) \equiv t$ and $[X^S{:=}t](Y^T) \equiv id \cdot Y^T$ for all other Y^T.[2]

Definition 26. Define a **substitution action** on terms by:

$$a\theta{\equiv}a \quad \mathsf{f}(r_1,\ldots,r_n)\theta{\equiv}\mathsf{f}(r_1\theta,\ldots,r_n\theta) \quad ([a]r)\theta{\equiv}[a](r\theta) \quad (\pi{\cdot}X^S)\theta{\equiv}\pi{\cdot}\theta(X^S)$$

Theorem 27. $fa(r\theta) \subseteq fa(r)$.

Proof. By induction on r.

- The case a. Since $a\theta \equiv a$.
- The case $\mathsf{f}(r_1,\ldots,r_n)$. We have:

$$\begin{array}{lcll} fa(\mathsf{f}(r_1,\ldots,r_n)\theta) & \equiv & fa(\mathsf{f}(r_1\theta,\ldots,r_n\theta)) & \text{Definition 26} \\ & = & fa(r_1\theta) \cup \ldots \cup fa(r_n\theta) & \text{Definition 9} \\ & \subseteq & fa(r_1) \cup \ldots \cup fa(r_n) & \text{Inductive hypotheses} \\ & = & fa(r_1,\ldots,r_n) & \text{Definition 9} \end{array}$$

  The result follows.
- The case $[a]r$. We have:

$$\begin{array}{lcll} fa(([a]r)\theta) & \equiv & fa([a]r\theta) & \text{Definition 26} \\ & = & fa(r\theta) \setminus \{a\} & \text{Definition 9} \\ & \subseteq & fa(r) \setminus \{a\} & \text{Inductive hypothesis} \\ & = & fa([a]r) & \text{Definition 9} \end{array}$$

  The result follows.
- The case $\pi \cdot X^S$. By Definition 9, $fa(\pi \cdot X^S) = \pi \cdot S$. By Definition 25, $fa(\theta(X^S)) \subseteq S$. Using Lemma 16, $fa(\pi \cdot \theta(X^S)) \subseteq \pi \cdot S$. The result follows.

□

Lemma 28. $\pi \cdot (r\theta) \equiv (\pi \cdot r)\theta$.

Proof. By induction on r.

- The case a. We have:

$$\begin{array}{lcll} \pi \cdot (a\theta) & \equiv & \pi \cdot a & \text{Definition 26} \\ & \equiv & \pi(a) & \text{Definition 7} \\ & \equiv & \pi(a)\theta & \text{Definition 26} \\ & \equiv & (\pi \cdot a)\theta & \text{Definition 7} \end{array}$$

  The result follows.
- The case $\mathsf{f}(r_1,\ldots,r_n)$ and $[a]r$. These are straightforward.

[2] '$fa(\theta(X^S)) \subseteq S$' looks absent in nominal terms theory ([27, Definition 2.13], [9, Definition 4]), yet it is there: see the conditions '$\nabla' \vdash \theta(\nabla)$' in Lemma 2.14, and '$\nabla \vdash a\#\theta(t)$' in Definition 3.1 of [27]. More on this in Section 4.

- The case $\pi \cdot X^S$.

$$\begin{array}{rcll} \pi \cdot ((\pi' \cdot X^S)\theta) & \equiv & \pi \cdot (\pi' \cdot \theta(X^S)) & \text{Definition 26} \\ & \equiv & (\pi \circ \pi') \cdot \theta(X^S) & \text{Lemma 15} \\ & \equiv & ((\pi \circ \pi') \cdot X^S)\theta & \text{Definition 26} \\ & \equiv & (\pi \cdot (\pi' \cdot X^S))\theta & \text{Lemma 15} \end{array}$$

  The result follows.

□

Lemma 29. *$fV(r[X^S{:=}s]) \subseteq fV(r) \cup fV(s)$.*

Proof. By induction on r.

- The case a. Since $a[X^S{:=}s] \equiv a$.
- The case $\mathsf{f}(r_1, \ldots, r_n)$. We have:

$$\begin{array}{rcll} fV(\mathsf{f}(r_1, \ldots, r_n)[X^S{:=}s]) & = & fV(\mathsf{f}(r_1[X^S{:=}s], \ldots, r_n[X^S{:=}s])) & \text{Definition 25} \\ & = & \bigcup_{1 \leq i \leq n} fV(r_i[X^S{:=}s]) & \text{Definition 10} \\ & \subseteq & \bigcup_{1 \leq i \leq n} fV(r_i) \cup fV(s) & \text{Inductive hypothesis} \\ & = & fV(\mathsf{f}(r_1, \ldots, r_n)) \cup fV(s) & \text{Definition 10} \end{array}$$

  The result follows.
- The case $[a]r$. We have:

$$\begin{array}{rcll} fV(([a]r)[X^S{:=}s]) & = & fV([a](r[X^S{:=}s])) & \text{Definition 25} \\ & = & fV(r[X^S{:=}s]) & \text{Definition 10} \\ & \subseteq & fV(r) \cup fV(s) & \text{Inductive hypothesis} \\ & = & fV([a]r) \cup fV(s) & \text{Definition 10} \end{array}$$

  The result follows.
- The case $\pi \cdot Y^T$. By Definition 26, $(\pi \cdot Y^T)[X^S{:=}s] \equiv \pi \cdot Y^T$. The result follows.
- The case $\pi \cdot X^S$. By Definition 26, $(\pi \cdot X^S)[X^S{:=}s] \equiv \pi \cdot s$. By Lemma 17, $fV(s) = fV(\pi \cdot s)$. The result follows.

□

Theorem 30. *If $X^S\theta_1 =_\alpha X^S\theta_2$ for all $X^S \in fV(r)$, then $r\theta_1 =_\alpha r\theta_2$.*

Proof. By induction on r.

- The case a. As $fV(a) = \varnothing$.
- The case $\mathsf{f}(r_1, \ldots, r_n)$. Suppose for every $X^S \in fV(r_i)$ for $1 \leq i \leq n$ we have $X^S\theta_1 =_\alpha X^S\theta_2$ and $r_i\theta_1 =_\alpha r_i\theta_2$ by hypothesis. Using $(=_\alpha \mathsf{f})$, $\mathsf{f}(r_1\theta_1, \ldots, r_n\theta_1) =_\alpha \mathsf{f}(r_1\theta_2, \ldots, r_n\theta_2)$. By Definition 26, $\mathsf{f}(r_1, \ldots, r_n)\theta_1 =_\alpha \mathsf{f}(r_1, \ldots, r_n)\theta_2$. The result follows.
- The case $[a]r$. Suppose, for every $X^S \in fa(r)$, we have $X^S\theta_1 =_\alpha X^S\theta_2$. By hypothesis, $r\theta_1 =_\alpha r\theta_2$. Using $(=_\alpha[\mathbf{a}])$, $[a](r\theta_1) =_\alpha [a](r\theta_2)$. The result follows.
- The case $\pi \cdot X^S$. By assumption, $X^S\theta_1 =_\alpha X^S\theta_2$. By Lemma 18, $\pi \cdot (X^S\theta_1) =_\alpha \pi \cdot (X^S\theta_2)$. By Lemma 28, $(\pi \cdot X^S)\theta_1 =_\alpha (\pi \cdot X^S)\theta_2$. The result follows.

□

Lemma 31. *If $r =_\alpha s$ then $r\theta =_\alpha s\theta$.*

Proof. By induction on the derivation of $r =_\alpha s$.

- The case $(=_\alpha \mathbf{aa})$. As $a\theta \equiv a$.

- The cases $(=_\alpha \mathsf{f})$ and $(=_\alpha [\mathbf{a}])$. These are immediate consequences of the inductive hypotheses.
- The case $(=_\alpha [\mathbf{b}])$. Suppose $(b\ a)\cdot r =_\alpha s$ with $b \notin fa(r)$. Then $((b\ a)\cdot r)\theta =_\alpha s\theta$ by assumption. By Lemma 28, $(b\ a)\cdot r\theta =_\alpha s\theta$. By Theorem 27, $b \notin fa(r\theta)$. Using $(=_\alpha [\mathbf{b}])$, $[a](r\theta) =_\alpha [b](s\theta)$. By Definition 25, $[a](r\theta) \equiv ([a]r)\theta$. The result follows.
- The case $(=_\alpha \mathbf{X})$. Suppose $\pi \cdot X^S =_\alpha \pi' \cdot X^S$ using $(=_\alpha \mathbf{X})$. Then $\pi|_S = \pi'|_S$ and as $fa(\theta(X^S)) \subseteq S$ by assumption. The result follows.

□

Definition 32. Define **composition** $\theta_1 \circ \theta_2$ by $(\theta_1 \circ \theta_2)(X^S) \equiv (\theta_1(X^S))\theta_2$.

Theorem 33. $(r\theta)\theta' \equiv r(\theta \circ \theta')$.

Proof. By induction on r.

- The cases a, $\mathsf{f}(r_1, \ldots, r_n)$ and $[a]r$. Straightforward.
- The case $[a]r$. We have:

$$\begin{array}{lcll} (([a]r)\theta)\theta' & \equiv & [a]((r\theta)\theta') & \text{Definition 26} \\ & \equiv & [a](r(\theta\circ\theta')) & \text{Inductive hypothesis} \\ & \equiv & ([a]r)(\theta\circ\theta') & \text{Definition 26} \end{array}$$

  The result follows.
- The case $\pi \cdot X^S$

$$\begin{array}{lcll} (\pi \cdot X^S)(\theta\circ\theta') & \equiv & \pi \cdot (\theta\circ\theta')(X^S) & \text{Definition 26} \\ & \equiv & \pi \cdot (\theta(X^S)\theta') & \text{Definition 32} \\ & \equiv & (\pi \cdot \theta(X^S))\theta' & \text{Lemma 28} \\ & \equiv & ((\pi \cdot X^S)\theta)\theta' & \text{Lemma 28} \end{array}$$

  The result follows.

□

4. Relation to nominal terms

Nominal terms are described fully elsewhere [27]. We inject 'nominal' into 'permissive'. Main results are Theorems 41, 42, and 49.

Fix a countably infinite set of **nominal atoms**, $\dot{\mathbb{A}}$. $\dot{a}, \dot{b}, \dot{c}, \ldots$ will range over distinct nominal atoms. Fix a bijection ι between $\dot{\mathbb{A}}$ and *comb* (Definition 2). Fix a countably infinite set of **nominal unknowns**. $\dot{X}, \dot{Y}, \dot{Z}, \ldots$ will range over distinct nominal unknowns. A **nominal permutation** is a bijection $\dot{\pi}$ on $\dot{\mathbb{A}}$ such that $dom(\dot{\pi})$ is finite. $\dot{\pi}, \dot{\pi}', \dot{\pi}'', \ldots$ will range over permutations.

Write $\dot{\pi}^{-1}$ for the inverse of $\dot{\pi}$, $\dot{id}$ for the identity permutation, and $\dot{\pi}\circ\dot{\pi}'$ for function composition, as is standard. For example, $(\dot{\pi}\circ\dot{\pi}')(\dot{a}) = \dot{\pi}(\dot{\pi}'(\dot{a}))$

Definition 34. Define **nominal terms** with the following grammar:

$$\dot{r}, \dot{s}, \dot{t} ::= \dot{a} \mid \dot{\pi} \cdot \dot{X} \mid [\dot{a}]\dot{r} \mid \mathsf{f}(\dot{r}_1, \ldots, \dot{r}_n)$$

Definition 35. Define a **permutation** action on nominal terms with the following rules:

$$\dot{\pi} \cdot \dot{a} \equiv \dot{\pi}(\dot{a}) \quad \dot{\pi} \cdot \mathsf{f}(\dot{r}_1, \ldots, r_n) \equiv \mathsf{f}(\dot{\pi} \cdot \dot{r}_1, \ldots, \dot{\pi} \cdot r_n) \quad \dot{\pi} \cdot [\dot{a}]\dot{r} \equiv [\dot{\pi}(\dot{a})](\dot{\pi} \cdot \dot{r})$$
$$\dot{\pi} \cdot (\dot{\pi}' \cdot \dot{X}) \equiv (\dot{\pi}\circ\dot{\pi}') \cdot \dot{X}$$

$$\frac{}{\Delta \vdash \dot{a}\#\dot{b}}\,(\#\dot{\mathbf{b}}) \qquad \frac{\Delta \vdash \dot{a}\#\dot{r}_i \quad (1 \le i \le n)}{\Delta \vdash \dot{a}\#\mathsf{f}(\dot{r}_1, \ldots, \dot{r}_n)}\,(\#\mathsf{f}) \qquad \frac{}{\Delta \vdash \dot{a}\#[\dot{a}]\dot{r}}\,(\#[\dot{\mathbf{a}}])$$

$$\frac{\Delta \vdash \dot{a}\#\dot{r}}{\Delta \vdash \dot{a}\#[\dot{b}]\dot{r}}\,(\#[\dot{\mathbf{b}}]) \qquad \frac{(\dot{\pi}^{-1}(\dot{a})\#\dot{X} \in \Delta)}{\Delta \vdash \dot{a}\#\dot{X}}\,(\#\dot{\mathbf{X}})$$

Figure 1: Derivable freshness on nominal terms

$$\frac{}{\Delta \vdash \dot{a} = \dot{a}}\,(=\dot{\mathbf{a}}) \qquad \frac{\Delta \vdash \dot{r}_i = \dot{s}_i \quad (1 \le i \le n)}{\Delta \vdash \mathsf{f}(\dot{r}_1, \ldots, \dot{r}_n) = \mathsf{f}(\dot{s}_1, \ldots, \dot{s}_n)}\,(=\mathsf{f}) \qquad \frac{\Delta \vdash \dot{r} = \dot{s}}{\Delta \vdash [\dot{a}]\dot{r} = [\dot{a}]\dot{s}}\,(=[\dot{\mathbf{a}}])$$

$$\frac{\Delta \vdash (\dot{b}\ \dot{a}) \cdot \dot{r} = \dot{s} \quad \Delta \vdash \dot{b}\#\dot{r}}{\Delta \vdash [\dot{a}]\dot{r} = [\dot{b}]\dot{s}}\,(=[\dot{\mathbf{b}}]) \qquad \frac{(\dot{a}\#\dot{X} \in \Delta \text{ for every } \dot{\pi}(\dot{a}) \neq \dot{\pi}'(\dot{a}))}{\Delta \vdash \dot{\pi} \cdot \dot{X} = \dot{\pi}' \cdot \dot{X}}\,(=\dot{\mathbf{X}})$$

Figure 2: Derivable equality on nominal terms

Write $\equiv$ for syntactic identity. f ranges over term-formers (Definition 1).

Definition 36. A **freshness** is a pair $\dot{a}\#\dot{r}$. A **freshness context** is a finite set of freshnesses of the form $\dot{a}\#\dot{X}$. Define **derivable freshness** on nominal terms by the rules in Figure 1.

Definition 37. A **equality** is a pair $\dot{r} = \dot{s}$. Define **derivable equality** on nominal terms by the rules in Figure 2.

Definition 38. Define a mapping $[\![\dot{\pi}]\!]$ from nominal permutations to permissive nominal permutations by $[\![\dot{\pi}]\!]\iota(\dot{a}) = \iota(\pi(\dot{a}))$ and $[\![\dot{\pi}]\!](c) = c$ for all other c. Define an **interpretation** $[\![\dot{r}]\!]_\Delta$ by:

$$[\![\dot{a}]\!]_\Delta \equiv \iota(\dot{a}) \qquad [\![\mathsf{f}(\dot{r}_1, \ldots, \dot{r}_n)]\!]_\Delta \equiv \mathsf{f}([\![\dot{r}_1]\!]_\Delta, \ldots, [\![\dot{r}_n]\!]_\Delta) \qquad [\![[\dot{a}]\dot{r}]\!]_\Delta \equiv [\iota(\dot{a})][\![\dot{r}]\!]_\Delta$$
$$[\![\dot{\pi} \cdot \dot{X}]\!]_\Delta \equiv [\![\dot{\pi}]\!] \cdot X^S \quad \text{where } S = \mathit{comb} \setminus \{\iota(\dot{a}) \mid \dot{a}\#\dot{X} \in \Delta\}$$

Here, we make a fixed but arbitrary choice of X^S for each $\dot{X}$, injectively so that $[\![\dot{X}]\!]_\Delta$ and $[\![\dot{Y}]\!]_\Delta$ are always distinct.

$[\![\dot{r}]\!]_\Delta$ commutes with permutation and it preserves and reflects freshness:

Lemma 39. $[\![\dot{\pi}]\!] \cdot [\![\dot{r}]\!]_\Delta \equiv [\![\dot{\pi} \cdot \dot{r}]\!]_\Delta$

Proof. By induction on $\dot{r}$.

- The case $\dot{a}$. Suppose $\dot{\pi} \cdot \dot{a} = \dot{b}$ and $[\![\dot{\pi}]\!](\iota(\dot{a})) = \iota(\dot{b})$. Then $\iota(\dot{\pi} \cdot \dot{a}) = \iota(\dot{b}) = [\![\dot{\pi}]\!](\iota(\dot{a}))$ and $[\![\dot{\pi}]\!](\iota(\dot{a})) = [\![\dot{\pi}]\!] \cdot [\![\dot{a}]\!]_\Delta$. The result follows.
- The case $\dot{\pi}' \cdot \dot{X}$.

$$\begin{array}{lcll} [\![\dot{\pi} \cdot \dot{\pi}' \cdot \dot{X}]\!]_\Delta & \equiv & [\![(\dot{\pi} \circ \dot{\pi}') \cdot \dot{X}]\!]_\Delta & \text{Definition 35} \\ & \equiv & [\![\dot{\pi} \circ \dot{\pi}']\!] \cdot X^S & \text{Definition 38, } S = \mathit{comb} \setminus \{\iota(\dot{a}) \mid \dot{a}\#\dot{X} \in \Delta\} \\ & \equiv & [\![\dot{\pi}]\!] \cdot ([\![\dot{\pi}']\!] \cdot X^S) & \text{Definition 35} \\ & \equiv & [\![\dot{\pi}]\!] \cdot ([\![\dot{\pi}' \cdot \dot{X}]\!]_\Delta) & \text{Lemma 15, Definition 38} \end{array}$$

- The case $[\dot{a}]\dot{r}$. We have:

$$\begin{array}{lcll} [\![\dot{\pi}\cdot[\dot{a}]\dot{r}]\!]_\Delta & \equiv & [\![[\dot{\pi}(\dot{a})]\dot{\pi}\cdot\dot{r}]\!]_\Delta & \text{Definition 35} \\ & \equiv & [\iota(\dot{\pi}\cdot\dot{a})][\![\dot{\pi}\cdot\dot{r}]\!]_\Delta & \text{Definition 38} \\ & \equiv & [\iota(\dot{\pi}\cdot\dot{a})]([\![\dot{\pi}]\!]\cdot[\![\dot{r}]\!]_\Delta) & \text{Inductive hypothesis} \\ & \equiv & [[\![\dot{\pi}]\!]\cdot\iota(\dot{a})]([\![\dot{\pi}]\!]\cdot[\![\dot{r}]\!]_\Delta) & \text{Definition 38} \\ & \equiv & [\![\dot{\pi}]\!]\cdot[\![[\dot{a}]\dot{r}]\!]_\Delta & \text{Definition 7} \end{array}$$

  The result follows.
- The case $\mathsf{f}(\dot{r}_1,\ldots,\dot{r}_n)$. This is routine.

□

Lemma 40. *$\iota(\dot{a})\notin fa([\![\dot{r}]\!]_\Delta)$ if and only if $\Delta\vdash\dot{a}\#\dot{r}$.*

Proof. We handle the two implications separately.

- The case $\iota(\dot{a})\notin fa([\![\dot{r}]\!]_\Delta)$ implies $\Delta\vdash\dot{a}\#\dot{r}$. We proceed by induction on $\dot{r}$.
  - The cases $\dot{b}$ and $\mathsf{f}(\dot{r}_1,\ldots,\dot{r}_n)$. Straightforward.
  - The case $[\dot{a}]\dot{r}$. There are two cases to consider:
    - The case $[\dot{a}]\dot{r}$. Using $(\#[\dot{\mathbf{a}}])$, $\Delta\vdash\dot{a}\#[\dot{a}]\dot{r}$ always.
    - The case $[\dot{b}]\dot{r}$. Suppose $\iota(\dot{a})\notin fa([\![[\iota(\dot{b})]\dot{r}]\!]_\Delta)$ and $\iota(\dot{a})\notin fa([\![\dot{r}]\!]_\Delta)\setminus\{\iota(\dot{b})\}$. Then $\iota(\dot{a})\notin fa([\![\dot{r}]\!]_\Delta)$, therefore $\Delta\vdash\dot{a}\#\dot{r}$ by hypothesis. Using $(\#[\dot{\mathbf{b}}])$, $\Delta\vdash\dot{a}\#[\dot{b}]\dot{r}$. The result follows.
  - The case $\dot{\pi}\cdot\dot{X}$. Suppose $\iota(\dot{a})\notin fa([\![\dot{\pi}\cdot\dot{X}]\!]_\Delta)$. Then $\iota(\dot{a})\notin[\![\dot{\pi}]\!]\cdot S$, where $S=comb\setminus\{\iota(\dot{a})\mid\dot{a}\#\dot{X}\in\Delta\}$. But $[\![\dot{\pi}]\!]\cdot comb\setminus\{\iota(\dot{a})\mid\dot{a}\#\dot{X}\in\Delta\}$ equivalent to $[\![\dot{\pi}]\!]\cdot comb\setminus[\![\dot{\pi}]\!]\cdot\{\iota(\dot{a})\mid\dot{a}\#\dot{X}\in\Delta\}$. Then $[\![\dot{\pi}]\!]\cdot\{\iota(\dot{a})\mid\dot{a}\#\dot{X}\in\Delta\}=\{[\![\dot{\pi}]\!]\cdot\iota(\dot{a})\mid\dot{a}\#\dot{X}\in\Delta\}$. By Definition 38, and the fact permutations are bijective, $\{[\![\dot{\pi}]\!]\cdot\iota(\dot{a})\mid a\#X\in\Delta\}=\{\iota(\dot{\pi}^{\text{-}1}\cdot\dot{a})\mid\dot{\pi}^{\text{-}1}\cdot\dot{a}\#X\in\Delta\}$. Using $(\#\dot{\mathbf{X}})$, $\Delta\vdash\dot{a}\#\dot{X}$. The result follows.
- The case $\Delta\vdash\dot{a}\#\dot{r}$ implies $\iota(\dot{a})\notin fa([\![\dot{r}]\!]_\Delta)$. We proceed by induction on the derivation of $\Delta\vdash\dot{a}\#\dot{r}$.
  - The cases $(\#\dot{\mathbf{b}})$ and $(\#\mathsf{f})$. Routine.
  - The case $(\#[\dot{\mathbf{a}}])$. Suppose $\Delta\vdash\dot{a}\#[\dot{a}]\dot{r}$ using $(\#[\dot{\mathbf{a}}])$. Then $[\![[\dot{a}]\dot{r}]\!]_\Delta\equiv[\iota(\dot{a})][\![\dot{r}]\!]_\Delta$. Further, $\iota(\dot{a})\notin fa([\![\dot{r}]\!]_\Delta)\setminus\{\iota(\dot{a})\}$. The result follows.
  - The case $(\#[\dot{\mathbf{b}}])$. Suppose $\Delta\vdash\dot{a}\#\dot{r}$ and $\iota(\dot{a})\notin fa(\dot{r})$ by assumption. Using $(\#[\dot{\mathbf{b}}])$, $\Delta\vdash\dot{a}\#[\dot{b}]\dot{r}$. Then, $fa([\![[\dot{b}]\dot{r}]\!]_\Delta)=fa([\![\dot{r}]\!]_\Delta)\setminus\{\iota(\dot{b})\}$. The result follows.
  - The case $(\#\dot{\mathbf{X}})$. Suppose $\dot{\pi}^{\text{-}1}(\dot{a})\#\dot{X}\in\Delta$, and $\Delta\vdash\dot{a}\#\dot{\pi}\cdot\dot{X}$ using $(\#\dot{\mathbf{X}})$. Then $[\![\dot{\pi}\cdot\dot{X}]\!]_\Delta=[\![\dot{\pi}]\!]\cdot X^S$ where $S=comb\setminus\{\iota(\dot{a})\mid\dot{a}\#\dot{X}\in\Delta\}$. Further, $fa([\![\dot{\pi}]\!]\cdot X^S)=[\![\dot{\pi}]\!]\cdot S$. The result follows by Definition 38.

□

Theorem 41. *$[\![\dot{r}]\!]_\Delta=_\alpha[\![\dot{s}]\!]_\Delta$ implies $\Delta\vdash\dot{r}=\dot{s}$.*

Proof. By induction on the derivation of $[\![\dot{r}]\!]_\Delta=_\alpha[\![\dot{s}]\!]_\Delta$.

- The case $\dot{a}$. We have $[\![\dot{a}]\!]_\Delta=\iota(\dot{a})$. Using $(=_\alpha\mathbf{aa})$, $s\equiv\dot{a}$ and $[\![\dot{a}]\!]_\Delta=_\alpha[\![\dot{a}]\!]_\Delta$. Using $(=\dot{\mathbf{a}})$, $\Delta\vdash\dot{a}=\dot{a}$. The result follows.
- The case $\mathsf{f}(\dot{r}_1,\ldots,\dot{r}_n)$. Suppose $[\![\mathsf{f}(\dot{r}_1,\ldots,\dot{r}_n)]\!]_\Delta=_\alpha[\![\mathsf{f}(\dot{s}_1,\ldots,\dot{s}_n)]\!]_\Delta$. Then $[\![\dot{r}_i]\!]_\Delta=_\alpha[\![\dot{s}_i]\!]_\Delta$ for $1\leq i\leq n$. By hypothesis, $\Delta\vdash\dot{r}_i=\dot{s}_i$ for $1\leq i\leq n$. Using $(=_\alpha\mathsf{f})$, $f([\![\dot{r}_1]\!]_\Delta,\ldots,[\![\dot{r}_n]\!]_\Delta)=_\alpha f([\![\dot{s}_1]\!]_\Delta,\ldots,[\![\dot{s}_n]\!]_\Delta)$. Then $f([\![\dot{r}_1]\!]_\Delta,\ldots,[\![\dot{r}_n]\!]_\Delta)=[\![\mathsf{f}(\dot{r}_1,\ldots,\dot{r}_n)]\!]_\Delta$. The result follows.
- The case $(=_\alpha[\mathbf{a}])$. Suppose $[\![\dot{r}]\!]_\Delta=_\alpha[\![\dot{s}]\!]_\Delta$ and $\Delta\vdash\dot{r}=\dot{s}$. Using $(=_\alpha[\mathbf{a}])$, $[\iota(\dot{a})][\![\dot{r}]\!]_\Delta=_\alpha[\iota(\dot{a})][\![\dot{s}]\!]_\Delta$. Using $(=[\dot{\mathbf{a}}])$, $\Delta\vdash[\dot{a}]\dot{r}=[\dot{a}]\dot{s}$. Then $[\iota(\dot{a})][\![\dot{r}]\!]_\Delta=[\![[\dot{a}]\dot{r}]\!]_\Delta$. The result follows.

- The case ($=_\alpha[\mathbf{b}]$). Suppose $(\iota(\dot{b})\ \iota(\dot{a}))\cdot[\![\dot{r}]\!]_\Delta =_\alpha [\![\dot{s}]\!]_\Delta$ and $\iota(\dot{b}) \notin fa([\![\dot{r}]\!]_\Delta)$. By Lemmas 39 and 40, $[\![(\dot{b}\ \dot{a})\cdot\dot{r}]\!]_\Delta =_\alpha [\![\dot{s}]\!]_\Delta$ and $\Delta \vdash \dot{b}\#\dot{r}$. By hypothesis, $\Delta \vdash (\dot{b}\ \dot{a})\cdot\dot{r} = \dot{s}$. Using ($=[\dot{\mathbf{b}}]$), $[\dot{a}]\dot{r} = [\dot{b}]\dot{s}$. The result follows.
- The case ($=_\alpha\mathbf{X}$). Suppose $[\![\dot{\pi}]\!]|_S = [\![\dot{\pi}']\!]|_S$ where $S = comb\backslash\{\iota(\dot{a}) \mid \dot{a}\#\dot{X}\in\Delta\}$. ι is injective, so $a\#\dot{X} \in \Delta$ for all $\dot{a}$ such that $\dot{\pi}(\dot{a}) \neq \dot{\pi}'(\dot{a})$. The result follows.

□

Theorem 42. *If $\Delta \vdash \dot{r} = \dot{s}$ then $[\![\dot{r}]\!]_\Delta =_\alpha [\![\dot{s}]\!]_\Delta$.*

Proof. By induction on the derivation of $\Delta \vdash \dot{r} = \dot{s}$.

- The case ($=\dot{\mathbf{a}}$). Straightforward.
- The case ($=\mathsf{f}$). Routine, by the inductive hypotheses.
- The case ($=[\dot{\mathbf{a}}]$). Suppose $\Delta \vdash \dot{r} = \dot{s}$ and $[\![\dot{r}]\!]_\Delta =_\alpha [\![\dot{s}]\!]_\Delta$. Using ($=[\dot{\mathbf{a}}]$), $\Delta \vdash [\dot{a}]\dot{r} = [\dot{a}]\dot{s}$. Then, $[\![[\dot{a}]\dot{r}]\!]_\Delta = [\iota(\dot{a})][\![\dot{r}]\!]_\Delta$ and $[\![[\dot{a}]\dot{s}]\!]_\Delta = [\iota(\dot{a})][\![\dot{s}]\!]_\Delta$. Using ($=_\alpha[\mathbf{a}]$), $[\iota(\dot{a})][\![\dot{r}]\!]_\Delta =_\alpha [\iota(\dot{a})][\![\dot{s}]\!]_\Delta$ whenever $[\![\dot{r}]\!]_\Delta =_\alpha [\![\dot{s}]\!]_\Delta$. The result follows.
- The case ($=[\dot{\mathbf{b}}]$). Suppose $\Delta \vdash (\dot{b}\ \dot{a})\cdot\dot{r} = \dot{s}$ and $\Delta \vdash \dot{b}\#\dot{r}$. By hypothesis and Lemma 39, $(\dot{b}\ \dot{a})\cdot[\![\dot{r}]\!]_\Delta =_\alpha [\![\dot{s}]\!]_\Delta$. By Lemma 40, $\iota(\dot{b}) \notin fa([\![\dot{r}]\!]_\Delta)$. The result follows by ($=_\alpha[\mathbf{b}]$).
- The case ($=\dot{\mathbf{X}}$). Recall that $[\![\dot{\pi}\cdot\dot{X}]\!]_\Delta = [\![\dot{\pi}]\!]\cdot X^S$ and $[\![\dot{\pi}'\cdot\dot{X}]\!]_\Delta = [\![\dot{\pi}']\!]\cdot X^S$ where $S = comb \setminus \{\iota(\dot{a}) \mid \dot{a}\#\dot{X} \in \Delta\}$. Suppose $\dot{\pi}(\dot{a}) \neq \dot{\pi}'(\dot{a})$ implies $\Delta \vdash \dot{a}\#\dot{X}$. By Lemma 40, $[\![\dot{\pi}]\!](\iota(\dot{a})) \neq [\![\dot{\pi}']\!](\iota(\dot{a}))$ implies $\iota(\dot{a}) \notin S$. The result follows by ($=_\alpha\mathbf{X}$).

□

Definition 43. A **substitution** $\dot{\theta}$ is a function from nominal unknowns to nominal terms such that $\{\dot{X} \mid \dot{\theta}(\dot{X}) \not\equiv \dot{id}\cdot\dot{X}\}$ is finite. $\dot{\theta}, \dot{\theta}', \dot{\theta}'', \ldots$ will range over nominal substitutions. Write $\dot{id}$ for the **identity**, mapping $\dot{X}$ to $\dot{id}\cdot\dot{X}$.

Definition 44. Define a **substitution action** on nominal terms by the following rules:

$$\dot{a}\dot{\theta} \equiv \dot{a} \quad \mathsf{f}(\dot{r}_1,\ldots,\dot{r}_n)\dot{\theta} \equiv \mathsf{f}(\dot{r}_1\dot{\theta},\ldots,\dot{r}_n\dot{\theta}) \quad ([\dot{a}]\dot{r})\dot{\theta} \equiv [\dot{a}](\dot{r}\dot{\theta}) \quad (\dot{\pi}\cdot\dot{X})\dot{\theta} \equiv \dot{\pi}\cdot\dot{\theta}(\dot{X})$$

Definition 45. A **unification problem** $\dot{Pr}$ is a finite multiset of freshnesses and equalities. A **solution** to $\dot{Pr}$ is a pair $(\Delta,\dot{\theta})$ such that $\Delta \vdash \dot{a}\#\dot{r}\dot{\theta}$ for every $\dot{a}\#\dot{r} \in \dot{Pr}$, and $\Delta \vdash \dot{r}\theta = \dot{s}\theta$ for every $\dot{r} = \dot{s} \in \dot{Pr}$. This follows [27, Definition 3.1]. We extend our interpretation to solutions by:

$$[\![(\Delta,\dot{\theta})]\!](X^S) \equiv [\![\dot{\theta}(X)]\!]_\Delta \text{ if } id\cdot X^S \equiv [\![X]\!]_\Delta \quad [\![(\Delta,\dot{\theta})]\!](Y^T) \equiv id\cdot Y^T \text{ otherwise}$$

Lemma 46. $[\![\dot{r}]\!]_\Delta[\![(\Delta,\dot{\theta})]\!] \equiv [\![\dot{r}\dot{\theta}]\!]_\Delta$.

Proof. By induction on $\dot{r}$.

- The cases $\dot{a}$ and $\mathsf{f}(\dot{r}_1,\ldots,\dot{r}_n)$. Routine.
- The case $[\dot{a}]\dot{r}$. We have:

$$\begin{array}{lcll}
[\![([\dot{a}]\dot{r})\dot{\theta}]\!]_\Delta & \equiv & [\![[\dot{a}]\dot{r}\dot{\theta}]\!]_\Delta & \text{Definition 44} \\
& \equiv & [\iota(\dot{a})][\![\dot{r}\dot{\theta}]\!]_\Delta & \text{Definition 38} \\
& \equiv & [\iota(\dot{a})][\![\dot{r}]\!]_\Delta[\![(\Delta,\dot{\theta})]\!] & \text{Inductive hypothesis} \\
& \equiv & ([\iota(\dot{a})][\![\dot{r}]\!]_\Delta)[\![(\Delta,\dot{\theta})]\!] & \text{Fact} \\
& \equiv & [\![[\dot{a}]\dot{r}]\!]_\Delta[\![(\Delta,\dot{\theta})]\!] & \text{Definition 38}
\end{array}$$

The result follows.

- The case $\dot{\pi} \cdot \dot{X}$. We have:

$$\begin{array}{lcll} [\![(\dot{\pi} \cdot \dot{X})\dot{\theta}]\!]_\Delta & \equiv & [\![\dot{\pi} \cdot \dot{\theta}(\dot{X})]\!]_\Delta & \text{Definition 44} \\ & \equiv & [\![\dot{\pi}]\!] \cdot [\![\dot{\theta}(\dot{X})]\!]_\Delta & \text{Definition 38} \\ & \equiv & [\![\dot{\pi}]\!] \cdot [\![\dot{\theta}]\!]([\![\dot{X}]\!]_\Delta) & \end{array}$$

  The result follows.

□

Definition 47. Define $[\![\dot{P}r]\!]_\Delta$ by mapping $\dot{r} = \dot{s}$ to $[\![\dot{r}]\!]_\Delta \;?{=}?\; [\![\dot{s}]\!]_\Delta$ and mapping $\dot{a}\#\dot{r}$ to $(b\ \iota(\dot{a})) \cdot [\![\dot{r}]\!]_\Delta \;?{=}?\; [\![\dot{r}]\!]_\Delta$, for some choice of fresh b (so $b \notin fa([\![\dot{r}]\!]_\Delta)$; in fact, it suffices to choose some $b \notin comb$).

Lemma 48. *Suppose $b \notin fa(r)$. Then $a \notin fa(r)$ if and only if $(b\ a) \cdot r =_\alpha r$.*

Proof. We handle the two implications separately.

- The case $a \notin fa(r)$ implies $(b\ a) \cdot r =_\alpha r$. We proceed by induction on r.
  - The case c. Straightforward.
  - The case $\mathsf{f}(r_1, \ldots, r_n)$. Easy, by the inductive hypotheses.
  - The cases $[a]r$ and $[b]r$. We handle only the $[a]r$ case, as the other is similar. We show $(b\ a) \cdot [a]r =_\alpha [a]r$ where $b \notin fa([a]r)$, hence $b \notin fa(r)$ and $a \notin fa(r)$. By Definition 7, $(b\ a) \cdot [a]r =_\alpha [b](b\ a) \cdot r$. By the rules in Definition 11, we must show $(a\ b) \cdot ((b\ a) \cdot r) =_\alpha r$ where $a \notin fa((b\ a) \cdot r)$. By Lemma 16, this is equivalent to $b \notin fa(r)$, which we have by assumption. By Lemma 15, $(a\ b) \cdot ((b\ a) \cdot r) =_\alpha ((a\ b) \circ (b\ a)) \cdot r$, and as $\pi = \pi^{-1}$, we have $r =_\alpha r$. The result follows from Theorem 24.
  - The case $[c]r$. Suppose $b \notin fa([c]r)$, $a \notin fa([c]r)$ and $a, b \notin fa(r)$. We show $(b\ a) \cdot [c]r =_\alpha [c]r$. By Definition 7, $(b\ a) \cdot [c]r \equiv [c](b\ a) \cdot r$. Using $(=_\alpha[\mathbf{a}])$ and the inductive hypothesis, $(b\ a) \cdot r =_\alpha r$.
  - The case $\pi \cdot X^S$. Suppose $b \notin fa(\pi \cdot X^S)$, $a \notin fa(\pi \cdot X^S)$ and $a, b \notin \pi \cdot S$. By Definition 7, $(b\ a) \cdot (\pi \cdot X^S) \equiv ((b\ a) \circ \pi) \cdot X^S$. Using $(=_\alpha\mathbf{X})$, $((b\ a) \circ \pi) \cdot X^S =_\alpha \pi \cdot X^S$ whenever $((b\ a) \circ \pi)|_S = \pi|_S$. As $a, b \notin \pi \cdot S$, $((b\ a) \circ \pi)|_S = \pi|_S$, and the result follows.
- The case $(b\ a) \cdot r =_\alpha r$ implies $a \notin fa(r)$. We proceed by induction on r.
  - The case a, b, c and $\mathsf{f}(r_1, \ldots, r_n)$. These are routine.
  - The case $\mathsf{f}(r_1, \ldots, r_n)$. By hypotheses, $(b\ a) \cdot r_1 =_\alpha r_1 \ldots (b\ a) \cdot r_n =_\alpha r_n$ implies $a \notin fa(r_1) \ldots a \notin fa(r_n)$. As $fa(\mathsf{f}(r_1, \ldots, r_n)) = fa(r_1) \cup \ldots \cup fa(r_n)$, the result follows.
  - The cases $[a]r$ and $[b]r$. We handle only the $[a]r$ case, as the other is similar. Suppose $(b\ a) \cdot [a]r =_\alpha [a]r$. By Definition 7, $(b\ a) \cdot [a]r \equiv [b](b\ a) \cdot r$. By the rules in Definition 11, $[b](b\ a) \cdot r =_\alpha [a]r$ whenever $(a\ b) \cdot ((b\ a) \cdot r) =_\alpha r$ with $a \notin fa((b\ a) \cdot r)$. By Definition 7, and as $\pi = \pi^{-1}$, $(a\ b) \cdot ((b\ a) \cdot r) \equiv r$. By assumption, $b \notin fa(r)$. By Lemma 16, $a \notin fa((b\ a) \cdot r)$. The result follows.
  - The case $[c]r$. By hypothesis, $(b\ a) \cdot r =_\alpha r$ implies $a \notin fa(r)$. Then $[c](b\ a) \cdot r \equiv (b\ a) \cdot [c]r$. The result follows.
  - The case $\pi \cdot X^S$. Suppose $(b\ a) \cdot \pi \cdot X^S =_\alpha \pi \cdot X^S$. By Definition 7, $(b\ a) \cdot \pi \cdot X^S \equiv ((b\ a) \circ \pi) \cdot X^S$. Using $(=_\alpha\mathbf{X})$, $((b\ a) \circ \pi) \cdot X^S =_\alpha \pi \cdot X^S$ whenever $(b\ a) \circ \pi|_S = \pi|_S$. However, $(b\ a) \circ \pi|_S = \pi|_S$ only when $b, a \notin \pi \cdot S$. The result follows.

□

No solutions go missing, moving from the nominal to the permissive world:

Theorem 49. $(\Delta, \dot\theta)$ *solves* $\dot Pr$ *if and only if* $[\![(\Delta, \dot\theta)]\!]$ *solves* $[\![\dot Pr]\!]_\Delta$.

Proof. We handle the two implications separately:

- The case $(\Delta, \dot\theta)$ solves $\dot Pr$ implies $[\![(\Delta, \dot\theta)]\!]$ solves $[\![\dot Pr]\!]_\Delta$. Suppose $\Delta \vdash \dot r\dot\theta = \dot s\dot\theta$. By Lemma 46 and Theorem 42, $[\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!] =_\alpha [\![\dot s]\!]_\Delta[\![(\Delta, \dot\theta)]\!]$.
  Suppose $\Delta \vdash a\#\dot r\dot\theta$. By Lemma 40, $\iota(\dot a) \notin fa([\![\dot r\dot\theta]\!]_\Delta)$. By Lemma 46, $\iota(\dot a) \notin fa([\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!])$. By Lemma 48, $(b\ \iota(\dot a)) \cdot [\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!] =_\alpha [\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!]$, where b is fresh (see Definition 47). By Lemma 28, $((b\ \iota(\dot a)) \cdot [\![\dot r]\!]_\Delta)[\![(\Delta, \dot\theta)]\!] =_\alpha [\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!]$. The result follows.
- The case $[\![(\Delta, \dot\theta)]\!]$ solves $[\![\dot Pr]\!]_\Delta$ implies $(\Delta, \dot\theta)$ solves $\dot Pr$. Suppose $[\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!] =_\alpha [\![\dot s]\!]_\Delta[\![(\Delta, \dot\theta)]\!]$. By Theorem 41, $\Delta \vdash r\theta = s\theta$.
  Suppose $((b\ \iota(\dot a)) \cdot [\![\dot r]\!]_\Delta)[\![(\Delta, \dot\theta)]\!] =_\alpha [\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!]$. By Lemma 28, $(b\ \iota(\dot a)) \cdot [\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!] =_\alpha [\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!]$. By Lemma 48, $\iota(\dot a) \notin fa([\![\dot r]\!]_\Delta[\![(\Delta, \dot\theta)]\!])$. By Lemma 46, $\iota(\dot a) \notin fa([\![\dot r\dot\theta]\!]_\Delta)$. By Lemma 40, $\Delta \vdash a\#\dot r\dot\theta$. The result follows.

□

5. Support inclusion problems

Nominal unification has 'freshness problems'; the algorithm of [27] solves these concurrently with equality problems. We prefer to factor the algorithm differently, so that problems to do with free atoms are solved separately from problems to do with equalities. The linkage is isolated in rule (**I3**) (Definition 75). Permissions sets differ finitely from *comb* or $\varnothing$, so although we are manipulating infinite sets, we may represent them as finite data structures in any implementation.

5.1. Simplification reduction and normal forms

Definition 50. A **support inclusion** is a pair $r \sqsubseteq T$ of a term and a permissions set. A **support inclusion problem** is a finite multiset of support inclusions; *Inc* will range over support inclusion problems. Call θ a **solution** to *Inc* when $fa(r\theta) \subseteq T$ for every $r \sqsubseteq T \in Inc$. Write $Sol(Inc)$ for the solutions of *Inc*. Call *Inc* **solvable** when $Sol(Inc) \neq \varnothing$.

Definition 51. Define a **simplification** rewrite relation by:

$(\sqsubseteq\mathbf{a})$	$a \sqsubseteq T, Inc \Longrightarrow Inc$	$(a \in T)$
$(\sqsubseteq\mathsf{f})$	$\mathsf{f}(r_1,\ldots,r_n) \sqsubseteq T, Inc \Longrightarrow r_1 \sqsubseteq T, \ldots, r_n \sqsubseteq T, Inc$	
$(\sqsubseteq[])$	$[a]r \sqsubseteq T, Inc \Longrightarrow r \sqsubseteq T \cup \{a\},\ Inc$	
$(\sqsubseteq\mathbf{X})$	$\pi \cdot X^S \sqsubseteq T, Inc \Longrightarrow X^S \sqsubseteq \pi^{-1} \cdot T,\ Inc$	$(S \not\subseteq \pi^{-1} \cdot T,\ \pi \neq id)$
$(\sqsubseteq\mathbf{X}')$	$\pi \cdot X^S \sqsubseteq T, Inc \Longrightarrow Inc$	$(S \subseteq \pi^{-1} \cdot T)$

Theorem 52. *If* $Inc \Longrightarrow Inc'$ *then* $Sol(Inc) = Sol(Inc')$.

Proof. First, we make the following claims:

Claim 1: If $a \in T$ then $fa(a\theta) \subseteq T$ always. As $fa(a\theta) = fa(a) = \{a\}$.

Claim 2: $fa(\mathsf{f}(r_1,\ldots,r_n)\theta) \subseteq T$ if and only if $fa(r_i\theta) \subseteq T$ for $1 \leq i \leq n$. As $fa(\mathsf{f}(r_1,\ldots,r_n)) = \bigcup_{1\leq i\leq n} fa(r_i)$, and $\mathsf{f}(r_1,\ldots,r_n)\theta \equiv \mathsf{f}(r_1\theta,\ldots,r_n\theta)$.

Claim 3: $fa(([a]s)\theta) \subseteq T$ if and only if $fa(s\theta) \subseteq T \cup \{a\}$. Suppose $fa(([a]s)\theta) \subseteq T$, therefore $fa([a]s\theta) \subseteq T$. Then $fa(s\theta) \setminus \{a\} \subseteq T$, therefore $fa(s\theta) \subseteq T \cup \{a\}$. The result follows. The reverse direction is similar.

Claim 4: $fa((\pi \cdot X^S)\theta) \subseteq T$ if and only if $fa(X^S\theta) \subseteq \pi^{-1} \cdot T$. We consider only one case. Suppose $\theta = [X^S{:=}t]$ and $fa(t) \subseteq S$, therefore $fa((\pi \cdot X^S)[X^S{:=}t]) = fa(\pi \cdot t)$ hence $fa(\pi \cdot t) \subseteq T$ by assumption. By Lemma 16, $\pi \cdot fa(t) \subseteq T$. By Lemmas 15 and 16, $(\pi^{-1} \circ \pi) \cdot fa(t) \subseteq \pi^{-1} \cdot T$. As $\pi^{-1} \circ \pi = id$, we have $fa(X^S[X^S{:=}t]) \subseteq \pi^{-1} \cdot T$. The result follows.
Alternatively, suppose $fa(t) \not\subseteq S$. Then $fa((\pi \cdot X^S)[X^S{:=}t]) = fa(\pi \cdot X^S)$. By Lemma 16, $\pi \cdot fa(X^S) \subseteq T$. By Lemmas 15 and 16, $fa(X^S[X^S{:=}t]) \subseteq \pi^{-1} \cdot T$. The result follows. The reverse direction is similar.

Claim 5: If $S \subseteq \pi^{-1} \cdot T$ then $fa(\pi \cdot X^S\theta) \subseteq T$ always. Note, $S \subseteq \pi^{-1} \cdot T$ if and only if $\pi \cdot S \subseteq T$ and $fa(\pi \cdot X^S) = \pi \cdot S$. By Lemmas 28 and 16, $fa(\pi \cdot X^S\theta) = \pi \cdot fa(\theta(X^S)) \subseteq \pi \cdot S$. Then, $fa(\pi \cdot X^S\theta) \subseteq T$. The result follows.

We proceed by case analysis on $Inc \Longrightarrow Inc'$ (Definition 51):

- The case ($\sqsubseteq\mathbf{a}$). Suppose $a \in T$. If $\theta \in Sol(a \sqsubseteq T, Inc')$ then $\theta \in Sol(Inc')$ and the result follows. Otherwise, suppose $\theta \in Sol(Inc')$. Using Claim 1, $fa(a\theta) \subseteq T$. The result follows.
- The case ($\sqsubseteq\mathsf{f}$). From Claim 2.
- The case ($\sqsubseteq[]$). If $\theta \in Sol(r \sqsubseteq T \cup \{a\}, Inc')$ then $fa(r\theta) \subseteq T \cup \{a\}$. By Claim 3, $fa([a](r\theta)) \subseteq T$. As $fa([a](r\theta)) = fa(([a]r)\theta)$ and $\theta \in Sol(Inc')$. The result follows. The reverse direction is similar.
- The case ($\sqsubseteq\mathbf{X}$). Suppose $S \not\subseteq \pi^{-1} \cdot T$, $\pi \neq id$ and $\theta \in Sol(\pi \cdot X^S \sqsubseteq T, Inc)$, so $fa((\pi \cdot X^S)\theta) \subseteq T$. By Claim 4, $fa(X^S\theta) \subseteq \pi^{-1} \cdot T$, and as $\theta \in Sol(Inc')$. The result follows. The reverse direction is similar.
- The case ($\sqsubseteq\mathbf{X}'$). Suppose $S \subseteq \pi^{-1} \cdot T$. If $\theta \in Sol(\pi \cdot X^S, Inc')$ then $\theta \in Sol(Inc')$ and the result follows. Otherwise, suppose $\theta \in Sol(Inc')$. By Claim 5, $fa(\pi \cdot X^S\theta) \subseteq T$. By Lemma 28, $fa((\pi \cdot X^S)\theta) \subseteq T$. The result follows.

□

Definition 53. Define the **size** of a support inclusion problem $size(Inc)$ to be a tuple (T, A, P, S), where:

- T is the number of term formers appearing within terms in Inc,
- A is the number of abstractions appearing within terms in Inc,
- P is the number of permutations, distinct from the identity permutation, appearing within terms in Inc, and
- S is the number of support inclusions within Inc.

Order tuples lexicographically.

Theorem 54. *Support inclusion problem simplication is strongly normalizing.*

Proof. By case analysis on $r \sqsubseteq T, Inc'$, showing that each rule reduces the measure (Definition 53).

- The case $a \sqsubseteq T, Inc'$. Suppose $a \in T$, $size(a \sqsubseteq T, Inc') = (T, A, P, S)$, and $a \sqsubseteq T, Inc' \Longrightarrow Inc'$ using ($\sqsubseteq\mathbf{a}$). Then $size(Inc') = (T, A, P, S - 1)$.
- The case $\mathsf{f}(r_1, \ldots, r_n) \sqsubseteq T, Inc'$. Suppose $size(\mathsf{f}(r_1, \ldots, r_n) \sqsubseteq T, Inc') = (T, A, P, S)$ and $\mathsf{f}(r_1, \ldots, r_n) \sqsubseteq T, Inc' \Longrightarrow r_1 \sqsubseteq T, \ldots, r_n \sqsubseteq T, Inc'$ by ($\sqsubseteq\mathsf{f}$). Then $size(r_1 \sqsubseteq T, \ldots, r_n \sqsubseteq T, Inc') = (T - 1, A, P, S + n - 1)$. The result follows from the ordering.
- The case $[a]r \sqsubseteq T, Inc'$. Suppose $size([a]r \sqsubseteq T, Inc') = (T, A, P, S)$ and $[a]r \sqsubseteq T, Inc' \Longrightarrow r \sqsubseteq T \cup \{a\}, Inc'$ using ($\sqsubseteq[]$). Then $size(r \sqsubseteq T \cup \{a\}, Inc') = (T, A - 1, P, S)$ and the result follows.
- The case $\pi \cdot X^S \sqsubseteq T, Inc'$. Suppose $size(\pi \cdot X^S \sqsubseteq T, Inc') = (T, A, P, S)$. Then, if $S \subseteq \pi^{-1} \cdot T$, we have $\pi \cdot X^S \sqsubseteq T, Inc' \Longrightarrow Inc'$ using ($\sqsubseteq\mathbf{X}'$), with measure

$(T, A, P, S-1)$. Otherwise, if we have $S \not\sqsubseteq \pi^{-1} \cdot T$ and $\pi \neq id$, we have $\pi \cdot X^S \sqsubseteq T, Inc' \Longrightarrow X^S \sqsubseteq \pi^{-1} \cdot T, Inc'$ with measure $(T, A, P-1, S)$.

□

Theorem 55. *Support inclusion problem simplification is strongly confluent.*

As an immediate corollary, support inclusion simplification is confluent.

Proof. We prove the result by induction on the cardinality of Inc.

- The case $Inc = \varnothing$. As no rewrites are applicable.
- The case $Inc = r \sqsubseteq T, Inc'$. Note for each $r \sqsubseteq T$ only one simplification rule may be applied. The result then follows from the inductive hypothesis.

The corollary follows, as all strongly confluent rewrite relations are confluent (see [1] for details). □

We conclude with a few useful observations:

Definition 56. Write $nf(Inc)$ for the unique $\Longrightarrow$-normal form of Inc, guaranteed to exist by Theorems 54 and 55.

Definition 57. Call Inc **consistent** when $a \sqsubseteq T \notin nf(Inc)$ for all atoms a and permissions sets T.

Lemma 58. *If Inc is consistent then all $inc \in nf(Inc)$ have the form $X^S \sqsubseteq T$ where $S \not\sqsubseteq T$.*

Proof. By inspection. □

5.2. Building solutions

Our main results are Theorems 62 and 70.

Definition 59. Define $fV(Inc)$ by $fV(Inc) = \bigcup\{fV(r) \mid \exists T.r \sqsubseteq T \in Inc\}$. ($fV(Inc)$ is 'the unknowns appearing in terms appearing in Inc'.)

Recall from Definition 3 that $\mathcal{V}$ ranges over finite sets of unknowns.

Definition 60. Suppose Inc is consistent. For every $X^S \in \mathcal{V}$ make a fixed but arbitrary choice of $X'^{S'}$ such that $X'^{S'} \notin \mathcal{V}$ and $S' = \bigcap\{T \mid X^S \sqsubseteq T \in nf(Inc)\}$.

We make our choice injectively; for distinct $X^S \in fV(Inc)$ and $Y^T \in fV(Inc)$, we choose $X'^{S'}$ and $Y'^{T'}$ distinct. It will be convenient to write $\mathcal{V}'^{\mathcal{V}}_{Inc}$ for the set of our choices $\{X'^{S'} \mid X^S \in \mathcal{V}\}$. Define a substitution $\rho^{\mathcal{V}}_{Inc}$ by:

$$\rho^{\mathcal{V}}_{Inc}(X^S) \equiv id \cdot X'^{S'} \text{ if } X^S \in \mathcal{V} \qquad \rho^{\mathcal{V}}_{Inc}(Y^T) \equiv id \cdot Y^T \text{ otherwise.}$$

It is easy to verify that $fa(\rho^{\mathcal{V}}_{Inc}(X^S)) \subseteq S$ always.

Lemma 61. *If Inc is consistent then $\rho^{\mathcal{V}}_{Inc} \in Sol(Inc)$. ('$\rho^{\mathcal{V}}_{Inc}$ solves Inc'.)*

Proof. Suppose Inc is a $\Longrightarrow$-normal form. If $X^S \sqsubseteq T \in Inc$ then $\rho^{\mathcal{V}}_{Inc}(X) = id \cdot X'^{S'}$ for an S' which satisfies $S' \subseteq T$. The result follows.

More generally, if Inc is not a $\Longrightarrow$-normal form, by Theorem 52 $Sol(Inc) = Sol(nf(Inc))$, and we use the previous paragraph. □

Theorem 62. *Inc is consistent (Definition 57) if and only if Inc is solvable (Definition 50).*

Proof. By Theorem 52 $Sol(Inc) = Sol(nf(Inc))$, so it suffices to show the result for the case when Inc is a $\Longrightarrow$-normal form.

Suppose Inc is inconsistent, so $nf(Inc)$ contains a support inclusions of the form $a \sqsubseteq T$ where $a \notin T$. Then $a\theta \equiv a$ always, so there is no substitution θ such that $a\theta \subseteq T$. Conversely, if Inc is consistent, the result follows by Lemma 61. □

Definition 63. Suppose that Inc is consistent, $fV(Inc) \subseteq \mathcal{V}$, and $\theta \in Sol(Inc)$. Define a substitution $\theta - \rho^{\mathcal{V}}_{Inc}$ by:

- $(\theta - \rho^{\mathcal{V}}_{Inc})(X'^{S'}) \equiv \theta(X^S)$ if $X^S \in \mathcal{V}$ and $\rho^{\mathcal{V}}_{Inc}(X^S) \equiv id \cdot X'^{S'}$.
- $(\theta - \rho^{\mathcal{V}}_{Inc})(X^S) \equiv \theta(X^S)$ if $X^S \notin \mathcal{V}$.

We check that Definition 63 is well-defined:

Lemma 64. *If $\theta - \rho^{\mathcal{V}}_{Inc}$ exists then it is well-defined.*

Proof. Suppose $\theta - \rho^{\mathcal{V}}_{Inc}$ exists. Then:

- Suppose $X^S \neq Y^T$, $X^S \notin \mathcal{V}$ and $Y^T \notin \mathcal{V}$. By Definition 63, $(\theta - \rho^{\mathcal{V}}_{Inc})(X^S) \equiv \theta(X^S)$ and $(\theta - \rho^{\mathcal{V}}_{Inc})(Y^T) \equiv \theta(Y^T)$. The result follows, as substitutions are well-defined.
- Suppose $X'^{S'} \neq Y'^{T'}$, $\rho^{\mathcal{V}}_{Inc}(X^S) \equiv id \cdot X'^{S'}$, $\rho^{\mathcal{V}}_{Inc}(Y^T) \equiv id \cdot Y'^{T'}$ and $X^S, Y^T \notin \mathcal{V}$. Then $(\theta - \rho^{\mathcal{V}}_{Inc})(X'^{S'}) \equiv \theta(X^S)$ and $(\theta - \rho^{\mathcal{V}}_{Inc})(Y'^{T'}) \equiv \theta(Y^T)$. Since $X^S \neq Y^T$, the result follows as substitutions are well-defined.
- Suppose $X'^{S'} \neq Y^T$, $\rho^{\mathcal{V}}_{Inc}(X^S) \equiv id \cdot X'^{S'}$, $X^S \notin \mathcal{V}$ and $Y^T \in \mathcal{V}$. Then $(\theta - \rho^{\mathcal{V}}_{Inc})(Y^T) \equiv \theta(Y^T)$ and $(\theta - \rho^{\mathcal{V}}_{Inc})(X'^{S'}) \equiv \theta(X^S)$. By Definition 60, $\rho^{\mathcal{V}}_{Inc}(X^S) \not\equiv id \cdot Y^T$ as $Y^T \in \mathcal{V}$. The result follows as substitutions are well-defined.

The case $Y'^{T'} \neq X^S$, $\rho^{\mathcal{V}}_{Inc}(Y^T) \equiv id \cdot Y'^{T'}$, $Y^T \notin \mathcal{V}$ and $X^S \in \mathcal{V}$ is similar to the case for $X'^{S'} \neq Y^T$, $\rho^{\mathcal{V}}_{Inc}(X^S) \equiv id \cdot X'^{S'}$, $X^S \notin \mathcal{V}$ and $Y^T \in \mathcal{V}$. □

Lemma 65. *If $\theta \in Sol(Inc)$ then $\rho^{\mathcal{V}}_{Inc}$ exists.*

Proof. By assumption, Inc is solvable. By Theorem 62, Inc is consistent. Using Definition 60, $\rho^{\mathcal{V}}_{Inc}$ exists. □

Lemma 66. *If $\rho^{\mathcal{V}}_{Inc}$ exists, then it is well-defined.*

Proof. Suppose $\rho^{\mathcal{V}}_{Inc}$ exists and $X^S \neq Y^T$. Then:

- $X^S \in \mathcal{V}$ and $Y^T \in \mathcal{V}$. Then $\rho^{\mathcal{V}}_{Inc}(X^S) = id \cdot X'^{S'}$ and $\rho^{\mathcal{V}}_{Inc}(Y^T) = id \cdot Y'^{T'}$. By Definition 60, $X'^{S'}$ and $Y'^{T'}$ are chosen so $X'^{S'} \neq Y'^{T'}$. The result follows.
- $X^S \notin \mathcal{V}$ and $Y^T \notin \mathcal{V}$. Then $\rho^{\mathcal{V}}_{Inc}(X^S) = id \cdot X^S$ and $\rho^{\mathcal{V}}_{Inc}(Y^T) = id \cdot Y^T$. The result follows.
- $X^S \in \mathcal{V}$ and $Y^T \notin \mathcal{V}$. Then $\rho^{\mathcal{V}}_{Inc}(Y^T) = id \cdot Y^T$ and $\rho^{\mathcal{V}}_{Inc}(X^S) = id \cdot X'^{S'}$ with $X'^{S'} \notin \mathcal{V}$. The result follows.
- The case $X^S \notin \mathcal{V}$ and $Y^T \in \mathcal{V}$ is similar to the case for $X^S \in \mathcal{V}$ and $Y^T \notin \mathcal{V}$.

□

Lemma 67. *If $Inc \Longrightarrow Inc'$ then $fV(Inc') \subseteq fV(Inc)$.*

Proof. By case analysis on the rules defining $\Longrightarrow$ (Definition 51).

- The case ($\sqsubseteq$**a**). Suppose $a \in T$ and $a \sqsubseteq T, Inc' \Longrightarrow Inc'$ using ($\sqsubseteq$**a**). By Definition 59, $fV(a \sqsubseteq T, Inc') = fV(Inc')$. The result follows.
- The case ($\sqsubseteq$f). Suppose $\mathsf{f}(r_1, \ldots, r_n) \sqsubseteq T, Inc' \Longrightarrow r_1 \sqsubseteq T, \ldots, r_n \sqsubseteq T, Inc'$ using ($\sqsubseteq$f). By Definition 59, $fV(\mathsf{f}(r_1, \ldots, r_n) \sqsubseteq T, Inc') = fV(r_1 \sqsubseteq T, \ldots, r_n \sqsubseteq T, Inc')$. The result follows.

- The case ($\sqsubseteq[]$). Suppose $[a]r \sqsubseteq T, Inc' \Longrightarrow r \sqsubseteq T \cup \{a\}, Inc'$ using ($\sqsubseteq[]$). By Definition 59, $fV([a]r) = fV(r)$. By Definition 10, $fV([a]r \sqsubseteq T, Inc') = fV(r \sqsubseteq T \cup \{a\}, Inc')$. The result follows.
- The case ($\sqsubseteq\mathbf{X}$). Suppose $S \not\subseteq \pi^{-1} \cdot T$, $\pi \neq id$ and $\pi \cdot X^S \sqsubseteq T, Inc' \Longrightarrow X^S \sqsubseteq \pi^{-1} \cdot T, Inc'$ using ($\sqsubseteq\mathbf{X}$). By Definition 10, $fV(\pi \cdot X^S) = X^S = fV(X^S)$. By Definition 59, $fV(\pi \cdot X^S \sqsubseteq T, Inc') = fV(X^S \sqsubseteq \pi^{-1} \cdot T, Inc')$. The result follows.
- The case ($\sqsubseteq\mathbf{X}'$). Suppose $S \subseteq \pi^{-1} \cdot T$ and $\pi \cdot X^S, Inc' \Longrightarrow Inc'$ using ($\sqsubseteq\mathbf{X}'$). By Definition 59, $fV(Inc') \subseteq fV(\pi \cdot X^S, Inc')$. The result follows.

□

Corollary 68. $fV(nf(Inc)) \subseteq fV(Inc)$

Proof. By Lemma 67. □

Lemma 69. *If $\theta \in Sol(Inc)$ and $fV(Inc) \subseteq \mathcal{V}$ then $\theta{-}\rho^{\mathcal{V}}_{Inc}$ is a substitution.*

Proof. By Lemma 65, $\rho^{\mathcal{V}}_{Inc}$ exists. We show $fa((\theta{-}\rho^{\mathcal{V}}_{Inc})(X'^{S'})) \subseteq S$ by cases:

- The case $id \cdot X'^{S'} \equiv \rho(X^S)$ for $X^S \in \mathcal{V}$.
  By Corollary 68, $fV(nf(Inc)) \subseteq fV(Inc)$. Then $fV(nf(Inc)) \subseteq \mathcal{V}$, as $fV(Inc) \subseteq \mathcal{V}$ by assumpyion. There are two sub-cases:
  - The case $X^S \notin fV(nf(Inc))$. Then $S = S'$ and $(\theta{-}\rho^{\mathcal{V}}_{Inc})(X'^S) = \theta(X^S)$ by Definition 63. By assumption, $fa(\theta(X^S)) \subseteq S$. The result follows.
  - The case $X^S \in fV(nf(Inc))$. By assumption, $\theta \in Sol(Inc)$ so $\theta \in Sol(nf(Inc))$ by Theorem 52. by Definition 50, $fa(\theta(X^S)) \subseteq T$ for every T such that $X^S \sqsubseteq T \in nf(Inc)$. By Definition 63, $S' = \bigcap\{T \mid X^S \sqsubseteq T \in nf(Inc)\}$. The result follows.
- Otherwise, $(\theta{-}\rho^{\mathcal{V}}_{Inc})(X^S) \equiv \theta(X^S)$ and $fa(\theta(X^S)) \subseteq S$ by assumption.

□

Theorem 70. *If $\theta \in Sol(Inc)$ and $fV(Inc) \subseteq \mathcal{V}$ then $\theta(X^S) \equiv (\rho^{\mathcal{V}}_{Inc} \circ (\theta{-}\rho^{\mathcal{V}}_{Inc}))(X^S)$ for every $X^S \in \mathcal{V}$.*

Proof. For some fresh $X'^S \notin \mathcal{V}$, $\rho(X^S) \equiv id \cdot X'^S$, and $(\theta{-}\rho^{\mathcal{V}}_{Inc})(X'^S) \equiv \theta(X^S)$. The result follows by Lemma 15. □

6. Permissive nominal unification problems

6.1. Problems, solutions, the unification algorithm

Definition 71. An **equality** is a pair $r \;?{=}?\; s$. A **problem** Pr is a finite multiset of equalities. Define $Pr\theta$ by:

$$Pr\theta = \{r\theta \;?{=}?\; s\theta \mid r \;?{=}?\; s \in Pr\}$$

Definition 72. θ **solves** Pr when $r \;?{=}?\; s \in Pr$ implies $r\theta =_\alpha s\theta$. Write $Sol(Pr)$ for the set of solutions to Pr. Call Pr **solvable** when $Sol(Pr)$ is non-empty.

A solution to Pr 'makes the equalities valid', as for first- and higher-order unification. This simplifies the nominal unification notion of solution (Definition 45 or [27, Definition 3.1]) based on 'a substitution + a freshness context'. We can do this, because in permissive nominal terms, freshness information is fixed. Lemma 73 will be useful:

Lemma 73. *$\theta \circ \theta' \in Sol(Pr)$ if and only if $\theta' \in Sol(Pr\theta)$.*

$$
\begin{array}{lll}
(?{=}_?\mathbf{a}) & \mathcal{V}; a \;?{=}_?\; a,\ Pr & \Longrightarrow \quad \mathcal{V}; Pr \\
(?{=}_?\mathsf{f}) & \mathcal{V}; \mathsf{f}(r_1, \ldots) \;?{=}_?\; \mathsf{f}(s_1, \ldots),\ Pr & \Longrightarrow \quad \mathcal{V}; r_1 \;?{=}_?\; s_1,\ \ldots,\ Pr \\
(?{=}_?[\mathbf{a}]) & \mathcal{V}; [a]r \;?{=}_?\; [a]s,\ Pr & \Longrightarrow \quad \mathcal{V}; r \;?{=}_?\; s,\ Pr \\
(?{=}_?[\mathbf{b}]) & \mathcal{V}; [a]r \;?{=}_?\; [b]s,\ Pr & \Longrightarrow \quad \mathcal{V}; (b\ a) \cdot r \;?{=}_?\; s,\ Pr \\
 & & \quad (b \notin fa(r)) \\
(?{=}_?\mathbf{X}) & \mathcal{V}; \pi \cdot X^S \;?{=}_?\; \pi \cdot X^S,\ Pr & \Longrightarrow \quad \mathcal{V}; Pr \\
(\mathbf{I1}) & \mathcal{V}; \pi \cdot X^S \;?{=}_?\; s,\ Pr & \overset{[X^S := \pi^{-1} \cdot s]}{\Longrightarrow} \quad \mathcal{V}; Pr[X^S := \pi^{-1} \cdot s] \\
 & & \quad (X^S \notin fV(s),\ fa(s) \subseteq \pi \cdot S) \\
(\mathbf{I2}) & \mathcal{V}; r \;?{=}_?\; \pi \cdot X^S,\ Pr & \overset{[X^S := \pi^{-1} \cdot r]}{\Longrightarrow} \quad \mathcal{V}; Pr[X^S := \pi^{-1} \cdot r] \\
 & & \quad (X^S \notin fV(r),\ fa(r) \subseteq \pi \cdot S) \\
(\mathbf{I3}) & \mathcal{V}; Pr & \overset{\rho^{\mathcal{V}}_{Pr_\sqsubseteq}}{\Longrightarrow} \quad \mathcal{V} \cup \mathcal{V}'^{\mathcal{V}}_{Pr_\sqsubseteq}; Pr(\rho^{\mathcal{V}}_{Pr_\sqsubseteq}) \\
 & & \quad (Pr_\sqsubseteq \text{ consistent and non-trivial})
\end{array}
$$

Figure 3: Simplification rules for problems

Proof. Suppose $\theta \circ \theta' \in Sol(Pr)$ and $r \;?{=}_?\; s \in Pr$. We have:

$$
\begin{array}{rcll}
(r\theta)\theta' & \equiv & r(\theta \circ \theta') & \text{Theorem 33} \\
 & =_\alpha & s(\theta \circ \theta') & \text{Assumption} \\
 & \equiv & (s\theta)\theta' & \text{Theorem 33}
\end{array}
$$

By Definition 72, $\theta' \in Sol(Pr\theta)$.

For the reverse implication, suppose $\theta' \in Sol(Pr\theta)$ and $r\theta \;?{=}_?\; s\theta \in Sol(Pr\theta)$. Then:

$$
\begin{array}{rcll}
r(\theta \circ \theta') & \equiv & (r\theta)\theta' & \text{Theorem 33} \\
 & =_\alpha & (s\theta)\theta' & \text{Assumption} \\
 & \equiv & s(\theta \circ \theta') & \text{Theorem 33}
\end{array}
$$

By Definition 72, $\theta \circ \theta' \in Sol(Pr)$. The result follows. □

Definition 74. If Pr is a problem, define a support inclusion problem $Pr_\sqsubseteq$ by:

$$Pr_\sqsubseteq = \{r \sqsubseteq fa(s),\ s \sqsubseteq fa(r) \mid r \;?{=}_?\; s \in Pr\}$$

Call a support inclusion problem Inc **non-trivial** when $nf(Inc) \neq \varnothing$.

Definition 75. Define a **simplification** rewrite relation $\mathcal{V}; Pr \Longrightarrow \mathcal{V}'; Pr'$ on unification problems by the rules in Figure 3.[3] Call $(?{=}_?\mathbf{a})$, $(?{=}_?\mathsf{f})$, $(?{=}_?[\mathbf{a}])$, $(?{=}_?[\mathbf{b}])$, and $(?{=}_?\mathbf{X})$ **non-instantiating rules**.

Call (**I1**), (**I2**), and (**I3**) **instantiating rules**. Write $\Longrightarrow\!\!\!\!\!^*$ for the transitive and reflexive closure of $\Longrightarrow$.

In (**I3**) we insist that $Pr_\sqsubseteq$ is non-trivial to avoid indefinite rewrites. We insist $Pr_\sqsubseteq$ is consistent so that $\rho^{\mathcal{V}}_{Pr_\sqsubseteq}$ exists. $\rho^{\mathcal{V}}_{Pr_\sqsubseteq}$ and $\mathcal{V}'^{\mathcal{V}}_{Pr_\sqsubseteq}$ are defined in Definition 60.

Lemma 76. *If $\mathcal{V}; Pr \Longrightarrow \mathcal{V}; Pr'$ by a non-instantiating rule then $Sol(Pr) = Sol(Pr')$.*

[3] Note to referees: an error in a previous version of this paper, which made the algorithm incomplete, has been corrected.

Proof. As the empty set cannot be simplified, it must be the case that $Pr = r \ {}_?{=}_? \ s, Pr'$. It suffices to perform case analysis on the simplification of $r \ {}_?{=}_? \ s$. We assume, without loss of generality, that Pr' has been simplified by non-instantiating rules as much as possible.

- The cases $({}_?{=}_?\mathbf{a})$, $({}_?{=}_?\mathsf{f})$ and $({}_?{=}_?\mathbf{X})$. Straightforward.
- The case $({}_?{=}_?[\mathbf{a}])$. Suppose $Pr = [a]r \ {}_?{=}_? \ [a]s, Pr'$ and $[a]r \ {}_?{=}_? \ [a]s, Pr' \Longrightarrow r \ {}_?{=}_? \ s, Pr'$ using $({}_?{=}_?[\mathbf{a}])$. Then:
  - Suppose $([a]r)\theta =_\alpha ([a]s)\theta$. By Definition 26, $[a](r\theta) =_\alpha [a](s\theta)$. By the rules in Definition 11, $r\theta =_\alpha s\theta$. The result follows.
  - Suppose $r\theta =_\alpha s\theta$. By the rules in Definition 26, $[a](r\theta) =_\alpha [a](s\theta)$. By Definition 26, $([a]r)\theta =_\alpha ([a]s)\theta$. The result follows.
- The case $({}_?{=}_?[\mathbf{b}])$. Suppose $Pr = [a]r \ {}_?{=}_? \ [b]s, Pr'$, $b \notin fa(r)$ and $Pr \Longrightarrow (b\ a) \cdot r \ {}_?{=}_? \ s, Pr'$ using $({}_?{=}_?[\mathbf{b}])$. Then:
  - Suppose $([a]r)\theta =_\alpha ([b]s)\theta$. By Definition 26, $[a](r\theta) =_\alpha [b](s\theta)$. By the rules in Definition 11, $(b\ a)\cdot(r\theta) =_\alpha s\theta$. By Lemma 28 and Theorem 24, $((b\ a)\cdot r)\theta =_\alpha s\theta$. The result follows.
  - Suppose $((b\ a)\cdot r)\theta =_\alpha s\theta$. By Lemma 28 and Theorem 24, $(b\ a)\cdot(r\theta) =_\alpha s\theta$. By Theorem 27, $b \notin fa(r\theta)$. Using $(=_\alpha[\mathbf{b}])$, $[a](r\theta) =_\alpha [b](s\theta)$. By Definition 26 $[a](r\theta) =_\alpha [b](s\theta)$. The result follows.

□

Definition 77. Define $fV(Pr) = \bigcup\{fV(r) \cup fV(s) \mid r \ {}_?{=}_? \ s \in Pr\}$.

Definition 78. Suppose $\mathcal{V}$ is a set of unknowns. Define $\theta|_{\mathcal{V}}$ by:[4]

$$\theta|_{\mathcal{V}}(X) \equiv \theta(X) \text{ if } X \in \mathcal{V} \qquad \theta|_{\mathcal{V}}(X) \equiv id \cdot X \text{ otherwise}$$

Definition 79. If Pr is a problem, define a **unification algorithm** by:

1. Rewrite $fV(Pr); Pr$ using the rules of Definition 75 as much as possible.
2. If we reduce to $\mathcal{V}'; \varnothing$, we succeed and return $\theta|_{\mathcal{V}}$ where θ is the functional composition of all the substitutions labelling rewrites (we take $\theta = id$ if there are none). Otherwise, we fail.

Lemma 80. *Suppose $\theta(X^S) =_\alpha \theta'(X^S)$ for all $X^S \in fV(Pr)$. Then $\theta \in Sol(Pr)$ if and only if $\theta' \in Sol(Pr)$.*

Proof. By Definition 72 it suffices to show $r\theta =_\alpha s\theta$ if and only if $r\theta' =_\alpha s\theta'$, for every $r \ {}_?{=}_? \ s \in Pr$. This is easy using Theorem 30 and the fact by construction (Definition 77) that $fV(r) \subseteq fV(Pr)$ and $fV(s) \subseteq fV(Pr)$. □

Definition 81. Write $\theta{-}X^S$ for the substitution such that:

$$(\theta{-}X^S)(X^S) \equiv id \cdot X^S \quad \text{and} \quad (\theta{-}X^S)(Y^T) \equiv \theta(Y^T) \text{ for all other } Y^T.$$

Theorem 82. *Suppose $X^S\theta =_\alpha s\theta$ and $X^S \notin fV(s)$. Then*

$$X^S\theta =_\alpha X^S([X^S{:=}s]\circ(\theta{-}X^S)) \quad \textit{and} \quad Y^T\theta =_\alpha Y^T([X^S{:=}s]\circ(\theta{-}X^S))$$

[4] We overload $|$, for technical convenience: $\pi|_S$ is partial and $\theta|_{\mathcal{V}}$ is total.

Proof. We reason as follows:

$$\begin{array}{lcll} X^S([X^S{:=}s]\circ(\theta{-}X^S)) & \equiv & s(\theta{-}X^S) & \text{Definition 26} \\ & \equiv & s\theta & X^S \notin fV(s),\ \text{Theorem 30} \\ & =_\alpha & X^S\theta & \text{Assumption} \\ \\ Y^T([X^S{:=}s]\circ(\theta{-}X^S)) & \equiv & Y^T(\theta{-}X^S) & \text{Definition 32} \\ & \equiv & Y^T\theta & \text{Definition 81} \end{array}$$

□

6.2. Simplification rewrites calculate principal solutions

Definition 83. Write $\theta_1 \leq \theta_2$ when there exists some θ' such that $X^S\theta_1 =_\alpha X^S(\theta_2 \circ \theta')$ always. Call $\leq$ the **instantiation ordering**.

Definition 84. A **principal** (or **most general**) solution to a problem Pr is a solution $\theta \in Sol(Pr)$ such that $\theta \leq \theta'$ for all other $\theta' \in Sol(Pr)$.

Our main results are Theorems 88 — the unification algorithm from Definition 79 calculates a solution — and 93 — the solution it calculates, is principal.

Lemma 85. *If $fV(Pr) \subseteq \mathcal{V}$ and $\mathcal{V}; Pr \Longrightarrow \mathcal{V}'; Pr'$ using a non-instantiating rule, then $fV(Pr') \subseteq \mathcal{V}$.*

Proof. As the empty set cannot be simplified, it must be that $Pr = r \ {}_?{=}_?\ s, Pr'$. Therefore, we perform case analysis on the simplification of $r \ {}_?{=}_?\ s$.

- The cases $({}_?{=}_?\mathbf{a})$, $({}_?{=}_?\mathbf{f})$ and $(=_\alpha\mathbf{X})$. Routine.
- The case $({}_?{=}_?[\mathbf{a}])$. Suppose $\mathcal{V}; [a]r \ {}_?{=}_?\ [a]s, Pr'$ and $fV([a]r \ {}_?{=}_?\ [a]s, Pr') \subseteq \mathcal{V}$, then $\mathcal{V}; [a]r \ {}_?{=}_?\ [a]s, Pr' \Longrightarrow \mathcal{V}; r \ {}_?{=}_?\ s, Pr'$ using $({}_?{=}_?[\mathbf{a}])$. By Definitions 10 and 77, $fV(r \ {}_?{=}_?\ s, Pr') \subseteq \mathcal{V}$. The result follows.
- The case $({}_?{=}_?[\mathbf{b}])$. Suppose $\mathcal{V}; [a]r \ {}_?{=}_?\ [b]s, Pr'$, $b \notin fa(r)$ with $fV([a]r \ {}_?{=}_?\ [b]s, Pr') \subseteq \mathcal{V}$, then $\mathcal{V}; [a]r \ {}_?{=}_?\ [b]s, Pr' \Longrightarrow \mathcal{V}; (b\ a)\cdot r \ {}_?{=}_?\ s, Pr'$ using $({}_?{=}_?[\mathbf{a}])$. By Definitions 10 and 77 and Lemma 17, $fV((b\ a)\cdot r) \subseteq \mathcal{V}$. The result follows.

□

Lemma 86. *If $fV(Pr) \subseteq \mathcal{V}$ and $\mathcal{V}; Pr \overset{\theta}{\Longrightarrow} \mathcal{V}'; Pr'\theta$ using an instantiating rule, then $fV(Pr'\theta) \subseteq \mathcal{V}$.*

Proof. There are two cases to consider:

- The cases (**I1**) and (**I2**). We handle the first case, the second is similar. Suppose $fV(\pi \cdot X^S \ {}_?{=}_?\ s, Pr') \subseteq \mathcal{V}$ and $\mathcal{V}; \pi \cdot X^S \ {}_?{=}_?\ s, Pr' \overset{[X^S:=\pi^{-1}\cdot s]}{\Longrightarrow} \mathcal{V}'; Pr'[X^S{:=}\pi^{-1}\cdot s]$ using (**I1**). By Definition 77 and Lemma 29, $fV(Pr'[X^S{:=}\pi^{-1}\cdot s]) \subseteq fV(Pr') \cup fV(\pi^{-1}\cdot s)$. The result follows.
- The case (**I3**). By Lemma 67.

□

Lemma 87. *If $X^S \in \mathcal{V}$ then $([X^S{:=}s]\circ\theta)|_\mathcal{V} = [X^S{:=}s]\circ(\theta|_\mathcal{V})$*

Proof. There are multiple cases to consider:

- The case X^S with $X^S \in \mathcal{V}$. We have:

$$\begin{array}{lcll} ([X^S{:=}s]\circ\theta)|_\mathcal{V}(X^S) & \equiv & ([X^S{:=}s]\circ\theta)(X^S) & \text{Definition 78, } X^S \in \mathcal{V} \\ & \equiv & s\theta & \text{Definition 32} \\ & \equiv & s\theta|_\mathcal{V} & \text{Definition 78} \end{array}$$

- The case Y^T with $Y^T \in \mathcal{V}$. We have:

$$\begin{array}{lcll} ([X^S{:=}s]\circ\theta)|_{\mathcal{V}}(Y^T) & \equiv & ([X^S{:=}s]\circ\theta)(Y^T) & \text{Definition 78, } Y^T \in \mathcal{V} \\ & \equiv & \theta(Y^T) & \text{Definition 32} \\ & \equiv & \theta|_{\mathcal{V}}(Y^T) & \text{Definition 78} \end{array}$$

- The case Y^T with $Y^T \notin \mathcal{V}$. Since $([X^S{:=}s]\circ\theta)|_{\mathcal{V}}(Y^T) \equiv id \cdot Y^T$ and $\theta|_{\mathcal{V}} \equiv id \cdot Y^T$.

□

Recall that $\theta|_{\mathcal{V}}$ is defined in Definition 78:

Theorem 88. *If $fV(Pr) \subseteq \mathcal{V}$ then $\mathcal{V}; Pr \stackrel{\theta}{\Longrightarrow}\!\!\!\!\!\!\ast\ \mathcal{V}'; \varnothing$ implies $\theta|_{\mathcal{V}} \in Sol(Pr)$.*

Proof. By induction on the length of the path in $\stackrel{\theta}{\Longrightarrow}\!\!\!\!\!\!\ast$.

- *Length* 0. Then $Pr = \varnothing$ and $\theta \equiv id$. The result follows.
- *Length* $k+1$. There are three cases:
  - *The non-instantiating case.* Suppose $\mathcal{V}; Pr \Longrightarrow \mathcal{V}; Pr'' \stackrel{\theta}{\Longrightarrow}\!\!\!\!\!\!\ast\ \mathcal{V}'; \varnothing$. By Lemma 85, $fV(Pr'') \subseteq \mathcal{V}$. By inductive hypothesis, $\theta \in Sol(Pr'')$. By Lemma 76, $\theta \in Sol(Pr)$. The result follows.
  - *The case of* (**I1**) *or* (**I2**). Suppose $\mathcal{V}; Pr \stackrel{\chi}{\Longrightarrow} \mathcal{V}; Pr\chi \stackrel{\theta'}{\Longrightarrow}\!\!\!\!\!\!\ast\ \mathcal{V}'; \varnothing$. By Lemma 86, $fV(Pr\chi) \subseteq \mathcal{V}$. By inductive hypothesis, $\theta'|_{\mathcal{V}} \in Sol(Pr\chi)$. By Lemma 87, $(\chi\circ\theta')|_{\mathcal{V}} = \chi\circ(\theta'|_{\mathcal{V}})$. By Lemma 73, $(\chi\circ\theta')|_{\mathcal{V}} \in Sol(Pr)$. The result follows.
  - *The case of* (**I3**). Suppose $\mathcal{V}; Pr \stackrel{\rho}{\Longrightarrow} \mathcal{V}'; Pr\rho \stackrel{\theta'}{\Longrightarrow}\!\!\!\!\!\!\ast\ \mathcal{V}''; \varnothing$. By Lemma 86, $fV(Pr\rho) \subseteq \mathcal{V}'$. By inductive hypothesis, $\theta'|_{\mathcal{V}'} \in Sol(Pr\rho)$. By Lemma 73, $\rho\circ(\theta'|_{\mathcal{V}'}) \in Sol(Pr)$. By Lemma 87, $\rho\circ(\theta'|_{\mathcal{V}'}) = (\rho\circ\theta')|_{\mathcal{V}'}$. By Lemma 80, $(\rho\circ\theta')|_{\mathcal{V}} \in Sol(Pr)$. The result follows.

□

We need some lemmas for Theorem 93:

Lemma 89. *If $\theta_1 \leq \theta_2$ then $\theta\circ\theta_1 \leq \theta\circ\theta_2$.*

Proof. By Definition 83, θ' exists such that $X^S\theta_1 =_\alpha X^S(\theta_2\circ\theta')$ always. Then:

$$\begin{array}{lcll} X^S(\theta\circ\theta_1) & \equiv & (X^S\theta)\theta_1 & \text{Theorem 33} \\ & =_\alpha & (X^S\theta)(\theta_2\circ\theta') & \text{Theorem 30} \\ & \equiv & X^S((\theta\circ\theta_2)\circ\theta') & \text{Theorem 33} \end{array}$$

The result follows. □

Lemma 90. *Suppose $X^S\theta_2 =_\alpha X^S\theta_2'$ always. Then $\theta_1 \leq \theta_2$ implies $\theta_1 \leq \theta_2'$.*

Proof. By a routine calculation using Definition 83 and using Theorem 24. □

Lemma 91. *If $\theta \in Sol(Pr)$ (Definition 72) then $\theta \in Sol(Pr_\sqsubseteq)$ (Definition 50).*

Proof. By a routine calculation, using Definitions 72 and 74, and Lemma 19. □

Lemma 92. *If $X^S \in \mathcal{V}$ then $(\theta|_{\mathcal{V}} - X^S) = (\theta - X^S)|_{\mathcal{V}}$.*

Proof. There are multiple cases to consider:

- The case X^S. Then $(\theta|_{\mathcal{V}} - X^S)(X^S) = id \cdot X^S$ and $(\theta - X^S)|_{\mathcal{V}}(X^S) = id \cdot X^S$. The result follows.

- The case Y^T with $Y^T \notin \mathcal{V}$. Then $(\theta|_{\mathcal{V}} - X^S)(Y^T) = id \cdot Y^T$ and $(\theta - X^S)|_{\mathcal{V}}(Y^T) = id \cdot Y^T$. The result follows.
- The case Y^T with $Y^T \in \mathcal{V}$. Then $(\theta|_{\mathcal{V}} - X^S)(Y^T) = \theta|_{\mathcal{V}}(Y^T)$ and $(\theta - X^S)|_{\mathcal{V}}(Y^T) = \theta(Y^T)$. As $\theta|_{\mathcal{V}}(Y^T) = \theta(Y^T)$ when $Y^T \in \mathcal{V}$. The result follows.

□

Theorem 93. *Suppose $fV(Pr) \subseteq \mathcal{V}$. If $\mathcal{V}; Pr \stackrel{\theta}{\Longrightarrow}{}^{*} \mathcal{V}'; \varnothing$ then $\theta|_{\mathcal{V}}$ is a principal solution to Pr (Definition 84).*

Proof. By Theorem 88, $\theta|_{\mathcal{V}} \in Sol(Pr)$. We prove $\theta|_{\mathcal{V}}$ is principal by induction on the path length of $\mathcal{V}; Pr \stackrel{\theta}{\Longrightarrow}{}^{*} \mathcal{V}'; \varnothing$.

- *Length* 0. So $Pr = \varnothing$ and $\theta = id|_{\mathcal{V}}$. By Definition 83, $id|_{\mathcal{V}} \leq \theta'|_{\mathcal{V}}$.
- *Length* $k+1$. We consider the rules in Definition 75.
  - The non-instantiating case. Suppose

$$\mathcal{V}; Pr \Longrightarrow \mathcal{V}; Pr' \stackrel{\theta}{\Longrightarrow}{}^{*} \mathcal{V}'; \varnothing$$

    where $\mathcal{V}; Pr \Longrightarrow \mathcal{V}; Pr'$ is a non-instantiating simplification rewrite. By inductive hypothesis, $\theta|_{\mathcal{V}}$ is a principal solution of Pr'. By Lemma 76, $\theta|_{\mathcal{V}}$ is a principal solution of Pr. The result follows.
  - The case (**I1**). Suppose $fa(s) \subseteq \pi \cdot S$ and $X^S \notin fV(s)$. Write $\chi = [X^S := \pi^{-1} \cdot s]$. Suppose $Pr = \pi \cdot X^S \ {}_?{=}_? \ s,\ Pr''$ so that

$$\mathcal{V}; \pi \cdot X^S \ {}_?{=}_? \ s,\ Pr'' \stackrel{\chi}{\Longrightarrow} \mathcal{V}; Pr''\chi \stackrel{\theta''}{\Longrightarrow}{}^{*} \mathcal{V}'; \varnothing.$$

    Further, suppose that $\theta'|_{\mathcal{V}} \in Sol(Pr)$.
    By Theorem 88, $\theta''|_{\mathcal{V}} \in Sol(Pr''\chi)$. By Lemma 86, $fV(Pr''\chi) \subseteq \mathcal{V}$. By Theorem 82 and Lemma 80, $\chi \circ (\theta'|_{\mathcal{V}} - X^S) \in Sol(Pr)$. By Lemma 92, $(\theta|_{\mathcal{V}} - X^S) = (\theta - X^S)|_{\mathcal{V}}$. By Lemma 73, $(\theta - X^S)|_{\mathcal{V}} \in Sol(Pr''\chi)$.
    By inductive hypothesis, $\theta''|_{\mathcal{V}} \leq (\theta' - X^S)|_{\mathcal{V}}$. By Lemma 89, $\chi \circ (\theta''|_{\mathcal{V}}) \leq \chi \circ (\theta' - X^S)|_{\mathcal{V}}$. By Lemma 87, $\chi \circ (\theta''|_{\mathcal{V}}) = (\chi \circ \theta'')|_{\mathcal{V}}$. By Lemma 92, $(\theta' - X^S)|_{\mathcal{V}} = \theta'|_{\mathcal{V}} - X^S$. By Theorem 82 and Lemma 90, $(\chi \circ \theta'')|_{\mathcal{V}} \leq \theta'|_{\mathcal{V}}$ as required.
  - The case (**I2**) is similar to the case of (**I1**).
  - The case (**I3**). Suppose $Pr_{\sqsubseteq}$ is consistent and non-trivial. Write $\rho = \rho^{\mathcal{V}}_{Pr_{\sqsubseteq}}$, so that

$$\mathcal{V}; Pr \stackrel{\rho}{\Longrightarrow} \mathcal{V}''; Pr\rho \stackrel{\theta''}{\Longrightarrow}{}^{*} \mathcal{V}'; \varnothing,$$

    and suppose that $\theta'|_{\mathcal{V}} \in Sol(Pr)$.
    By Theorem 88, $\theta''|_{\mathcal{V}''} \in Sol(Pr\rho)$. It is a fact that $\mathcal{V}'' = \mathcal{V} \cup \mathcal{V}'^{\mathcal{V}}_{Pr_{\sqsubseteq}}$, so $fV(Pr\rho) \subseteq \mathcal{V}''$. By Lemma 91, $\theta'|_{\mathcal{V}} \in Sol(Pr_{\sqsubseteq})$. By Theorem 70 and Lemma 80, $\rho \circ (\theta'|_{\mathcal{V}} - \rho) \in Sol(Pr)$. By Lemma 73, $\theta'|_{\mathcal{V}} - \rho \in Sol(Pr\rho)$.
    By inductive hypothesis, $\theta''|_{\mathcal{V}} \leq \theta'|_{\mathcal{V}} - \rho$. By Lemma 89, $\rho \circ \theta''|_{\mathcal{V}} \leq \rho \circ (\theta'|_{\mathcal{V}} - \rho)$. It is a fact that $\rho \circ (\theta''|_{\mathcal{V}}) = (\rho \circ \theta'')|_{\mathcal{V}}$. By Theorem 70 and Lemma 90, $(\rho \circ \theta'')|_{\mathcal{V}} \leq \theta'|_{\mathcal{V}}$ as required.

□

Theorem 94. *Given a problem Pr, if the algorithm of Definition 79 succeeds then it returns a principal solution; if it fails then there is no solution.*

Proof. If the algorithm succeeds we use Theorem 93. Otherwise, the algorithm generates an element of the form $\mathsf{f}(r_1,\ldots,r_n) \;?{=}?\; \mathsf{f}(r'_1,\ldots,r'_{n'})$ where $n \neq n'$, $\mathsf{f}(\ldots) \;?{=}?\; \mathsf{g}(\ldots)$, $\mathsf{f}(\ldots) \;?{=}?\; [a]s$, $\mathsf{f}(\ldots) \;?{=}?\; a$, $[a]r =_\alpha a$, $[a]r =_\alpha b$, $a \;?{=}?\; b$, a Pr such that $Pr_{\sqsubseteq}$ is inconsistent, or $\pi \cdot X^S \;?{=}?\; r$ or $r \;?{=}?\; \pi \cdot X^S$ where $X^S \in fV(r)$. It is clear that no solution to Pr exists. □

7. The λ-calculus

Definition 95. Let $X, Y, Z, \ldots$ range over distinct unknowns.

Define **λ-terms** by:

$$g, h, \ldots ::= a \mid X \mid \mathsf{f} \mid \lambda a.g \mid g'g$$

Here f ranges over term-formers, and a ranges over atoms (see Definition 1). g, h, k will range over λ-terms.

Definition 96. Define a **permutation action** by:

$$\pi{\cdot}a \equiv \pi(a) \quad \pi{\cdot}X \equiv X \quad \pi{\cdot}\mathsf{f} \equiv \mathsf{f} \quad \pi{\cdot}(\lambda a.g) \equiv \lambda\pi(a).(\pi{\cdot}g) \quad \pi{\cdot}(g'g) \equiv (\pi{\cdot}g')(\pi{\cdot}g)$$

Write $\pi \circ \pi'$ for the **composition** of permutations π and π', and id for the **identity** permutation on λ-terms.

Definition 97. Define **free atoms** by:

$$fa(a) = \{a\} \quad fa(X) = \varnothing \quad fa(\mathsf{f}) = \varnothing \quad fa(\lambda a.g) = fa(g)\backslash\{a\} \quad fa(g'g) = fa(g') \cup fa(g)$$

Definition 98. Let α-equivalence $=_\alpha$ be the least relation on λ-terms such that:

$$\frac{}{a =_\alpha a}\,(\lambda{=_\alpha}\mathbf{a}) \qquad \frac{g =_\alpha h}{\lambda a.g =_\alpha \lambda a.h}\,(\lambda{=_\alpha}\lambda\mathbf{aa}) \qquad \frac{(b\ a)\cdot g =_\alpha h \quad b \notin fa(g)}{\lambda a.g =_\alpha \lambda b.h}\,(\lambda{=_\alpha}\lambda\mathbf{ab})$$

$$\frac{}{\mathsf{f} =_\alpha \mathsf{f}}\,(\lambda{=_\alpha}\mathsf{f}) \qquad \frac{}{X =_\alpha X}\,(\lambda{=_\alpha}\mathbf{X}) \qquad \frac{g =_\alpha g' \quad h =_\alpha h'}{gh =_\alpha g'h'}\,(\lambda{=_\alpha}\mathbf{p})$$

It is not hard to prove that Definition 98 does indeed specify the usual α-equivalence relation on λ-terms. Our definition is designed to match the definition of α-equivalence on nominal terms (Definition 11). This makes later results easier to prove (for example Theorem 126).

Lemma 99 to Theorem 107 mirror similar results for permissive nominal terms.

Lemma 99. *If* $\pi|_{fa(g)} = \pi'|_{fa(g)}$ *then* $\pi \cdot g =_\alpha \pi' \cdot g$.

Proof. By induction on g.

- The cases a, f and X. Routine.
- The case $\lambda a.g$. We wish to show $\lambda\pi(a).\pi \cdot g =_\alpha \lambda\pi'(a).\pi' \cdot g$. There are two cases to consider:
  - The case $\pi(a) = \pi'(a)$. By inductive hypothesis.
  - The case $\pi(a) \neq \pi'(a)$. We wish to show $\lambda\pi(a).g =_\alpha \lambda\pi'(a).h$. Using $(\lambda{=_\alpha}\lambda\mathbf{a})$, this is equivalent to showing $(\pi'(a)\ \pi(a)) \cdot g =_\alpha h$ with $\pi'(a) \notin fa(g)$. If $\pi'(a) \notin fa(g)$ then $\pi(a) \notin fa(g)$, which holds by assumption. Therefore there is nothing to prove.
- The case $g'g$. Routine.

□

Lemma 100. $\pi \cdot (\pi' \cdot g) \equiv (\pi \circ \pi') \cdot g$

Proof. By induction on g.

- The case a. We have:

$$\begin{array}{rcll} \pi \cdot (\pi' \cdot a) & \equiv & \pi \cdot \pi'(a) & \text{Definition 96} \\ & \equiv & \pi(\pi'(a)) & \text{Definition 96} \\ & \equiv & (\pi \circ \pi') \cdot a & \text{Definition 96} \end{array}$$

- The cases X, f and $g'g$. These are routine.
- The case $\lambda a.g$. We have:

$$\begin{array}{rcll} \pi \cdot (\pi' \cdot \lambda a.g) & \equiv & \pi \cdot \lambda \pi'(a).(\pi' \cdot g) & \text{Definition 96} \\ & \equiv & \lambda \pi(\pi'(a)).(\pi \cdot (\pi' \cdot g)) & \text{Definition 96} \\ & \equiv & \lambda \pi(\pi'(a)).((\pi \circ \pi') \cdot g) & \text{Inductive hypothesis} \\ & \equiv & (\pi \circ \pi') \cdot \lambda a.g & \text{Definition 96} \end{array}$$

The result follows.

□

Lemma 101. $fa(\pi \cdot g) = \pi \cdot fa(g)$.

Proof. By induction on g.

- The case a. We have:

$$\begin{array}{rcll} \pi \cdot fa(a) & = & \pi \cdot \{a\} & \text{Definition 97} \\ & = & \{\pi(a)\} & \text{Definition 8} \\ & = & fa(\pi(a)) & \text{Definition 97} \\ & = & fa(\pi \cdot a) & \text{Definition 96} \end{array}$$

- The case X and f. These are straightforward.
- The case $g'g$. We have:

$$\begin{array}{rcll} \pi \cdot fa(g'g) & = & \pi \cdot (fa(g') \cup fa(g)) & \text{Definition 97} \\ & = & \pi \cdot fa(g') \cup \pi \cdot fa(g) & \text{Fact} \\ & = & fa(\pi \cdot g') \cup fa(\pi \cdot g) & \text{Inductive hypothesis} \\ & = & fa((\pi \cdot g')(\pi \cdot g)) & \text{Definition 97} \\ & = & fa(\pi \cdot g'g) & \text{Definition 96} \end{array}$$

The result follows.

- The case $\lambda a.g$. We have:

$$\begin{array}{rcll} \pi \cdot fa(\lambda a.g) & = & \pi \cdot (fa(g) \setminus \{a\}) & \text{Definition 97} \\ & = & \pi \cdot fa(g) \setminus \pi \cdot \{a\} & \text{Fact} \\ & = & fa(\pi \cdot g) \setminus \{\pi(a)\} & \text{Inductive hypothesis, Definition 8} \\ & = & fa(\lambda \pi(a).(\pi \cdot g)) & \text{Definition 97} \\ & = & fa(\pi \cdot \lambda a.g) & \text{Definition 97} \end{array}$$

The result follows.

□

Lemma 102. *$g =_\alpha h$ implies $\pi \cdot g =_\alpha \pi \cdot h$.*

Proof. By induction on the derivation of $g =_\alpha h$.

- The case $(\lambda =_\alpha \mathbf{a})$. Using $(\lambda =_\alpha \mathbf{a})$, $\pi(a) =_\alpha \pi(a)$.
- The case $(\lambda =_\alpha \mathbf{X})$ and $(\lambda =_\alpha \mathsf{f})$. Routine.
- The case $(\lambda =_\alpha \mathbf{p})$. By inductive hypothesis, $\pi \cdot g =_\alpha \pi \cdot g'$ and $\pi \cdot h =_\alpha \pi \cdot h'$. Using $(\lambda =_\alpha \mathbf{p})$, $(\pi \cdot g)(\pi \cdot h) =_\alpha (\pi \cdot g')(\pi \cdot h')$. By Definition 96, $(\pi \cdot g)(\pi \cdot h) \equiv \pi \cdot gh$. The result follows.
- The case $(\lambda =_\alpha \lambda\mathbf{aa})$. By inductive hypothesis, $\pi \cdot g =_\alpha \pi \cdot h$. Using $(\lambda =_\alpha \lambda\mathbf{aa})$, $\lambda\pi(a).(\pi \cdot g) =_\alpha \lambda\pi(a).(\pi \cdot h)$. By Definition 96, $\lambda\pi(a).\pi \cdot g \equiv \pi \cdot \lambda a.g$. The result follows.
- The case $(\lambda =_\alpha \lambda\mathbf{ab})$. By inductive hypothesis, $\pi \cdot ((b\ a) \cdot g) =_\alpha \pi \cdot h$. By Lemma 100, $\pi \cdot ((b\ a) \cdot g) \equiv (\pi \circ (b\ a)) \cdot g$. It is a fact that $\pi \circ (b\ a) = (\pi(b)\ \pi(a)) \circ \pi$. By Lemma 100, $(\pi(b)\ \pi(a)) \cdot (\pi \cdot g) =_\alpha \pi \cdot h$. By Lemma 101, $\pi(b) \notin fa(\pi \cdot g)$. Using $(\lambda =_\alpha \lambda\mathbf{ab})$, $\lambda\pi(a).(\pi \cdot g) =_\alpha \lambda\pi(b).(\pi \cdot h)$. The result follows by Definition 96.

□

Lemma 103. *If $g =_\alpha h$ then $fa(g) = fa(h)$.*

Proof. By induction on the derivation of $g =_\alpha h$.

- The cases $(\lambda =_\alpha \mathbf{a})$, $(\lambda =_\alpha \mathbf{X})$ and $(\lambda =_\alpha \mathsf{f})$. Straightforward.
- The case $(\lambda =_\alpha \mathbf{p})$. By inductive hypothesis, $fa(g') = fa(g)$ and $fa(h') = fa(h)$. As $fa(g'g) = fa(g') \cup fa(g)$, the result follows.
- The case $(\lambda =_\alpha \lambda\mathbf{aa})$. By inductive hypothesis, $fa(g) = fa(h)$, hence $fa(g) \setminus \{a\} = fa(h) \setminus \{a\}$. The result follows.
- The case $(\lambda =_\alpha \lambda\mathbf{ab})$. Suppose $\lambda a.g =_\alpha \lambda b.h$ using $(\lambda =_\alpha \lambda\mathbf{ab})$, with $b \notin fa(g)$. We aim to show $fa(\lambda a.g) = fa(\lambda b.h)$, that is, $fa(g) \setminus \{a\} = fa(h) \setminus \{b\}$. As $b \notin fa(g)$, $fa(g) \setminus \{a\} = (b\ a) \cdot fa(g) \setminus \{b\}$. By Lemma 101, $(b\ a) \cdot fa(g) \setminus \{b\} = fa((b\ a) \cdot g) \setminus \{b\}$. By inductive hypothesis, $fa((b\ a) \cdot g) = fa(s)$. The result follows.

□

Definition 104. Define a notion of **size** on λ-terms by:

$$size(a) = 0 \quad size(X) = 0 \quad size(\mathsf{f}) = 0 \quad size(g'g) = size(g') + size(g)$$
$$size(\lambda a.g) = 1 + size(g)$$

Lemma 105. *For every lambda-term g, the set $\{size(h) \mid h \text{ is a subterm of } g\}$ is well-ordered.*

Proof. Since the set $\{size(h) \mid h \text{ is a subterm of } g\}$ forms a subset of the natural numbers. □

Lemma 106. $size(g) = size(\pi \cdot g)$

Proof. By induction on g.

- The cases a, X and f. Straightforward.
- The case $g'g$. We have:

$size(g'g)$	$=$	$size(g') + size(g)$	Definition 104
	$=$	$size(\pi \cdot g') + size(\pi \cdot g)$	Inductive hypothesis
	$=$	$size((\pi \cdot g')(\pi \cdot g))$	Definition 104
	$=$	$size(\pi \cdot g'g)$	Definition 96

The result follows.

- The case $\lambda a.g$. We have:

$$\begin{array}{rcll} size(\lambda a.g) & = & 1 + size(g) & \text{Definition 104} \\ & = & 1 + size(\pi \cdot g) & \text{Inductive hypothesis} \\ & = & size(\lambda \pi(a).(\pi \cdot g)) & \text{Definition 104} \\ & = & size(\pi \cdot \lambda a.g) & \text{Definition 96} \end{array}$$

  The result follows.

□

Theorem 107. *$=_\alpha$ is transitive, reflexive, and symmetric.*

Proof. We handle the three cases separately.

- The reflexivity case, $g =_\alpha g$. We proceed by induction on g.
  - The case a, X and f. Routine.
  - The case $g'g$. By hypothesis, $g' =_\alpha g'$ and $g =_\alpha g$. Using $(\lambda =_\alpha \mathbf{p})$, $g'g =_\alpha g'g$. The result follows.
  - The case $\lambda a.g$. By hypothesis, $g =_\alpha g$. Using $(\lambda =_\alpha \lambda\mathbf{aa})$, $\lambda a.g =_\alpha \lambda a.g$. The result follows.
- The symmetry case, $g =_\alpha h$ implies $h =_\alpha g$. We proceed by induction on the derivation of $g =_\alpha h$.
  - The cases $(\lambda =_\alpha \mathbf{a})$, $(\lambda =_\alpha \mathbf{X})$ and $(\lambda =_\alpha \mathsf{f})$. Routine.
  - The case $(\lambda =_\alpha \mathbf{p})$. By inductive hypotheses, $g' =_\alpha g$ and $h' =_\alpha h$. Using $(\lambda =_\alpha \mathbf{p})$, $g'h' =_\alpha gh$. The result follows.
  - The case $(\lambda =_\alpha \lambda\mathbf{aa})$. By inductive hypothesis, $h =_\alpha g$. Using $(\lambda =_\alpha \lambda\mathbf{aa})$, $\lambda a.h =_\alpha \lambda a.g$. The result follows.
  - The case $(\lambda =_\alpha \lambda\mathbf{ab})$. Suppose $(b\ a) \cdot g =_\alpha h$ with $b \notin fa(g)$. By inductive hypothesis, $h =_\alpha (b\ a) \cdot h$. By Lemma 103, $b \notin fa(h)$. By Lemma 102, $(b\ a) \cdot h =_\alpha (b\ a) \cdot ((b\ a) \cdot g)$. By Lemma 100, $(b\ a) \cdot h =_\alpha ((b\ a) \circ (b\ a)) \cdot g$, therefore $(b\ a) \cdot h =_\alpha g$. By Lemma 101, $a \notin fa((b\ a) \cdot h)$. Using $(\lambda =_\alpha \lambda[\mathbf{b}])$, $\lambda b.h =_\alpha \lambda a.g$. The result follows.
- The transitivity case, $g =_\alpha h$ and $h =_\alpha i$ imply $g =_\alpha i$. Following Lemma 105, we proceed by induction on $size(g)$.
  - The cases a, X and f. Straightforward.
  - The case $g'g$. By the inductive hypotheses.
  - The case $\lambda a.g$. There are multiple cases to consider. We consider the most difficult, the case where all abstractions are named apart.
    Suppose $\lambda a.g =_\alpha \lambda b.h$ and $\lambda b.h =_\alpha \lambda c.k$. We aim to show $\lambda a.g =_\alpha \lambda c.k$. Suppose $(b\ a) \cdot g =_\alpha h$ and $(c\ b) \cdot h =_\alpha k$ with $b \notin fa(g)$ and $c \notin fa(h)$. By Lemma 100, $(c\ b) \cdot ((b\ a) \cdot g) =_\alpha (c\ b) \cdot h$. By Lemma 106, $(c\ b) \cdot ((b\ a) \cdot g) =_\alpha k$, equivalent to $(c\ a) \cdot g =_\alpha k$. By Lemma 103, $c \notin fa((b\ a) \cdot g)$. By Lemma 102, $c \notin fa(g)$. Using $(\lambda =_\alpha \lambda\mathbf{ab})$, the result follows.

□

Definition 108. Let β-equivalence $=_{\alpha\beta}$ be the least relation such that $(\lambda a.g)h =_{\alpha\beta} g[h/a]$ and closed under the rules of Definition 98.

Definition 109. Call a function σ from unknowns to λ-terms a (λ-calculus) **substitution**. σ will range over substitutions (and later so will ρ; Definition 142).

Definition 110. Define the **capture-avoiding substitution** action $g\sigma$ on λ-terms by:

$$a\sigma{\equiv}a \quad X\sigma{\equiv}\sigma(X) \quad \mathsf{f}\sigma{\equiv}\mathsf{f} \quad (g'g)\sigma{\equiv}(g'\sigma)(g\sigma) \quad (\lambda a.g)\sigma{\equiv}\lambda a.(g\sigma) \quad (a \notin fa(g\sigma))$$
$$(\lambda a.g)\sigma \equiv \lambda b.((b\ a)\cdot g\sigma) \quad (a \in fa(g\sigma),\ b \text{ fresh})$$

In the final clause, 'b fresh' denotes a fixed but arbitrary choice of b such that $b \notin fa(g\sigma) \cup fa(g)$.

Lemma 111. $\pi \cdot (g\sigma) =_\alpha (\pi \cdot g)\sigma$

Proof. By induction on $size(g)$.

- The cases a, f and X. Straightforward.
- The case $g'g$. We have:

$$\begin{array}{lcll}
(\pi \cdot g'g)\sigma & \equiv & ((\pi \cdot g')(\pi \cdot g))\sigma & \text{Definition 96} \\
& \equiv & ((\pi \cdot g')\sigma)((\pi \cdot g)\sigma) & \text{Definition 110} \\
& \equiv & (\pi \cdot g'\sigma)(\pi \cdot (g\sigma)) & \text{Inductive hypothesis} \\
& \equiv & \pi \cdot (g'\sigma)(g\sigma) & \text{Definition 96} \\
& \equiv & \pi \cdot ((g'g)\sigma) & \text{Definition 110}
\end{array}$$

  The result follows.
- The case $\lambda a.g$ with $a, \pi(a) \notin fa(rng(\sigma))$. We have:

$$\begin{array}{lcll}
(\pi \cdot \lambda a.g)\sigma & \equiv & (\lambda \pi(a).(\pi \cdot g))\sigma & \text{Definition 96} \\
& \equiv & \lambda \pi(a).((\pi \cdot g)\sigma) & \text{Definition 110} \\
& \equiv & \lambda \pi(a).(\pi \cdot (g\sigma)) & \text{Inductive hypothesis} \\
& \equiv & \pi \cdot \lambda a.(g\sigma) & \text{Definition 96} \\
& \equiv & \pi \cdot ((\lambda a.g)\sigma) & \text{Definition 110}
\end{array}$$

  The result follows.
- The case $\lambda a.g$ with $a \in fa(rng(\sigma))$ or $\pi(a) \in fa(rng(\sigma))$. We have:

$$\begin{array}{lcll}
(\pi \cdot \lambda a.g)\sigma & =_\alpha & (\pi \cdot \lambda b.((b\ a)\cdot g))\sigma & b \text{ fresh} \\
& \equiv & (\lambda \pi(b).(\pi \cdot ((b\ a)\cdot g)))\sigma & \text{Definition 96} \\
& \equiv & (\lambda \pi(b).((\pi \circ (b\ a))\cdot g))\sigma & \text{Lemma 100} \\
& \equiv & \lambda \pi(b).((\pi \circ (b\ a))\cdot g)\sigma) & \text{Definition 110} \\
& =_\alpha & \lambda \pi(b).((\pi \circ (b\ a))\cdot g\sigma) & \text{Lemma 106, Inductive hypothesis} \\
& \equiv & \lambda \pi(b).(\pi \cdot ((b\ a)\cdot (g\sigma))) & \text{Lemma 100} \\
& \equiv & \pi \cdot \lambda b.((b\ a)\cdot (g\sigma)) & \text{Definition 96} \\
& \equiv & \pi \cdot \lambda b.(((b\ a)\cdot g)\sigma) & \text{Lemma 106, Inductive hypothesis} \\
& \equiv & \pi \cdot ((\lambda b.(b\ a)\cdot g)\sigma) & \text{Definition 110} \\
& =_\alpha & \pi \cdot (\lambda a.g)\sigma & b \text{ fresh}
\end{array}$$

  The result follows.

□

Definition 112 is an analogue of the substitution action on permissive nominal terms from Definition 32:

Definition 112. Define **composition** $\sigma \circ \sigma'$ by: $(\sigma \circ \sigma')(X) \equiv (\sigma(X))\sigma'$.

Lemma 113. $g\sigma\sigma' =_\alpha g(\sigma \circ \sigma')$

Proof. By induction on $size(g)$.

- The case a. Since $a\sigma \equiv a$.
- The case X. By Definition 112.
- The case f. Since $\pi \cdot \mathsf{f} \equiv \mathsf{f}$ and $\mathsf{f}\sigma \equiv \mathsf{f}$.
- The case $g'g$. We have:

$$\begin{array}{rcll} (g'g)\sigma\sigma' & \equiv & (g'\sigma\sigma')(g\sigma\sigma') & \text{Definition 110} \\ & \equiv & (g'(\sigma\circ\sigma'))(g(\sigma\circ\sigma')) & \text{Inductive hypothesis} \\ & \equiv & (g'g)(\sigma\circ\sigma') & \text{Definition 110} \end{array}$$

  The result follows.
- The case $\lambda a.g$ with $a \notin fa(rng(\sigma)) \cup fa(rng(\sigma'))$. We have:

$$\begin{array}{rcll} (\lambda a.g)\sigma\sigma' & =_\alpha & (\lambda b.(b\ a)\cdot g)\sigma\sigma' & b \text{ fresh} \\ & \equiv & \lambda b.((b\ a)\cdot g)\sigma\sigma' & \text{Definition 110} \\ & =_\alpha & \lambda b.((b\ a)\cdot g)(\sigma\circ\sigma') & \text{Lemma 106, Inductive hypothesis} \\ & \equiv & (\lambda b.((b\ a)\cdot g)(\sigma\circ\sigma') & \text{Definition 110} \\ & =_\alpha & (\lambda a.g)(\sigma\circ\sigma') & b \text{ fresh} \end{array}$$

  The result follows.

□

We define unification problems as usual and write '$g \stackrel{?}{=}_? h$' for an equality considered as part of a unification problem. σ solves a problem when $g\sigma =_{\alpha\beta} h\sigma$ for every $g \stackrel{?}{=}_? h$ in the problem, as usual.

We conclude with definions of *pattern* and *pattern substitution* [22, 21]. Recall that, unlike [19], we work in an untyped λ-calculus.

Definition 114. Let ϕ map each unknown X to a natural number which we call its **arity**. Define **ϕ-patterns**, a subset of λ-terms, by:

$$q, r, \ldots ::= a \mid Xa_1 \ldots a_{\phi(X)} \mid \mathsf{f}q_1 \ldots q_n \mid \lambda a.q$$

Call q a **pattern** when it is a ϕ-pattern for some ϕ. $q, r, \ldots$ will range over patterns.

Call σ a **ϕ-pattern substitution** when every $\sigma(X)$ is a ϕ-pattern. Call σ a **pattern substitution** when σ is a ϕ-pattern substitution for some ϕ.

So g is a pattern when every X in g occurs as $Xa_1 \ldots a_{\phi(X)}$, for some $\phi(X)$.

8. Translating nominal terms into the λ-calculus

8.1. The translation $[\![\text{-}]\!]^D$, and its soundness

Definition 115. Call a finite list of distinct atoms a **vector**. C, D range over vectors. Write $[a_1, \ldots, a_n]$ for the vector containing a_1, ..., a_n in that order.

Definition 116. Suppose $A \subseteq \mathbb{A}$. Write $C \cap A$ for the vector of atoms in C that occur in A, in order; thus $[a_1, a_2, a_3] \cap \{a_1, a_3, a_5\} = [a_1, a_3]$. Write $C \subseteq A$ when every atom in C is in A. Write $A \subseteq C$ when every atom in A is in C.

Definition 117. Translate a nominal term r to a λ-term $[\![r]\!]^D$ by:

$$[\![a]\!]^D \equiv a \quad [\![\pi \cdot X^S]\!]^D \equiv X^S\pi(d_1)\ldots\pi(d_n) \quad ([d_1, \ldots, d_n] = D \cap S)$$
$$[\![[a]r]\!]^D \equiv \lambda a.[\![r]\!]^D \quad [\![\mathsf{f}(r_1, \ldots, r_n)]\!]^D \equiv \mathsf{f}[\![r_1]\!]^D \ldots [\![r_n]\!]^D$$

Lemma 118. $[\![\pi \cdot r]\!]^D \equiv \pi \cdot [\![r]\!]^D$

Proof. By induction on r.

- The case a. We have:

$$\begin{array}{lcll} [\![\pi \cdot a]\!]^D & \equiv & [\![\pi(a)]\!]^D & \text{Definition 7} \\ & \equiv & \pi \cdot a & \text{Definition 117} \\ & \equiv & \pi \cdot [\![a]\!]^D & \text{Definition 96} \end{array}$$

  The result follows.

- The case $\mathsf{f}(r_1, \ldots, r_n)$. We have:

$$\begin{array}{lcll} [\![\pi \cdot \mathsf{f}(r_1, \ldots, r_n)]\!]^D & \equiv & [\![\mathsf{f}(\pi \cdot r_1, \ldots, \pi \cdot r_n)]\!]^D & \text{Definition 7} \\ & \equiv & \mathsf{f}[\![\pi \cdot r_1]\!]^D \ldots [\![\pi \cdot r_n]\!]^D & \text{Definition 117} \\ & \equiv & \mathsf{f}\pi \cdot [\![r_1]\!]^D \ldots \pi \cdot [\![r_n]\!]^D & \text{Inductive hypothesis} \\ & \equiv & \pi \cdot \mathsf{f}[\![r_1]\!]^D \ldots [\![r_n]\!]^D & \text{Definition 96} \\ & \equiv & \pi \cdot [\![\mathsf{f}(r_1, \ldots, r_n)]\!]^D & \text{Definition 117} \end{array}$$

  The result follows.

- The case $[a]r$. We have:

$$\begin{array}{lcll} \pi \cdot [\![[a]r]\!]^D & \equiv & \pi \cdot \lambda a.[\![r]\!]^D & \text{Definition 117} \\ & \equiv & \lambda\pi(a).(\pi \cdot [\![r]\!]^D) & \text{Definition 96} \\ & \equiv & [\![[\pi(a)](\pi \cdot r)]\!]^D & \text{Definition 117} \\ & \equiv & [\![\pi \cdot [a]r]\!]^D & \text{Definition 7} \end{array}$$

  The result follows.

- The case $\pi' \cdot X^S$.

$$\begin{array}{lcll} [\![\pi \cdot (\pi' \cdot X^S)]\!]^D & \equiv & [\![(\pi \circ \pi') \cdot X^S]\!]^D & \text{Definition 7} \\ & \equiv & X^S(\pi \circ \pi')(c_1) \ldots (\pi \circ \pi')(c_n) & \text{Definition 117} \\ & \equiv & \pi \cdot (X^S \pi'(c_1) \ldots \pi'(c_n)) & \text{Fact} \\ & \equiv & \pi \cdot [\![\pi' \cdot X^S]\!]^D & \text{Definition 117} \end{array}$$

  The result follows.

□

Lemma 119 is useful for the proof of Theorem 120:

Lemma 119. $fa([\![r]\!]^D) \subseteq fa(r)$.

Proof. By induction on r.

- The cases a and $\mathsf{f}(r_1, \ldots, r_n)$. Routine.
- The case $\mathsf{f}(r_1, \ldots, r_n)$. We have:

$$\begin{array}{lcll} fa([\![\mathsf{f}(r_1, \ldots, r_n)]\!]^D) & = & fa(\mathsf{f}[\![r_1]\!]^D \ldots [\![r_n]\!]^D) & \text{Definition 117} \\ & = & fa([\![r_1]\!]^D) \cup \ldots \cup fa([\![r_n]\!]^D) & \text{Definition 97} \\ & \subseteq & fa(r_1) \cup \ldots \cup fa(r_n) & \text{Inductive hypothesis} \\ & = & fa(\mathsf{f}(r_1, \ldots, r_n)) & \text{Definition 9} \end{array}$$

  The result follows.

- The case $[a]r$. We have:

$$\begin{array}{rcll} fa([\![[a]r]\!]^D) & = & fa(\lambda a.[\![r]\!]^D) & \text{Definition 117} \\ & = & fa([\![r]\!]^D) \setminus \{a\} & \text{Definition 97} \\ & \subseteq & fa(r) \setminus \{a\} & \text{Inductive hypothesis} \\ & = & fa([a]r) & \text{Definition 9} \end{array}$$

  The result follows.
- The case $\pi \cdot X^S$. We have:

$$\begin{array}{rcll} fa([\![\pi \cdot X^S]\!]^D) & = & fa(X^S\pi(d_1)\ldots\pi(d_n)) & \text{Definition 117} \\ & = & fa(\pi(d_1)) \cup \ldots \cup fa(\pi(d_n)) & \text{Definition 97} \\ & = & \pi \cdot (fa(d_1) \cup \ldots \cup fa(d_n)) & \text{Fact} \\ & \subseteq & \pi \cdot fa(X^S) & \text{Definition 9} \\ & = & fa(\pi \cdot X^S) & \text{Lemma 16} \end{array}$$

  The result follows.

□

Theorem 120 (Soundness)**.** *If* $r =_\alpha s$ *then* $[\![r]\!]^D =_\alpha [\![s]\!]^D$.

Proof. By induction on the size of r. We reason by cases on the last rule in the derivation of $r =_\alpha s$:

- The cases $(=_\alpha \mathbf{a})$, $(=_\alpha \mathsf{f})$ and $(=_\alpha []\mathbf{aa})$. Straightforward.
- The case $(=_\alpha \mathbf{X})$. There are two cases to consider:
  - The case $D \cap S = []$. Then $[\![\pi \cdot X^S]\!]^D = [\![\pi' \cdot X^S]\!]^D = X^S$. Using $(\lambda{=_\alpha}\mathbf{X})$, the result follows.
  - The case $D \cap S = [d_1, \ldots, d_n]$ and $n \geq 1$. By assumption, $\pi|_{\delta(X)} = \pi'|_{\delta(X)}$. Then $\pi(d_i) = \pi'(d_i)$ for $1 \leq i \leq n$ and $[\![\pi \cdot X^S]\!]^D \equiv [\![\pi' \cdot X^S]\!]^D \equiv X^S\pi(d_1)\ldots\pi(d_n)$. Using $(\lambda{=_\alpha}\mathbf{p})$, $(\lambda{=_\alpha}\mathbf{X})$, and $(\lambda{=_\alpha}\mathbf{a})$, the result follows.
- The case $(=_\alpha []\mathbf{ab})$. By assumption, $(b\ a) \cdot r =_\alpha s$ and $b \notin fa(r)$. Choose fresh c, so $c \notin fa(r) \cup fa(s)$. By Lemma 18, $(c\ a) \cdot r =_\alpha (c\ b) \cdot s$. By inductive hypothesis, $[\![(c\ a) \cdot r]\!]^D =_\alpha [\![(c\ b) \cdot s]\!]^D$. Using $(\lambda{=_\alpha}\lambda\mathbf{aa})$, $\lambda c.[\![(c\ a) \cdot r]\!]^D =_\alpha \lambda c.[\![(c\ a) \cdot s]\!]^D$. By Lemma 118, $\lambda c.((c\ a) \cdot [\![r]\!]^D) =_\alpha \lambda c.((c\ b) \cdot [\![s]\!]^D)$. By Lemma 119, $c \notin fa([\![r]\!]^D) \cup fa([\![s]\!]^D)$. Using $(\lambda{=_\alpha}\lambda\mathbf{ab})$, $\lambda c.((c\ a) \cdot [\![r]\!]^D) =_\alpha \lambda a.[\![r]\!]^D$ and $\lambda c.((c\ b) \cdot [\![s]\!]^D) =_\alpha \lambda b.[\![s]\!]^D$. By Theorem 107, $\lambda a.[\![r]\!]^D =_\alpha \lambda b.[\![s]\!]^D$. By Definition 117, $[\![[a]r]\!]^D =_\alpha [\![[b]s]\!]^D$. The result follows.

□

8.2. Capturable atoms; injectivity and minimality

The main results of this subsection are Theorems 126 and 128, and also Definition 121.

$[\![r]\!]^D$ (Definition 117) is parameterised by a vector D. Levy and Villaret introduced a similar translation [19, Definition 2]; they used *all* the atoms in r. We now show that the smaller set of *capturable* atoms in r (Definition 121) is consistent with injectivity of the translation (Theorem 126), and that it is minimal (Theorem 128).

Definition 121. Define the **capturable atoms** of a term (with respect to a set of atoms) $capt_A(r)$ inductively by:

$$capt_A(a) = \varnothing \quad capt_A(\pi \cdot X^S) = (dom(\pi) \cup A) \cap S \quad capt_A([a]r) = capt_{A \cup \{a\}}(r)$$

$$capt_A(\mathsf{f}(r_1, \ldots, r_n)) = \bigcup_{1 \leq i \leq n} capt_A(r_i)$$

Write $capt_{\varnothing}(r)$ as $capt(r)$.

For instance, if $S = (comb \cup \{a\}) \setminus \{b\}$, then $capt([a][b]X^S) = \{a\}$ and $capt((b\ a) \cdot X^S) = \{a\}$. We now prove that $capt$ respects α-equivalence:

Lemma 122. *If $a \notin fa(r)$ then $capt_A(r) = capt_{A \cup \{a\}}(r)$.*

Proof. By induction on r.

- The case b. Routine.
- The case $\mathsf{f}(r_1, \ldots, r_n)$. If $a \notin fa(\mathsf{f}(r_1, \ldots, r_n))$, then by Definition 9, $a \notin fa(r_i)$ for $1 \leq i \leq n$. We have:

$$\begin{array}{lcll} capt_A(\mathsf{f}(r_1, \ldots, r_n)) & = & \bigcup_{1 \leq i \leq n} capt_A(r_i) & \text{Definition 121} \\ & = & \bigcup_{1 \leq i \leq n} capt_{A \cup \{a\}}(r_i) & \text{Inductive hypothesis} \\ & = & capt_{A \cup \{a\}}(\mathsf{f}(r_1, \ldots, r_n)) & \text{Definition 121} \end{array}$$

  The result follows.
- The case $[b]r$. If $a \notin fa([b]r)$, then by Definition 9, $a \notin fa(r)$. We have:

$$\begin{array}{lcll} capt_A([b]r) & = & capt_{A \cup \{b\}}(r) & \text{Definition 121} \\ & = & capt_{A \cup \{b\} \cup \{a\}}(r) & \text{Inductive hypothesis} \\ & = & capt_{A \cup \{a\}}([b]r) & \text{Definition 121} \end{array}$$

  The result follows.
- The case $\pi \cdot X^S$. If $a \notin fa(\pi \cdot X^S)$, then $a \notin \pi \cdot S$. By Definition 121, $capt_A(\pi \cdot X^S) = (dom(\pi) \cup A) \cap S$. Then, $capt_{A \cup \{a\}}(\pi \cdot X^S) = (dom(\pi) \cup A \cup \{a\}) \cap S$. If $\pi(a) = a$, then $a \notin S$. If $\pi(a) \neq a$, then $a \in dom(\pi)$. The result follows.

□

Lemma 123. *If $dom(\pi) \subseteq A$ then $capt_A(\pi \cdot r) = capt_A(r)$.*

Proof. By induction on r.

- The cases a and $\mathsf{f}(r_1, \ldots, r_n)$. Straightforward.
- The case a. Since $capt_A(\pi(a)) = \varnothing = capt_A(a)$.
- The case $\mathsf{f}(r_1, \ldots, r_n)$. We have:

$$\begin{array}{lcll} capt_A(\pi \cdot \mathsf{f}(r_1, \ldots, r_n)) & = & capt_A(\mathsf{f}(\pi \cdot r_1, \ldots, \pi \cdot r_n)) & \text{Definition 7} \\ & = & capt_A(\pi \cdot r_1) \cup \ldots \cup capt_A(\pi \cdot r_n) & \text{Definition 121} \\ & = & capt_A(r_1) \cup \ldots \cup capt_A(r_n) & \text{Inductive hypotheses} \\ & = & capt_A(\mathsf{f}(r_1, \ldots, r_n)) & \text{Definition 121} \end{array}$$

  The result follows.
- The case $[a]r$. We have:

$$\begin{array}{lcll} capt_A(\pi \cdot [a]r) & = & capt_A([\pi(a)](\pi \cdot r)) & \text{Definition 7} \\ & = & capt_{A \cup \{\pi(a)\}}(\pi \cdot r) & \text{Definition 121} \end{array}$$

  There are two cases to consider:
  - The case $\pi(a) = a$. Then:

$$\begin{array}{lcll} capt_{A \cup \{\pi(a)\}}(\pi \cdot r) & = & capt_{A \cup \{a\}}(\pi \cdot r) & \text{Assumption} \\ & = & capt_{A \cup \{a\}}(r) & \text{Inductive hypothesis} \\ & = & capt_A([a]r) & \text{Definition 121} \end{array}$$

    The result follows.

- The case $\pi(a) \neq a$. Then:

$$\begin{array}{rcll} capt_{A\cup\{\pi(a)\}}(\pi\cdot r) & = & capt_A(\pi\cdot r) & \text{Assumption, } \pi(a)\neq a \\ & = & capt_A(r) & \text{Inductive hypothesis} \\ & = & capt_A([a]r) & \text{Definition 121} \end{array}$$

  The result follows.
- The case $\pi'\cdot X^S$. We have:

$$\begin{array}{rcll} capt(\pi\cdot(\pi'\cdot X^S)) & = & capt_A((\pi\circ\pi')\cdot X^S) & \text{Lemma 15} \\ & = & (dom(\pi\circ\pi')\cup A)\cap S & \text{Definition 121} \\ & = & (dom(\pi)\cup dom(\pi')\cup A)\cap S & \text{Fact} \\ & = & (dom(\pi')\cup A)\cap S & \text{Assumption} \\ & = & capt_A(\pi'\cdot X^S) & \text{Definition 121} \end{array}$$

  The result follows.

□

Corollary 124. *If $a\notin fa(r)$ then $capt_A([b]r) = capt_A([a](b\ a)\cdot r)$.*

Proof. We have:

$$\begin{array}{rcll} capt_A([b]r) & = & capt_{A\cup\{b\}}(r) & \text{Definition 121} \\ & = & capt_{A\cup\{a,b\}}(r) & \text{Lemma 122, } a\notin fa(r) \\ & = & capt_{A\cup\{a,b\}}((b\ a)\cdot r) & \text{Lemma 123} \\ & = & capt_{A\cup\{a\}}((b\ a)\cdot r) & \text{Lemmas 122 and 16} \\ & = & capt_A([a](b\ a)\cdot r) & \text{Definition 121} \end{array}$$

The result follows. □

Lemma 125. *If $r =_\alpha s$ then $capt_A(r) = capt_A(s)$.*

Proof. By induction on the derivation of $r =_\alpha s$.

- The case $(=_\alpha \mathbf{aa})$. Straightforward.
- The case $(=_\alpha \mathsf{f})$. Suppose $r_1 =_\alpha s_1 \ldots r_n =_\alpha s_n$. By hypothesis, $capt_A(r_1) = capt_A s_1 \ldots capt_A(r_n) = capt_A s_n$. Using $(=_\alpha \mathsf{f})$, $\mathsf{f}(r_1,\ldots r_n) =_\alpha \mathsf{f}(s_1,\ldots,s_n)$. Then, $capt_A(\mathsf{f}(r_1,\ldots,r_n)) = capt_A(r_1)\cup\ldots\cup capt_A(r_n)$. The result follows.
- The case of $(=_\alpha[\mathbf{a}])$. Suppose $r =_\alpha s$. By Definition 121, $capt_A([a]r) = capt_{A\cup\{a\}}(r)$, similarly for s. By inductive hypothesis, $capt_{A\cup\{a\}}(r) = capt_{A\cup\{a\}}(s)$. The result follows.
- The case of $(=_\alpha[\mathbf{b}])$. Suppose $b\notin fa(r)$, $(b\ a)\cdot r =_\alpha s$, and $s\equiv[b](b\ a)\cdot r$. The result follows by Corollary 124.
- The case of $(=_\alpha \mathbf{X})$. Suppose $\pi|_S = \pi'|_S$. Then, $dom(\pi)\cap S = dom(\pi')\cap S$. The result follows.

□

Theorem 126 (Injectivity)**.** *Let D be a vector. Let r and s be nominal terms and let $A, B\subseteq\mathbb{A}$ be finite. Suppose $capt_A(r)\cup capt_B(s)\subseteq D$. Then*

$$[\![r]\!]^D =_\alpha [\![s]\!]^D \quad \textit{implies} \quad r =_\alpha s.$$

As a corollary, if $capt(r)\cup capt(s)\subseteq D$ and $[\![r]\!]^D =_\alpha [\![s]\!]^D$ then $r =_\alpha s$ and $capt_A(r) = capt_A(s)$ for all A.

Proof. For the first part, we work by induction on the size of r, reasoning by cases on the last rule in the derivation of $[\![r]\!]^D =_\alpha [\![s]\!]^D$:

- The cases ($\lambda{=_\alpha}\mathbf{a}$) and ($\lambda{=_\alpha}\lambda\mathbf{aa}$). Routine.
- The case ($\lambda{=_\alpha}\lambda\mathbf{ab}$). Suppose $(b\ a)\cdot[\![r]\!]^D =_\alpha [\![s]\!]^D$, $b \notin fa([\![r]\!]^D)$ and $capt_A([a]r) \cup capt_B([b]s) \subseteq D$.
  Choose fresh c, so $c \notin fa(r) \cup fa(s)$ and $c \notin fa([\![r]\!]^D) \cup fa([\![s]\!]^D)$. By Lemma 99, $(c\ a)\cdot[\![r]\!]^D =_\alpha (c\ b)\cdot[\![s]\!]^D$. By Lemma 118, $[\![(c\ a)\cdot r]\!]^D =_\alpha [\![(c\ b)\cdot s]\!]^D$. By Corollary 124 and Definition 121, $capt_{A\cup\{c\}}((c\ a)\cdot r) \cup capt_{B\cup\{c\}}((c\ b)\cdot s) \subseteq D$. By hypothesis, $(c\ a)\cdot r =_\alpha (c\ b)\cdot s$. Using ($=_\alpha[]\mathbf{ab}$), and by Theorem 107, $[a]r =_\alpha [b]s$. The result follows.
- The case ($\lambda{=_\alpha}\mathbf{app}$). By Definition 117, there are two cases:
  - The case $\mathsf{f}r_1 \ldots r_n$ and $\mathsf{f}s_1 \ldots s_n$ and $[\![r_i]\!]^D =_\alpha [\![s_i]\!]^D$ for $1 \le i \le n$. By hypothesis, $r_i =_\alpha s_i$ for $1 \le i \le n$. Using ($=_\alpha\mathsf{f}$), the result follows.
  - The case $X^S\pi(d_1)\ldots\pi(d_n)$ and $X^S\pi'(d_1)\ldots\pi'(d_n)$ with $[d_1,\ldots,d_n] = D \cap S$ and $\pi(d_i) =_\alpha \pi'(d_i)$ for $1 \le i \le n$.
    Then $\pi|_{D\cap S} = \pi'|_{D\cap S}$ follows immediately. By assumption, $capt(\pi\cdot X^S) \subseteq D$. By definition, $\pi|_S = \pi'|_S$. Using ($=_\alpha\mathbf{X}$), the result follows.
- The case ($\lambda{=_\alpha}\mathbf{X}$). From the form of the translation, $r \equiv \pi\cdot X^S$ and $s \equiv \pi'\cdot X^S$ and $dom(\pi)\cap S = \varnothing = dom(\pi')\cap S$. Using ($=_\alpha\mathbf{X}$), the result follows.

The corollary follows from the first part and Lemma 125. □

Lemma 127. *$a \in capt_A(r)$ implies $X^S \in fV(r)$ exists such that $a \in S$.*

Proof. By induction on r.

- The cases a and b. Since $capt_A(a) = \varnothing$.
- The case $\mathsf{f}(r_1,\ldots,r_n)$. Suppose $a \in capt_A(\mathsf{f}(r_1,\ldots,r_n))$, with $a \in capt_A(r_i)$ for some i with $1 \le i \le n$. By hypothesis, $X^S \in fV(r_i)$ with $a \in S$. Then $fV(\mathsf{f}(r_1,\ldots,r_n)) = fV(r_1)\cup\ldots\cup fV(r_n)$. The result follows.
- The cases $[a]r$ and $[b]s$. We handle the first case, the second is similar.
  Suppose $a \in capt_A([a]r)$. Then $a \in capt_{A\cup\{a\}}(r)$. By hypothesis, $X^S \in fV(r)$ exists with $a \in S$. As $fV([a]r) = fV(r)$, the result follows.
- The $\pi\cdot X^S$. By Definition 121.

□

Theorem 128 (Minimality)**.** *If $capt(r) \not\subseteq D$ then there exists some s such that $r \neq_\alpha s$ and $[\![r]\!]^D =_\alpha [\![s]\!]^D$.*

Proof. Suppose $a \in capt(r)$ and $a \notin D$. By Lemma 127, $X^S \in fV(r)$ exists with $a \in S$. Choose fresh c, so $c \notin fa(r)\cup D$, and take $s \equiv r[X^S{:=}(c\ a)\cdot X^S]$. It is a fact that $X^S \neq_\alpha (c\ a)\cdot X^S$ whilst $[\![X^S]\!]^D =_\alpha [\![(c\ a)\cdot X^S]\!]^D$. An easy calculation shows $r \neq_\alpha r[X^S{:=}(c\ a)\cdot X^S]$ and $[\![r]\!]^D =_\alpha [\![r[X^S{:=}(c\ a)\cdot X^S]]\!]^D$. □

9. Translating substitutions; relating solutions of nominal and pattern unification problems

9.1. Translating substitutions

The main result of this subsection is Theorem 131.

We extend the translation to substitutions, to then prove that if a substitution solves a nominal unification problem, then its translation solves the translation of the problem. This raises a difficulty: θ may solve Pr but in substituting it may introduce

new capturable atoms (consider $\theta = [X^S{:=}[c]Z^S]$ solving $\{X^S \ ?{=}? \ X^S\}$, where $c \in S$). This motivates introducing another vector E, to account for the capturable atoms 'after' the substitution. Accordingly, we will introduce another vector E that contains at least the capturable atoms of θ.

Definition 129. Define $[\![\theta]\!]^E_D$ by:

$$[\![\theta]\!]^E_D(X^S) = \lambda d_1 \ldots \lambda d_n.[\![\theta(X^S)]\!]^E \text{ where } [d_1, \ldots, d_n] = D \cap S.$$

Lemma 130 is useful in the proof of Theorem 131:

Lemma 130. *$dom(\pi) \subseteq \{d_1, \ldots, d_n\}$ implies $(\lambda d_1 \ldots \lambda d_n.g)\pi(d_1) \ldots \pi(d_n) =_{\alpha\beta} \pi \cdot g$.*

Proof. By induction on g.

- The case a. Suppose $\pi(a) = a$ so $a \notin \{d_1, \ldots, d_n\}$, therefore $a[\pi(d_i)/d_i] \equiv a$ for $1 \leq i \leq n$, as required. Otherwise, suppose $\pi(a) \neq a$, so $a \in \{d_1, \ldots, d_n\}$. Then $a[\pi(d_i)/d_i] \equiv \pi(d_i)$ for some $1 \leq i \leq n$. The result follows.
- The case X, f and $g'g$. Routine.
- The case $\lambda a.g$. Suppose $\pi(a) = a$, so $a \notin \{d_1, \ldots, d_n\}$ therefore $a \notin \{\pi(d_1), \ldots, \pi(d_n)\}$. Write h for $(\lambda d_1 \ldots \lambda d_n.\lambda a.g)\pi(d_1) \ldots \pi(d_n)$. Then:

$$\begin{array}{lll}
h & =_{\alpha\beta} & (\lambda a.g)[\pi(d_1)/d_1] \ldots [\pi(d_n)/d_n] \quad \text{Definition 108} \\
& =_{\alpha\beta} & \lambda a.(g[\pi(d_1)/d_1] \ldots [\pi(d_n)/d_n]) \quad a \notin \{d_1, \ldots, d_n\} \\
& =_{\alpha\beta} & \lambda a.(\pi \cdot g) \quad \text{Inductive hypothesis} \\
& \equiv & \lambda \pi(a).(\pi \cdot g) \quad \pi(a) = a \\
& \equiv & \pi \cdot \lambda a.g \quad \text{Definition 96}
\end{array}$$

  The result follows.
  Otherwise, suppose $\pi(a) \neq a$ so $a \in \{d_1, \ldots, d_n\}$ and therefore $\pi(a) \in \{\pi(d_1), \ldots, \pi(d_n)\}$. Assume $a = d_i$ for some d_i and write h as shorthand for $(\lambda d_1 \ldots \lambda d_n.\lambda a.g)\pi(d_1) \ldots \pi(d_n)$. Then:

$$\begin{array}{lll}
h & =_{\alpha\beta} & (\lambda a.g)[\pi(d_1)/d_1] \ldots [\pi(d_n)/d_n] \quad \text{Definition 108} \\
& =_{\alpha\beta} & \lambda b.((b\ a) \cdot g[\pi(d_1)/d_1] \ldots [\pi(d_n)/d_n]) \quad b \notin fa(g),\ b \notin \{d_1, \ldots, d_n\} \\
& =_{\alpha\beta} & \lambda b.((b\ a) \cdot (\pi \cdot g)) \quad \text{Inductive hypothesis} \\
& =_{\alpha} & \lambda \pi(b).((b\ a) \cdot (\pi \cdot g)) \quad \pi(b) = b \\
& \equiv & \pi \cdot \lambda a.g \quad \text{Definitions 98 and 96}
\end{array}$$

  The result follows.

□

Theorem 131. *If $capt(r) \subseteq D$ then $[\![r\theta]\!]^E =_{\alpha\beta} [\![r]\!]^D[\![\theta]\!]^E_D$.*

Proof. By induction on r.

- The cases a. Routine.
- The case $\mathsf{f}(r_1, \ldots, r_n)$. We have:

$$\begin{array}{lll}
[\![\mathsf{f}(r_1, \ldots, r_n)\theta]\!]^E & \equiv & [\![\mathsf{f}(r_1\theta, \ldots, r_n\theta)]\!]^E \quad \text{Definition 43} \\
& \equiv & \mathsf{f}[\![r_1\theta]\!]^E \ldots [\![r_n\theta]\!]^E \quad \text{Definition 117} \\
& =_{\alpha\beta} & \mathsf{f}([\![r_1]\!]^D[\![\theta]\!]^E_D) \ldots ([\![r_n]\!]^D[\![\theta]\!]^E_D) \quad \text{Inductive hypothesis} \\
& \equiv & (\mathsf{f}[\![r_1]\!]^D \ldots [\![r_n]\!]^D)[\![\theta]\!]^E_D \quad \text{Properties of } \lambda\text{-calculus} \\
& \equiv & [\![\mathsf{f}(r_1, \ldots, r_n)]\!]^D[\![\theta]\!]^E_D \quad \text{Definition 117}
\end{array}$$

  The result follows.

- The case $\pi \cdot X^S$. Let $d_1, \ldots, d_n$ be $D \cap S$. By Definition 129, $[\![\theta]\!]^E_D(X^S) = \lambda d_1 \ldots \lambda d_n . [\![\theta(X^S)]\!]^E$. Then:

$$\begin{array}{lcll} [\![(\pi \cdot X^S)\theta]\!]^E & \equiv & [\![\pi \cdot \theta(X^S)]\!]^E & \text{Definition 43} \\ & \equiv & \pi \cdot [\![\theta(X^S)]\!]^E & \text{Lemma 118} \\ & =_{\alpha\beta} & (\lambda d_1 \ldots \lambda d_n . [\![\theta(X^S)]\!]^E)\pi(d_1)\ldots\pi(d_n) & \text{Lemma 130} \\ & \equiv & (X^S\pi(d_1)\ldots\pi(d_n))[\![\theta]\!]^E_D & \text{Definition 129} \\ & \equiv & [\![\pi \cdot X^S]\!]^D[\![\theta]\!]^E_D & \text{Definition 117} \end{array}$$

  The use of Lemma 130 above is valid, as $capt(\pi \cdot X^S) \subseteq D$, therefore $dom(\pi) \cap S \subseteq D \cap S$ by Definition 121. The result follows.
- The case $[a]r$. Choose b fresh, so $b \notin fa([\![\theta(X^S)]\!]^E_D)$ for every $X^S \in fV(r)$ and $b \notin fa(r)$. Then:

$$\begin{array}{lcll} [\![([a]r)\theta]\!]^E & =_{\alpha} & [\![([b]((b\ a) \cdot r))\theta]\!]^E & \text{Definition 11, Theorem 120, Lemma 31} \\ & \equiv & \lambda b.([\![((b\ a) \cdot r)\theta)]\!]^E & \text{Definitions 43 and 117} \\ & =_{\alpha\beta} & \lambda b.([\![(b\ a) \cdot r]\!]^D)[\![\theta]\!]^E_D & \text{Inductive hypothesis} \\ & \equiv & (\lambda b.[\![(b\ a) \cdot r]\!]^D)[\![\theta]\!]^E_D & b \text{ fresh} \\ & \equiv & [\![[b]((b\ a) \cdot r)]\!]^D[\![\theta]\!]^E_D & \text{Definition 117} \\ & =_{\alpha} & [\![[a]r]\!]^D[\![\theta]\!]^E_D & \text{Definition 11, Theorem 120, Lemma 31} \end{array}$$

  The result follows.

□

Recall the instantiation ordering $\theta_1 \leq \theta_2$ from Definition 83. Similarly:

Definition 132. Write $\sigma_1 \leq \sigma_2$ when there exists some σ' such that $X\sigma_1 =_{\alpha\beta} X(\sigma_2 \circ \sigma')$, for any X. Call $\leq$ the **instantiation ordering**.

We can leverage Theorem 131 to prove a corollary, describing a sense in which the instantiation ordering $\theta_1 \leq \theta_2$ of Definition 83 translates to the instantiation ordering of Definition 132:

Corollary 133. *Suppose* $\bigcup_{X^S} capt(\theta_2(X^S)) \subseteq E$.
If $\theta_1 \leq \theta_2$ *then* $[\![\theta_1]\!]^E_D \leq [\![\theta_2]\!]^E_D$.

Proof. Suppose $\theta_1 \leq \theta_2$. By Definition 83, there exists θ' such that $X^S\theta_1 =_{\alpha} X^S(\theta_2 \circ \theta')$ always. We reason as follows, for any unknown X^S:

$$\begin{array}{lcll} [\![X^S]\!]^D[\![\theta_1]\!]^E_D & =_{\alpha\beta} & [\![X^S\theta_1]\!]^E & \text{Theorem 131} \\ & =_{\alpha} & [\![X^S(\theta_2 \circ \theta')]\!]^E & \text{Theorem 120} \\ & \equiv & [\![(X^S\theta_2)\theta']\!]^E & \text{Theorem 33} \\ & =_{\alpha\beta} & [\![X^S\theta_2]\!]^E[\![\theta']\!]^E_E & \text{Theorem 131, } capt(\theta_2(X^S)) \subseteq E \\ & =_{\alpha\beta} & ([\![X^S]\!]^D[\![\theta_2]\!]^E_D)[\![\theta']\!]^E_E & \text{Theorem 131} \\ & \equiv & [\![X^S]\!]^D([\![\theta_2]\!]^E_D \circ [\![\theta']\!]^E_E) & \text{Lemma 113} \end{array}$$

The result follows. □

In Corollary 133, the precondition $\bigcup_{X^S} capt(\theta_2(X^S)) \subseteq E$ is necessary to prevent θ_2 from introducing infinitely many capturable atoms. The 'complexity' of θ_1 is unconstrained. In practice it is likely that we will care about a particular finite set of unknowns $\mathcal{V}$ (for example, $fV(Pr)$ for some Pr), and the precondition can be correspondingly refined to consider just $X^S \in \mathcal{V}$.

9.2. Reducing permissive nominal unification to pattern unification; soundness, weak completeness

The main result of this subsection is Theorem 141. It says that if D and E are 'large enough', then θ solves Pr if and only if $[\![\theta]\!]^E_D$ solves $[\![Pr]\!]^D$.

Definition 134. An **equation** is a pair $r \ {}_?{=}_? \ s$. A **unification problem** Pr is a finite set of equations. A **solution** to Pr is a θ such that $r\theta =_\alpha s\theta$ for all $r {}_?{=}_? s \in Pr$.

Definition 135. If $D = [d_1, \ldots, d_n]$ and $Pr = \{r_1 \ {}_?{=}_? \ s_1, \ldots\}$ then define $[\![Pr]\!]^D$ by:

$$[\![Pr]\!]^D = \left\{\lambda d_1 \ldots \lambda d_n.[\![r]\!]^D \ {}_?{=}_? \ \lambda d_1 \ldots \lambda d_n.[\![s]\!]^D \mid r \ {}_?{=}_? \ s \in Pr\right\}$$

So if $Pr = \{X^S \ {}_?{=}_? \ \mathsf{f}(Y^S, a, Z^S)\}$ where $S = comb \cup \{a, b\}$, then the translation $[\![Pr]\!]^{[a]} = \{\lambda a.(X^S a) \ {}_?{=}_? \ \lambda a.(\mathsf{f} \ (Y^S a) \ a \ (Z^S a))\}$.

Lemma 136. *If $A \subseteq B$ then $capt_A(r) \subseteq capt_B(r)$.*
As a corollary, $capt_A(r) \subseteq capt_A([a]r)$.

Proof. By induction on r.

- The case a. As $capt_A(a) = \varnothing$.
- The case $\mathsf{f}(r_1, \ldots, r_n)$. We have:

$$\begin{array}{lcll} capt_A(\mathsf{f}(r_1, \ldots, r_n)) & = & capt_A(r_1) \cup \ldots \cup capt_A(r_n) & \text{Definition 121} \\ & \subseteq & capt_B(r_1) \cup \ldots \cup capt_B(r_n) & \text{Inductive hypotheses} \\ & = & capt_B(\mathsf{f}(r_1, \ldots, r_n)) & \text{Definition 121} \end{array}$$

  The result follows.
- The case $[a]r$. We have:

$$\begin{array}{lcll} capt_A([a]r) & = & capt_{A\cup\{a\}}(r) & \text{Definition 121} \\ & \subseteq & capt_{B\cup\{a\}}(r) & \text{Inductive hypothesis} \\ & = & capt_B([a]r) & \text{Definition 121} \end{array}$$

  The result follows.
- The case $\pi \cdot X^S$. We have:

$$\begin{array}{lcll} capt_A(\pi \cdot X^S) & = & (dom(\pi) \cup A) \cap S & \text{Definition 121} \\ & \subseteq & (dom(\pi) \cup B) \cap S & \text{Assumption} \\ & = & capt_B(\pi \cdot X^S) & \text{Definition 121} \end{array}$$

  The result follows.

As $capt_A([a]r) = capt_{A\cup\{a\}}(r)$, the corollary follows. □

We need Lemma 137 to prove Lemma 138:

Lemma 137. $capt_A(\pi \cdot r) \subseteq ((dom(\pi) \cup A) \cap fa(r)) \cup capt(r)$.

Proof. By induction on r.

- The cases a and $\mathsf{f}(r_1, \ldots, r_n)$. Routine.
- The case $[a]r$. Suppose $\pi(a) = a$. Then:

$$\begin{array}{lcll} capt_A(\pi \cdot [a]r) & = & capt_A([a](\pi \cdot r)) & \text{Definition 7, } \pi(a) = a \\ & = & capt_{A\cup\{a\}}(\pi \cdot r) & \text{Definition 121} \\ & \subseteq & (dom(\pi) \cap fa(r)) \cup capt(r) & \text{Inductive hypothesis} \\ & \subseteq & (dom(\pi) \cap fa(r)) \cup capt([a]r) & \text{Lemma 136} \\ & = & (dom(\pi) \cap fa([a]r)) \cup capt([a]r) & a \notin dom(\pi) \end{array}$$

Conversely, suppose $\pi(a) \neq a$. Choose fresh b, so $b \notin dom(\pi) \cup fa(r)$. Then, set $\pi' = (b\ a) \circ \pi \circ (b\ a)$. By Lemma 20, $\pi \cdot [a]r =_\alpha \pi' \cdot [a]r$. By similar reasoning as above,

$$capt_A(\pi' \cdot [a]r) \subseteq (dom(\pi') \cup A) \cap fa([a]r)) \cup capt([a]r)$$

By Definition 9, $a \notin fa([a]r)$, and the result follows by sets calculations.

- The case $\pi' \cdot X^S$. Then:

$$\begin{array}{rcl} capt_A((\pi \circ \pi') \cdot X^S) & = & (dom(\pi \circ \pi') \cup A) \cap S \\ & \subseteq & (dom(\pi) \cup dom(\pi') \cup A) \cap S \\ & = & (((dom(\pi) \cup A) \setminus dom(\pi')) \cup dom(\pi')) \cap S \\ & = & (((dom(\pi) \cup A) \setminus dom(\pi')) \cap S) \cup capt(\pi' \cdot X^S) \\ & \subseteq & ((dom(\pi) \cup A) \cap \pi' \cdot S) \cup capt(\pi' \cdot X^S) \\ & = & ((dom(\pi) \cup A) \cap fa(\pi' \cdot X^S)) \cup capt(\pi' \cdot X^S) \end{array}$$

The result follows.

□

Lemma 138. *$fa(t) \subseteq S$ implies $capt_A(r[X^S{:=}t]) \subseteq capt_A(r) \cup capt(t)$. (We* really *do mean '$capt(t)$', and not '$capt_A(t)$'.)*

$capt(r\theta) \subseteq \bigcup_{X^S \in fV(r)} capt(\theta(X^S)) \cup capt(r)$ always.

Proof. The first part is by induction on r.

- The cases a and $\mathsf{f}(r_1, \ldots, r_n)$. Straightforward.
- The case $[a]r$. Then:

$$\begin{array}{rcll} capt_A([a](r[X^S{:=}t])) & = & capt_{A \cup \{a\}}(r[X^S{:=}t]) & \text{Definition 121} \\ & \subseteq & capt_{A \cup \{a\}}(r) \cup capt(t) & \text{Inductive hypothesis} \\ & = & capt_A([a]r) \cup capt(t) & \text{Definition 121} \end{array}$$

The result follows.

- The case $\pi \cdot X^S$. As $(\pi \cdot X^S)[X^S{:=}t] \equiv \pi \cdot t$, we reason as follows:

$$\begin{array}{rcll} capt_A(\pi \cdot t) & \subseteq & ((dom(\pi) \cup A) \cap fa(t)) \cup capt(t) & \text{Lemma 137} \\ & \subseteq & ((dom(\pi) \cup A) \cap S) \cup capt(t) & fa(t) \subseteq S \\ & = & capt_A(\pi \cdot X^S) \cup capt(t) & \text{Definition 121} \end{array}$$

The result follows.

The second part follows from the first. □

Remark 139. $capt(r\theta) \subseteq \bigcup_{fV(r)} capt(\theta(X^S))$ is not true in general. For example if $a \in S$ and $b \in S$ then $capt([a]X^S) = \{a\}$ and $capt([X^S{:=}[b]X^S]) = \{b\}$, and $capt(\theta([a]X^S)) = \{a, b\} \not\subseteq \{b\}$.

Lemma 140. *Suppose $capt(Pr) \subseteq D$ and $capt(Pr\theta) \subseteq E$. Then θ solves Pr if and only if $[\![\theta]\!]^E_D$ solves $[\![Pr]\!]^D$.*

Proof. Suppose $r \ ?{=}?\ s \in Pr$. By Definition 72, $r\theta =_\alpha s\theta$. By Theorems 120 and 126, $[\![r\theta]\!]^E =_\alpha [\![s\theta]\!]^E$. By Theorem 131, $[\![r]\!]^D[\![\theta]\!]^E_D =_{\alpha\beta} [\![s]\!]^D[\![\theta]\!]^E_D$. It is a fact of the λ-calculus that this is equivalent to $\lambda d_1 \ldots \lambda d_n.[\![r]\!]^D[\![\theta]\!]^E_D =_{\alpha\beta} \lambda d_1 \ldots \lambda d_n.[\![s]\!]^D[\![\theta]\!]^E_D$. As no atom of D is free in $[\![\theta]\!]^E_D$, $(\lambda d_1 \ldots \lambda d_n.[\![r]\!]^D)[\![\theta]\!]^E_D =_{\alpha\beta} (\lambda d_1 \ldots \lambda d_n.[\![s]\!]^D)[\![\theta]\!]^E_D$, as required.

The reverse implication uses the same results, in reverse order. □

Theorem 141 (Soundness and weak completeness)**.** *Suppose* $capt(Pr) \subseteq D$*, and* $\bigcup_{X^S \in fV(Pr)} capt(\theta(X^S)) \subseteq E$*, with* $D \subseteq E$*. Then* θ *solves* Pr *if and only if* $[\![\theta]\!]^E_D$ *solves* $[\![Pr]\!]^D$*.*

Proof. An immediate consequence of Lemmas 140 and 138. □

$Pr = \{X^S \ ?=_? \ \mathsf{f}(Y^S, a, Z^S)\}$ where $S = comb \cup \{a, b\}$ translates to $[\![Pr]\!]^{[a]} = \{\lambda a.(X^S\, a) = \lambda a.(\mathsf{f}\,(Y^S\, a)\, a\, (Z^S\, a))\}$.

The solution $[X^S{:=}\mathsf{f}(W^S, a, b), Y^S{:=}W^S, Z^S{:=}b]$ with $S = comb \cup \{a, b\}$ translates to the solution $[\![\theta]\!]^{[a,b]}_{[a]} = [X^S{:=}\lambda a.(\mathsf{f}\,(W^S\, a\, b)\, a\, b), Y^S{:=}\lambda a.(W^S\, a\, b), Z^S{:=}\lambda a.b]$.

9.3. Strong Completeness

The main result of this subsection is Theorem 155. This strengthens the completeness result of Theorem 141, in a certain sense, by expressing that a class of σ solving $[\![Pr]\!]^D$ all originate from θ solving Pr, in a suitable formal sense.

Definition 142. Call a bijection on unknowns a **renaming**. ρ will range over renamings. Each X is also a λ-term (Definition 95), so each ρ is also a substitution (Definition 109).

Lemma 143. $fa(g) = fa(g\rho)$

Proof. By induction on g.

- The case a. Since $a\rho \equiv a$.
- The case X. Since $fa(X) = \varnothing$ and ρ is a bijection on unknowns.
- The case f. Since $\mathsf{f}\rho \equiv \mathsf{f}$.
- The case $g'g$. By hypothesis, $fa(g'\rho) = fa(g')$ and $fa(g\rho) = fa(g)$. As $fa(g'g) = fa(g') \cup fa(g) = fa(g'\rho) \cup fa(g\rho) = fa((g'\rho)g\rho)$, and $(g'\rho)g\rho \equiv (g'g)\rho$, the result follows.
- The case $\lambda a.g$. As $fa((\lambda a.g)\rho) = fa(\lambda a.(g\rho))$ we have $fa(\lambda a.(g\rho)) = fa(g\rho) \setminus \{a\}$. By hypothesis, $fa(g\rho) = fa(g)$. The result follows.

□

Lemma 144. $g =_\alpha h$ *if and only if* $g\rho =_\alpha h\rho$*.*

Proof. The left to right implication is by induction on the derivation of $g =_\alpha h$; right to left is by induction on the derivation of $g\rho =_\alpha h\rho$.

- The cases $(\lambda{=_\alpha}\mathbf{aa})$, $(\lambda{=_\alpha}\mathbf{X})$ and $(\lambda{=_\alpha}\mathsf{f})$. Routine.
- The cases $(\lambda{=_\alpha}\mathbf{p})$ and $(\lambda{=_\alpha}\lambda\mathbf{aa})$. By the inductive hypotheses.
- The case $(\lambda{=_\alpha}\lambda\mathbf{ab})$. For the left to right implication, by hypothesis, $((b\ a)\cdot g)\rho =_\alpha h\rho$ with $b \notin fa(g)$. By Theorem 107 and Lemma 111, $(b\ a)\cdot g\rho =_\alpha h\rho$. By Lemma 143, $b \notin fa(g\rho)$. Using $(\lambda{=_\alpha}\lambda\mathbf{ab})$, $\lambda a.(g\rho) =_\alpha \lambda b.(h\rho)$. The result follows.
  For the right to left implication, suppose $((b\ a)\cdot g)\rho =_\alpha h\rho$ with $b \notin fa(g\rho)$. By Theorem 107 and Lemma 111, $(b\ a)\cdot g\rho =_\alpha h\rho$. By Lemma 143, $b \notin fa(g)$. Using $(\lambda{=_\alpha}\lambda\mathbf{ab})$, $\lambda a.g =_\alpha \lambda b.h$. The result follows.

□

Definition 145. Define the substitution $\pi \cdot \sigma$ by: $(\pi \cdot \sigma)(X) \equiv \pi \cdot \sigma(X)$.

Note that $\pi \cdot \sigma$ is a substitution. $g(\pi \cdot \sigma)$ is not a shorthand for $\pi \cdot (g\sigma)$, and the two are not equal in general.

Lemma 146. *If* $dom(\pi) \cap fa(g) = \varnothing$ *then* $g(\pi \cdot \sigma) =_\alpha \pi \cdot (g\sigma)$.

Proof. By induction on $size(g)$.

- The case a. By assumption, $\pi \cdot a \equiv a$ and $a\sigma \equiv a$.
- The case X. Since $X(\pi \cdot \sigma) \equiv \pi \cdot \sigma(X)$ by Definition 145.
- The case f. Since $\pi \cdot \mathsf{f} \equiv \mathsf{f}$ and $\mathsf{f}\sigma \equiv \mathsf{f}$.
- The case $g'g$. If $dom(\pi) \cap fa(g'g) = \varnothing$, then $dom(\pi) \cap fa(g') = \varnothing$ and $dom(\pi) \cap fa(g) = \varnothing$. Then:

$$\begin{array}{rcll} (g'g)(\pi \cdot \sigma) & \equiv & g'(\pi \cdot \sigma)(g(\pi \cdot \sigma)) & \text{Definition 110} \\ & =_\alpha & (\pi \cdot (g'\sigma))(\pi \cdot (g\sigma)) & \text{Inductive hypotheses} \\ & \equiv & \pi \cdot ((g'g)\sigma) & \text{Definition 96} \\ & \equiv & \pi \cdot ((g'g)\sigma) & \text{Definition 110} \end{array}$$

- The case $\lambda a.g$. There are multiple cases to consider:
  - The case $a \notin fa(g\sigma)$, $a \notin fa(\pi \cdot \sigma)$ and $\pi(a) = a$. Then:

$$\begin{array}{rcll} (\lambda a.g)(\pi \cdot \sigma) & \equiv & \lambda a.(g(\pi \cdot \sigma)) & \text{Definition 110} \\ & =_\alpha & \lambda a.(\pi \cdot (g\sigma)) & \text{Inductive hypothesis} \\ & \equiv & \lambda\pi(a).(\pi \cdot (g\sigma)) & \text{Assumption} \\ & \equiv & \pi \cdot \lambda a.(g\sigma) & \text{Definition 96} \\ & \equiv & \pi \cdot (\lambda a.g)\sigma & \text{Definition 110} \end{array}$$

    The result follows.
  - The case $a \notin fa(g\sigma)$, $a \notin fa(\pi \cdot \sigma)$ and $\pi(a) \neq a$. Pick fresh b, so $b \notin fa(g)$, $b \notin fa(g\sigma)$, $b \notin fa(\pi \cdot \sigma)$ and $\pi(b) \neq b$. Every permutation has finite support, so b is guaranteed to exist. Then:

$$\begin{array}{rcll} (\lambda a.g)(\pi \cdot \sigma) & =_\alpha & (\lambda b.((b\ a) \cdot g))(\pi \cdot \sigma) & \text{Definition 98} \\ & \equiv & \lambda b.((b\ a) \cdot g)(\pi \cdot \sigma)) & \text{Definition 110} \\ & =_\alpha & \lambda b.(\pi \cdot (((b\ a) \cdot g)\sigma)) & \text{Inductive hypothesis} \\ & \equiv & \lambda b.((\pi(b)\ \pi(a)) \cdot ((\pi \cdot g)\sigma)) & \text{Fact} \\ & \equiv & \lambda\pi(b).((\pi(b)\ \pi(a)) \cdot ((\pi \cdot g)\sigma)) & \text{Assumption} \\ & \equiv & \lambda\pi(a).(\pi \cdot (g\sigma)) & \text{Definition 96} \\ & \equiv & ((\pi(b)\ \pi(a)) \circ \pi) \cdot (\lambda a.(g\sigma)) & \text{Definition 96, Lemma 15} \\ & \equiv & \pi \cdot ((b\ a) \cdot (\lambda a.g)\sigma) & \text{Definition 110} \\ & \equiv & \pi \cdot ((\lambda b.((b\ a) \cdot g))\sigma) & \text{Definition 96} \\ & =_\alpha & \pi \cdot (\lambda a.g)\sigma & \text{Definition 98} \end{array}$$

    The result follows.

  All other cases are similar to the case for $a \notin fa(g\sigma)$, $a \notin fa(\pi \cdot \sigma)$ and $\pi(a) \neq a$. The result follows.

□

Lemma 147. *σ solves $[\![Pr]\!]^D$ if and only if $\sigma \circ \rho$ does.*

Suppose $dom(\pi) \cap (fa(r) \cup fa(s)) = \varnothing$ for every $r \ ?{=}? \ s \in Pr$. Then σ solves $[\![Pr]\!]^D$ if and only if $\pi \cdot \sigma$ does.

Proof. For the first part, we have two cases:

- The case σ solves $[\![Pr]\!]^D$ implies $\sigma \circ \rho$ solves $[\![Pr]\!]^D$. Suppose $g \ ?{=}? \ h \in [\![Pr]\!]^D$ and σ solves $[\![Pr]\!]^D$. Then $g\sigma =_\alpha h\sigma$. By Lemma 144, $g\sigma\rho =_\alpha h\sigma\rho$. By Lemma 113, $g(\sigma \circ \rho) =_\alpha h(\sigma \circ \rho)$. The result follows.

- The case $\sigma\circ\rho$ solves $[\![Pr]\!]^D$ implies σ solves $[\![Pr]\!]^D$. Suppose $g \ {}_?{=}_?\ h \in [\![Pr]\!]^D$ and $\sigma\circ\rho$ solves $[\![Pr]\!]^D$. Then $g(\sigma\circ\rho) =_\alpha h(\sigma\circ\rho)$. By Lemma 113, $g\sigma\rho =_\alpha h\sigma\rho$. By Lemma 144, $g\sigma =_\alpha h\sigma$. The result follows.

For the second part, suppose $dom(\pi) \cap (fa(r) \cup fa(s)) = \varnothing$ for every $r \ {}_?{=}_?\ s \in Pr$ and $D = [d_1, \ldots, d_n]$. Then $[\![Pr]\!]^D = \{\lambda d_1 \ldots \lambda d_n.[\![r]\!]^D \ {}_?{=}_?\ \lambda d_1 \ldots \lambda d_n.[\![s]\!]^D \mid r \ {}_?{=}_?\ s \in Pr\}$. By Lemma 118, $dom(\pi) \cap (fa([\![r]\!]^D) \cup fa([\![s]\!]^D)) = \varnothing$. By Lemma 146, Theorem 107 and Lemma 102, the result follows. □

Remark 148. Lemma 147 expresses an intuition that 'names of atoms and unknowns on the right in a solution, do not matter', which also underlies the π and ρ in Theorem 155. ρ is the price we pay for using the same unknowns in Definitions 95 and 6: This design decision makes Definition 117 compact, but it causes technical problems in Lemma 154, because $\sigma(X)$ can introduce new unknowns over whose permission sorts (back in the nominal world) we have no control. ρ lets us rename those new unknowns as convenient. As for π, we discuss it below.

Another design decision is to work with an untyped λ-calculus. This simplifies our presentation and makes our results slightly more powerful (because they apply to more substitutions), but we cannot be *too* liberal: Suppose σ solves $[\![Pr]\!]^D$. Examining Definition 117, if X occurs in $[\![Pr]\!]^D$ then it is applied to a number of atoms equal to the length of $D \cap S$. So, we will only be interested in σ that respect this fragment of typability ($\mathcal{V}$ will be $fV(Pr)$):

Definition 149. Let $\mathcal{V}$ be a finite set of unknowns. Call σ **D-consistent on $\mathcal{V}$** when for every $X \in \mathcal{V}$, $\sigma(X) =_\alpha \lambda a_1 \ldots \lambda a_k.q$ where k is the length of $D \cap S$. (So $\sigma(X)$ starts with 'at least' length-$D \cap S$-many λ-abstractions.)

Call σ **strictly** D-consistent when also, for every $X \in \mathcal{V}$, $fa(\sigma(X)) \cap D = []$.

Remark 150. Strictness is motivated by the following examples: Take $D = [a]$.

Take $Pr = \{X^S \ {}_?{=}_?\ \mathsf{f}([a]Y^S, Y^S)\}$ with $S = comb$. Then the problem $[\![Pr]\!]^D = \{\lambda a.(X^S\, a) \ {}_?{=}_?\ \lambda a.(\mathsf{f}\,(\lambda a.(Y^S a))\,(Y^S a))\}$ has the solution $\sigma = [X^S{:=}\lambda c.(\mathsf{f}\,(\lambda c.a)\,a)$, $Y^S{:=}\lambda c.a]$. $(\sigma\circ\rho)(Y^S) =_\alpha [\![\theta]\!]^E_D(Y^S)$ is impossible for any ρ, since $\lambda c.a =_\alpha \lambda a.[\![\theta(Y^S)]\!]^E$ is impossible.

Take $Pr = \{X^S \ {}_?{=}_?\ \mathsf{f}([a]Y^T, Y^T)\}$ with $S = comb$ and $T = comb \setminus \{a\}$. Then $[\![Pr]\!]^D = \{\lambda a.(X^S\, a) \ {}_?{=}_?\ \lambda a.(\mathsf{f}\,(\lambda a.Y^T)\,Y^T)\}$ has the solution $\sigma = [X^S{:=}\lambda c.(\mathsf{f}\,(\lambda c.a)\,a)$, $Y^T{:=}a]$. $(\sigma\circ\rho)(Y^T) =_\alpha [\![\theta]\!]^E_D(Y^T)$ is impossible, since $a \in fa(a)$ whereas $a \notin fa([\![\theta(Y^T)]\!]^E)$ by Lemma 119.

The a in $\sigma(Y^T)$ for the two σ considered above, has nothing to do with the a in D. We can regard this as an unfortunate 'name-clash' which Lemma 147 allows us to eliminate with a permutation π.

More on this in Theorem 155. We continue with the proofs:

Definition 151. Define the **arguments of unknowns** in a pattern q by:

$$args(a) = \varnothing \quad args(X) = \varnothing \quad args(Xa_1 \ldots a_n) = \{a_1, \ldots, a_n\}$$
$$args(\mathsf{f}q_1 \ldots q_n) = \bigcup_{1 \le i \le n} args(q_i) \quad args(\lambda a.q) = args(q)$$

$q =_\alpha r$ does not imply $args(q) = args(r)$. This is by design.

Definition 152. Suppose q is a ϕ-pattern and $args(q) \subseteq E$. Define a nominal term $q^{\text{-}E}$ by:

$$a^{\text{-}E} \equiv a \quad (Xb_1 \ldots b_{\phi(X)})^{\text{-}E} \equiv \pi\cdot X^S \quad (\lambda a.q)^{\text{-}E} \equiv [a]q^{\text{-}E} \quad (\mathsf{f}q_1 \ldots q_n)^{\text{-}E} \equiv \mathsf{f}(q_1{}^{\text{-}E}, \ldots, q_n{}^{\text{-}E})$$

Here π is a fixed but arbitrary choice of permutation of the atoms in E, mapping the i^{th} element of $E \cap S$ (Definition 116) to b_i for $1 \leq i \leq \phi(X)$.

Lemma 153. *$args(q) \subseteq E$ implies $[\![q^{\text{-}E}]\!]^E \equiv q$.*

Proof. By induction on q.

- The cases a and $\mathtt{f} q_1 \dots q_n$. Routine.
- The case $\lambda a.g$. Suppose $args(\lambda a.q) \subseteq E$ so that $args(q) \subseteq E$. By hypothesis, $[\![q^{\text{-}E}]\!]^E \equiv q$. The result now follows.
- The case $X b_1 \dots b_n$. Then $q^{\text{-}E} = \pi \cdot X$ and $[\![q^{\text{-}E}]\!]^E \equiv X\pi(x_1) \dots \pi(x_n)$, where $[x_1, \dots, x_n] = E \cap \delta(X)$ and $\pi(x_i) = b_i$. The result follows.

□

Lemma 154. *Suppose $\mathcal{V}$ is a finite set of unknowns and σ is a ϕ-pattern substitution, strictly D-consistent on $\mathcal{V}$.*

Then there exist ρ, θ, and E, such that $D \subseteq E$, $\bigcup_{X \in \mathcal{V}} capt(\theta(X)) \subseteq E$, and $(\sigma \circ \rho)(X) =_\alpha [\![\theta]\!]^E_D(X)$ for every $X \in \mathcal{V}$.

Proof. Take any $E = [e_1, ..., e_p]$ which includes all atoms in D and in $\{\sigma(X) \mid X \in \mathcal{V}\}$. Define $\mathcal{V}' = \bigcup_{X \in \mathcal{V}} fV(\sigma(X))$ ('the unknowns in $\sigma(X)$ for $X \in \mathcal{V}$'). For each $Y \in \mathcal{V}'$ choose a fresh Y' such that the length of $E \cap fa(Y')$ is equal to $\phi(Y)$. We do this injectively, so that for distinct $Y, Z \in \mathcal{V}'$, Y' and Z' are also distinct. Let ρ be any renaming such that $\rho(Y) \equiv Y'$ for all $Y \in \mathcal{V}'$.

By assumption $\sigma(X) =_\alpha \lambda a_1. \dots \lambda a_n.q$ for a ϕ-pattern q, where $[a_1, \dots, a_n] = D \cap S$. Take $\theta(X) \equiv (q\rho)^{\text{-}E}$.

We can verify that $\bigcup_{X \in \mathcal{V}} capt(\theta(X)) \subseteq E$. We then reason as follows:

$$\begin{array}{lcll} [\![\theta]\!]^E_D(X) & \equiv & \lambda a_1. \dots \lambda a_n.[\![(q\rho)^{\text{-}E}]\!]^E & \text{Definition 129} \\ & \equiv & \lambda a_1. \dots \lambda a_n.(q\rho) & \text{Lemma 153} \\ & \equiv & (\lambda a_1. \dots \lambda a_n.q)\rho & \text{Fact of } \lambda\text{-calculus} \\ & =_\alpha & (\sigma \circ \rho)(X) & \text{By construction} \end{array}$$

□

Theorem 155. *Suppose $capt(Pr) \subseteq D$.*

For σ strictly D-consistent on $fV(Pr)$ solving $[\![Pr]\!]^D$ there are ρ, θ, and E, such that $(\sigma \circ \rho)(X) =_\alpha [\![\theta]\!]^E_D(X)$ for all $X \in fV(Pr)$ and θ solves Pr.

For σ D-consistent on $fV(Pr)$ solving $[\![Pr]\!]^D$ there are π, ρ, θ, and E, such that $\pi \cdot (\sigma \circ \rho)(X) =_\alpha [\![\theta]\!]^E_D(X)$ for all $X \in fV(Pr)$ and θ solves Pr.

Proof. By Lemma 154, there are ρ, θ, and E, such that $(\sigma \circ \rho)(X) =_\alpha [\![\theta]\!]^E_D(X)$ for all $X \in fV(Pr)$, $D \subseteq E$ and $\bigcup_{X \in fV(Pr)} capt(\theta(X)) \subseteq E$. $capt(Pr) \subseteq D$ and $D \subseteq E$, so $capt(Pr) \subseteq E$. By Theorem 141, θ solves Pr.

For the second part, write $D = [d_1, \dots, d_n]$, choose $D' = [d'_1, \dots, d'_n]$ fresh (so d'_i is not in D, Pr, or $\sigma(X)$ for any $X \in fV(Pr)$), and take $\pi = (d'_1\ d_1) \dots (d'_n\ d_n)$. $\pi \cdot \sigma$ is strictly D-consistent and the result follows from the first part and Lemma 147. □

10. Conclusions

Nominal contrasted with permissive nominal terms. Permissive nominal terms come closer to first- and higher-order terms than nominal terms do, but they are a special

case of neither. The idea of associating permissions sets to unknowns is mentioned already in [27, Remark 2.6]. What really makes that idea come alive, in this paper, is the use of fixed permissions sets of *co-infinite* sets of atoms. This has beneficial technical repercussions which go well beyond 'just tweaking nominal terms'. We recover Theorem 13 and Corollary 14, α-equivalence is a property of terms (of course; there are no longer freshness contexts) — and the notions of unification problem and solution are based on equality (rather than equality-and-freshness-context) with no loss of expressivity.

Permissive nominal terms do not obsolete nominal terms; if we want to talk about 'an arbitrary term', then a nominal terms unknown $\dot{X}$ is more directly useful than a permissive nominal terms unknown X^{comb} (which means 'an arbitrary term, mentioning atoms in *comb*').

In Section 4 we connected the 'permissive' and the 'nominal' worlds in some technical detail. In nominal terms, if we need a fresh name then we *can* enrich the freshness context (consider [12, Figure 2, axiom (**fr**)] and [13, e.g. Lemma 25 and Theorem 33]). One nice way to view the interpretation of Section 4 is that *comb* plays the rôle of 'the atoms we had so far' and $\mathbb{A} \setminus comb$ that of 'the atoms we will generate fresh in the future'.

Related work on unification. Patterns emerged by studying Skolemisation of unification problems [22]; they proved useful in the unification of higher-order abstract syntax terms [21]. Cheney proposed a two-stage translation of higher-order to nominal unification [3], first by exhibiting a translation of higher-order pattern unification to nominal pattern unification (where nominal patterns are a variant of nominal terms, with a *concretion* operator, where unknowns have empty support), followed by a translation between nominal pattern unification and nominal unification. Levy and Villaret's translation [19], of nominal unification to higher-order patterns, crystallised an intuition that pattern unification is exactly what is needed to unify encodings of nominal terms. Their encoding is not minimal and addresses unifiability rather than individual solutions. In Section 8 we refined their encoding, using $capt(r)$ (Definition 121) to obtain one that is minimal, and in Section 9 we established a precise sense in which solutions correspond across the translations.

Hamana's β_0 unification of λ-terms with holes adds a capturing substitution [16]. Level 2 variables (which are instantiated) are annotated with level 1 variable symbols that *may* appear in them; permissive nominal terms move in this direction in the sense that permissions sorts also describe which level 1 variable symbols (we call them atoms in this paper) may appear in them, though with our permissions sorts there are infinitely many that may, and infinitely many that may not. Another significant difference is that the treatment of α-equivalence in Hamana's system is not nominal (not based on permutations) and unlike our systems, Hamana's does not have most general unifiers. Similarly, Qu-Prolog [23] adds level 2 variables, but does not manage α-conversion in nominal style, and, for better or for worse, the system is more ambitious in what it expresses, and thus loses mathematical properties (unification is semi-decidable, most general unifiers need not exist).

Future work. We noted, in Definition 2, that *comb* is incompatible with the finite-support property of nominal sets [15, Definition 3.1]. This matters because permissive nominal terms can be directly quotiented by α-equivalence, so it could be useful to apply the Gabbay-Pitts model of abstract syntax up to α-equivalence [14]. We hypothesise that this can be overcome by using generalisations of nominal sets by the second author

[10] or by Cheney ([2, Section 3], or [4]). We also hypothesise a theory of rewriting could be developed similarly to [9].

Via the interpretation in Section 4 this extends to solutions of 'ordinary' nominal unficiation problems. We have begun to apply permissive nominal terms to construct novel logics and λ-calculi, taking advantage of their properties to simplify the theory — we find it very useful to reason on terms (without a freshness context), to have an inexhaustible supply of fresh names, and to be able to quotient by α-equivalence.

Nominal terms come with a denotation in nominal sets [14]. These are based on the idea of giving names a denotational reality as *urelemente* [28] (the *atoms* in this paper can be considered urelemente of a sets universe; this is a reason that nominal terms retain a first-order flavour). Famously, nominal sets exclude sets like permission sorts S, because they do not have finite support. This is fully consistent with our use of permission sorts here; in this paper we are working at the meta-level where we can talk about any sets of atoms that we like. However, generalisations of nominal sets exist [10, 2] and we believe that permissive nominal terms can use them for denotation. Checking this is future work.

References

[1] Franz Baader and Tobias Nipkow. *Term rewriting and all that*. Cambridge University Press, Great Britain, 1998.

[2] James Cheney. *Nominal Logic Programming*. PhD thesis, Cornell University, August 2004.

[3] James Cheney. Relating nominal and higher-order pattern unification. In *Proceedings of the 19th International Workshop on Unification (UNIF 2005)*, pages 104–119, 2005.

[4] James Cheney. Completeness and Herbrand theorems for nominal logic. *Journal of Symbolic Logic*, 71:299–320, 2006.

[5] Gilles Dowek and Murdoch J. Gabbay. Relating the solutions of nominal unification and pattern unification. 2009. Available online at `http://www.gabbay.org.uk/papers/relsnp.pdf`.

[6] Gilles Dowek, Murdoch J. Gabbay, and Dominic P. Mulligan. Permissive nominal terms and their unification. 2009. Available online at `http://www.gabbay.org.uk/papers/perntu.pdf`.

[7] Gilles Dowek, Murdoch J. Gabbay, and Dominic P. Mulligan. Permissive nominal terms and their unification. 2009. Available online at `http://www.gabbay.org.uk/papers/perntu-jv.pdf`.

[8] Gilles Dowek, Thérèse Hardin, and Claude Kirchner. Binding logic: Proofs and models. In *LPAR '02: Proceedings of the 9th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning*, pages 130–144, London, UK, 2002. Springer.

[9] Maribel Fernández and Murdoch J. Gabbay. Nominal rewriting (journal version). *Information and Computation*, 205(6):917–965, 2007.

[10] Murdoch J. Gabbay. A General Mathematics of Names. *Information and Computation*, 205(7):982–1011, July 2007.

[11] Murdoch J. Gabbay and Stéphane Lengrand. The lambda-context calculus. *Electronic Notes in Theoretical Computer Science*, 196:19–35, 2008.

[12] Murdoch J. Gabbay and Aad Mathijssen. A formal calculus for informal equality with binding. In *Proceedings of* 14^{th} *Workshop on Logic, Language and Information in Computation (WoLLIC 2007)*, volume 4576 of *Lecture Notes in Computer Science*, pages 162–176, 2007.

[13] Murdoch J. Gabbay and Dominic P. Mulligan. Two-and-a-halfth Order Lambda-calculus. *Electronic Notes in Theoretical Computer Science*, 2009. To appear.

[14] Murdoch J. Gabbay and A. M. Pitts. A New Approach to Abstract Syntax with Variable Binding (journal version). *Formal Aspects of Computing*, 13(3–5):341–363, 2001.

[15] Murdoch J. Gabbay and Andrew M. Pitts. A New Approach to Abstract Syntax Involving Binders. In *14th Annual Symposium on Logic in Computer Science*, pages 214–224. IEEE Computer Society Press, 1999.

[16] Makoto Hamana. A logic programming language based on binding algebras. In *TACS'01*, volume 2215 of *Lecture Notes in Computer Science*, pages 243–262. Springer, 2001.

[17] Robert Harper, Furio Honsell, and Gordon Plotkin. A framework for defining logics. In *Proc. 2nd Annual IEEE Symposium on Logic in Computer Science, LICS'87*, pages 194–204. IEEE Computer Society Press, 1987.

[18] J.-W. Klop, V. van Oostrom, and F. van Raamsdonk. Combinatory reduction systems. *Theoretical Computer Science*, 121:279–308, 1993.

[19] Jordi Levy and Mateu Villaret. Nominal unification from a higher-order perspective. In *Proceedings of RTA'08*, volume 5117 of *Lecture Notes in Computer Science*. Springer, 2008.

[20] Richard Mayr and Tobias Nipkow. Higher-order rewrite systems and their confluence. *Theoretical Computer Science*, 192:3–29, 1998.

[21] Dale Miller. A logic programming language with lambda-abstraction, function variables, and simple unification. *Journal of Logic and Computation*, 1(4):497 – 536, 1991.

[22] Dale Miller. Unification under a mixed prefix. *Journal of Symbolic Computation*, 14(4):321–358, 1992.

[23] Peter Nickolas and Peter J. Robinson. The Qu-Prolog unification algorithm: formalisation and correctness. *Theoretical Computer Science*, 169(1):81–112, 1996.

[24] Lawrence C. Paulson. Isabelle: the next 700 theorem provers. In P. Odifreddi, editor, *Logic and Computer Science*, pages 361–386. Academic Press, 1990.

[25] F. Pfenning and C. Elliot. Higher-order abstract syntax. In *PLDI (Programming Language design and Implementation)*, pages 199–208. ACM Press, 1988.

[26] M. R. Shinwell, A. M. Pitts, and Murdoch J. Gabbay. FreshML: Programming with Binders Made Simple. In *ICFP'03*, volume 38, pages 263–274. ACM Press, 2003.

[27] Christian Urban, Andrew M. Pitts, and Murdoch J. Gabbay. Nominal Unification. *Theoretical Computer Science*, 323(1–3):473–497, 2004.

[28] Ernst Zermelo. Untersuchungen über die Grundlagen der Mengenlehre. *Mathematische Annalen*, 65:261–281, 1908.