THE ELIMINABILITY PROPERTY IN EXTENSIONS BY SORT

María Claudia MERÉ
e-mail: mere@inf.puc-rio.br
Paulo A. S. VELOSO
e-mail: veloso@inf.puc-rio.br

Abstract

The process of implementing a formal specification on another one is very important in formal software development and can be described in terms of simple logical concepts as an interpretation into a conservative extension. An extension consists of the addition of new symbols and axioms, a simple, but important, special case being the so-called extensions by definitions. The new symbols to be introduced correspond to sorts, functions and predicates. Extensions by the addition of function and predicate symbols are well studied in the literature.

Our purpose here is to analyse the eliminability property in conservative extensions that introduce new sorts. This is of importance because it occurs often in implementing formal specifications, when new sorts are "constructed" from the concrete ones.

We view some known sort introduction constructs, akin those found in programming languages (namely cartesian product, discriminated union, subsort and quotient), like extensions of a base language with the given sorts and study some common properties to these extensions.

Key words: Formal specifications, sort constructors, extensions, implementation, abstract data types, parameterized specifications, formal program development.

Resumo

O processo de implementar uma especificação formal em outra é de grande importância no desenvolvimento formal de programas, podendo ser descrito em termos de conceitos lógicos simples como uma interpretação em uma extensão conservativa. Uma extensão consiste do acréscimo de novos símbolos e axiomas, um caso especial simples, porém importante, sendo as extensões por definições. Os novos símbolos a serem introduzidos correspondem a sortes, funções ou predicados. Extensões por acréscimo de símbolos de função e de predicado estão bem estudadas na literatura.

Nosso objetivo aqui é analisar a propriedade de eliminabilidade em extensões conservativas que introduzem novos sortes. Isto é importante pois ocorre com freqüência na implementação de especificações formais, quando novos sortes são "construídos" a partir dos concretos.

Vemos alguns construtores de introdução de sortes, similares aqueles encontrados nas linguagens de programação (produto cartesiano, união discriminada, subsorte e quociente), como extensões de uma linguagem base com os sortes dados e estudamos algumas propriedades comuns de estas extensões.

Palavras chave: Especificações formais, construtores de sortes, extensões, implementação, tipos abstratos de dados, especificações parametrizadas, desenvolvimento formal de programas.

Contents

1 Introduction **2**

2 Eliminability for new sorts **3**
2.1 Cartesian Product 4
2.2 Subsort 4
2.3 Discriminated Union 5
2.4 Quotient sort 5

3 Extensions by Product Sorts behave like Definitions **6**

4 Extensions by Subsorts behave like Definitions **7**

5 Extensions by Quotient Sorts behave like Definitions **9**

6 Extensions by Discriminated Union Sorts behave like Definitions **10**

7 Conclusions **12**

References **13**

A Appendix **14**

B Appendix **18**

C Appendix **22**

D Appendix **25**

1 Introduction

The process of implementing a formal specification on another one is very important in formal software development and it can be described in terms of simple basic logical concepts.

Formal specifications are presentations of theories in many-sorted logic, and an *implementation* of a formal specification A on another formal specification C amounts to an interpretation of A into a conservative extension B of C. The idea here is that the 'concrete' specification C is extended so as to incorporate versions to which the abstract symbols of A can then be mapped, [TM87, Vel87].

An *extension* consists of the addition of new symbols and axioms; a simple, but important, special case being the so-called *extensions by definitions*. The new symbols to be introduced correspond to sorts, functions or predicates. Extensions by the addition of function and predicate symbols are well studied in the literature. In particular, extensions by definitions of function and predicate symbols are characterized by the property of *unique expandability*, which entails *conservativeness* and *eliminability* [Sho67].

Following a previous work [MV92], our purpose here is to analyse **conservative extensions that introduce new sorts**. This is of importance because it occurs often in implementing formal specifications, when new sorts are "constructed" from the concrete ones. We shall first remember the specification of some well-known sort introduction constructs ([MV92]), akin to those found in programming languages ([Hoa74]), namely cartesian product, discriminated union, subsort and quotient. This last construct, quotient, even though not prevalent in programming languages, corresponds to the representation of the equality on an abstract sort, which is necessary for a correct implementation.

In [MV92] we centered our efforts in axiomatising the specifications and establishing their adequacy in the sense of exhibiting the intented behaviour and in [VM94] we viewed these specifications as parametrized specifications by a base language, where the concretes sorts live. In this work we will center our attention in establishing that these extensions are like extensions by definition of a new sort, even though in [MV92] we proved that we have only unique expandability up to isomorphism. When performing an extension one expects an increase in both deductive and expressive powers. On the other hand, one often does not wish to disturb the smaller theory. When the extension only adds properties of the new symbols, without disturbing the old ones, one has a conservative extension, which does not increase the deductive power for the properties of the old symbols. Similarly, the increase in expressive power may be inessential, in which case our extension has the property of eliminability.

Extensions by definition of predicate or function symbols have some nice properties.

Consider such an extension of theory T over language L to T' over L'. We then

have **Eliminability**. Every formula ρ of L' is equivalent in T' to a formula φ of L. Consequently, we have the following properties:

- **Interpretability**. T' can be interpreted back into L.
- On the model-theoretic side, we have **Invariance under elementary equivalence**. Two models of T' are elementarily equivalent whenever their reducts to L are so.
- **Definability**. A relation is definable in a model of T' iff it is definable in its reduct to L.

In the case of introduction of a sort, we have new variables over the new sort, and new universes realising the new sort. To cope with these, we use two ideas

- **connection**: a local translation of variables ranging over the new sort.
- **internalisation**: a view of relations involving the new sorts.

With these ideas we will be able to generalise the above properties of extensions by definitions of predicate or function symbols to our sort introducing constructs.

The next four sections have the same structure: for each sort introduction construct we prove same interested properties, that corroborate the intuitive feeling that these amount to extensions by definition of a new sort. Actually, we will present a generalised eliminability property that account for the new variables ranging over the sort introduced.

In section 7 we present some conclusions and some sugestions of future works that will deal with internalisation and definability.

2 Eliminability for new sorts

As mentioned in the introduction, an extension causes increases in both deductive and expressive powers. But, one often does not wish to disturb the smaller theory. When the extension does not increase the deductive power for the properties of the old symbols, one has a conservative extension. Similarly, the increase in expressive power may be inessential, in which case our extension has the property of eliminability.

Extensions by definition of predicate or function symbols have some nice properties.

Consider such an extension of theory T over language L to T' over L'. We then have **Eliminability**. Every formula ψ of L' is equivalent in T' to a formula φ of L. Consequently, we have the following properties:

- **Interpretability**. T' can be interpreted back into L.
- On the model-theoretic side, we have **Invariance under elementary equivalence**. Two models of T' are elementarily equivalent whenever their reducts to L are so.

- **Definability**. A relation is definable in a model of T' iff it is definable in its reduct to L.

The above properties are well known in the case of extensions by additions of predicate or function symbols; in fact conservativeness and eliminability characterise when such extensions are by definition. In the case of introduction of a sort, we have new variables over the new sort, and new universes realising the new sort. To cope with these, we use two ideas

- **connection**: a local translation of variables ranging over the new sort.
- **internalisation**: a view of relations involving the new sorts.

With these ideas we will be able to generalise the above properties of extensions by definitions of predicate or function symbols to our sort introducing constructs.

We shall be dealing with extensions by introductin of a new sort. It is convenient to establish some notational conventions. Consider an extension of theory T over language L to theory T' over language L'. We shall generally use letters like v, u, and x (perhaps with subscripts or superscripts) for variables of the smaller language L. Similarly, we shall employ letters like y and z (perhaps with subscripts or superscripts) for variables ranging over the new sort: new variables of the larger language L'. We also use the notational device of an overbar to indicate a set of variables, e. g. $\overline{v}$ for a set of variables of the smaller language L and $\overline{y}$ for a set of variables ranging over the new sort. Generally when we display a term t as $t(\overline{v}, \overline{y})$ we mean that t has no occurrences of variables other than those in $\overline{v} \cup \overline{y}$. Similarly, when we display a formula φ as $\varphi(\overline{v}, \overline{y})$ we mean that φ has no free occurrences of variables other than those in $\overline{v} \cup \overline{y}$.

2.1 Cartesian Product

Let T be a specification with two sorts S_1, S_2. Then, a specification of the cartesian product of these two sorts is:

T' = T +

sort	P	
operations	$p_1 : P \to S_1$	
	$p_2 : P \to S_2$	
axioms		

- $\forall x_1 : S_1\ \forall x_2 : S_2\ \ \exists y : P \quad p_1(y) = x_1 \wedge p_2(y) = x_2$
- $\forall y_1, y_2 : P \quad p_1(y_1) = p_1(y_2) \wedge p_2(y_1) = p_2(y_2) \Longrightarrow y_1 = y_2$

These axioms make P behave as the cartesian product of S_1 and S_2, as was shown in [MV92].

2.2 Subsort

Let T be a specification with a sort S and an unary (relativisation) predicate r over S (**r** points which elements of **S** are in the subset **N** of **S**) such that $\mathrm{T} \models \exists x : S\ r(x)$.

Then, a specification of the subsort N of the sort S in according with r is:
T' = T +

sort N
operations $j : N \to S$
axioms

- $\forall x : S \quad r(x) \iff \exists y : N \quad x = j(y)$
- $\forall y_1, y_2 : N \quad j(y_1) = j(y_2) \Longrightarrow y_1 = y_2$

These axioms force N to behave as the set of elements of S selected by r.

2.3 Discriminated Union

Let T be a specification with two sorts S_1 and S_2. Then, a specification of the (discriminated) union of these two sorts is:
T' = T +

sort U
operations $i_1 : S_1 \to U$
$i_2 : S_2 \to U$
axioms

- $\forall y : U \quad (\exists x_1 : S_1 \quad y = i_1(x_1)) \vee (\exists x_2 : S_2 \quad y = i_2(x_2))$
- $\forall x_1 : S_1 \; \forall x_2 : S_2 \; \neg(i_1(x_1) = i_2(x_2))$
- $\forall x_1, x_1' : S_1 \quad i_1(x_1) = i_1(x_1') \Rightarrow x_1 = x_1'$
- $\forall x_2, x_2' : S_2 \quad i_2(x_2) = i_2(x_2') \Rightarrow x_2 = x_2'$

These axioms make U exhibit the behaviour of the discriminated union of S_1 and S_2, as will be shown.

2.4 Quotient sort

Let T be a specification with a sort S and a binary predicate q over S, which is proved in T to be an equivalence relation. Then, a specification of the quotient sort Q of the sort S by the relation q is:
T' = T +

sort Q
operations $p : S \to Q$
axioms

- $\forall x_1, x_2 : S \quad q(x, x') \iff p(x_1) = p(x_2)$
- $\forall y : Q \quad \exists x : S \quad p(x) = y$

The axioms state that q is the kernel of the function p and also p is onto.

3 Extensions by Product Sorts behave like Definitions

Given theory T with language L we wish to add a new sort P to behave as the product of sorts S_1 and S_2 of L.

Call $T' = T \cup Prod$ the extension of T by introduction of the new sort, where $Prod$ is the axiomatisation for product extension given in section 2. We shall first show that we have a generalised eliminability: eliminability under connection. For this purpose we start by analysing the terms and atomic formulae of the larger language L'.

The following characterisation of terms of L' should be clear, since the product extension adds no new function (or constant) symbols having sort P as target, the other assertion following from the substitutivity property of equality.

Lemma 3.1 (Characterisation of terms of L') .

1. The terms of L' of sort P are the variables of sort P.
2. Each term $t'(\overline{v}, \overline{y})$[1] of L' with sort in L is the result $t(\overline{v}, \overline{p_i(y)})$ of replacing in term $t(\overline{v}, \overline{u})$ of L every occurrence of a new variable u^k in $\overline{u}$ of sort S_i by $(p_i(y))^k$, and $\vdash \overline{u = p_i(y)} \rightarrow t'(\overline{v}, \overline{y}) = t(\overline{v}, \overline{u})$. (Notice that the case $|\overline{u}| = |\overline{y}| = 0$ accounts for the terms of L and that $\overline{u} \cap \overline{v} = \emptyset$.)

The following characterisation of atomic formulae of L' should also be clear, since the product extension adds no new predicate symbols, the other assertion again following from the substitutivity of equality.

Lemma 3.2 (Characterisation of atomic formulae of L') *Consider an atomic formula α' of L'. Then*

1. *either α' is an equality $y = y'$ between variables y and y' of P;*
2. *or else, $\alpha' = \alpha'(\overline{v}, \overline{y})$ is the result $\alpha(\overline{v}, \overline{p_i(y)})$ of replacing in atomic formula $\alpha(\overline{v}, \overline{u})$ of L every occurrence of a new variable u^k in $\overline{u}$ of sort S_i by $(p_i(y))^k$, and $\vdash \overline{u = p_i(y)} \rightarrow \alpha'(\overline{v}, \overline{y}) = \alpha(\overline{v}, \overline{u})$.*

We present a generalised version of the concept of eliminability when we extend our theory with a product sort. In this case we have to face the fact that we now have variables over the product sort. This can be overcome by means of a connection, which indicates how to translate variables ranging over the sort P as pairs of variables over sorts S_1 and S_2.

Consider sets of variables $\overline{y} = \{y^1, \ldots, y^n\}$ over sort P and $\overline{v}$ of sorts in L. By a (product) connection for $\overline{y}$ over $\overline{v}$ we mean sets of variables $\overline{x_1} = \{x_1^1, \ldots, x_1^n\}$ with x_1^k

[1] In the notation $t'(\overline{v}, \overline{y})$, $\overline{v}$ means that the variables of the language L that occur in the term t' are in the set $\overline{v}$ and $\overline{y}$ means that the variables of the language $L' - L$ that occur in the term t' are in the set $\overline{y}$.

of sort S_1 and $\overline{x_2} = \{x_2^1, \ldots, x_2^n\}$ with x_2^k of sort S_2 such that $\overline{v} \cap \overline{x_1} = \emptyset = \overline{v} \cap \overline{x_2}$. We also form the conjunction $K(\overline{y}, \overline{x_1}, \overline{x_2})$ of $p_1(y^k) = x_1^k \wedge p_2(y^k) = x_2^k$, for $k = 1, \ldots, n$. Call such conjunction $K(\overline{y}, \overline{x_1}, \overline{x_2})$ the **content** of the connection.

Such a (product) connection provides enough information for eliminating formulae of the larger language in favour of formulae of the smaller language. The proof of the next proposition provides a constructive procedure, by induction on the structure of formulae, for such elimination under connection.

Proposition 3.1 (Eliminability under connection) *Consider a (product) connection for $\overline{y}$ over $\overline{v}$ with content $K(\overline{y}, \overline{x_1}, \overline{x_2})$.*

For every formula $\rho(\overline{v}, \overline{y})$ of L' having as its free variables those in $\overline{v} \cup \overline{y}$, there exists a formula $\varphi(\overline{v}, \overline{x_1}, \overline{x_2})$ of L with free variables in $\overline{v} \cup \overline{x_1} \cup \overline{x_2}$, such that $T' \vdash K(\overline{y}, \overline{x_1}, \overline{x_2}) \rightarrow (\rho(\overline{v}, \overline{y}) \leftrightarrow \varphi(\overline{v}, \overline{x_1}, \overline{x_2}))$.

Call such a formula φ of L the **reflection** *of ρ into L (via K), and denote it by ρ^K.*

Two immediate consequences of this result are the usual eliminability (without connection) for the case of formulae without free variables over the new sort and the invariance under elementary equivalence.

Corollary 3.1 (Eliminability of formulae without new free variables) *For each formula $\rho(\overline{v})$ of L' without free variables of sort P, there exists a formula $\varphi(\overline{v})$ of L with the same free variables, such that $T' \vdash \rho(\overline{v}) \leftrightarrow \varphi(\overline{v})$.*

Proposition 3.2 (Invariance under elementary equivalence) *Given $\mathcal{A}'_1, \mathcal{A}'_2 \models T'$ if their reducts $\mathcal{A}_1, \mathcal{A}_2$ to L are elementarily equivalent then so are $\mathcal{A}'_1$ and $\mathcal{A}'_2$.*

4 Extensions by Subsorts behave like Definitions

Given theory T with language L we wish to add a new sort N to behave as the subsort of sort S of L according to a predicate r. In T we have that $T \vdash \exists x\, r(x)$.

Call $T' = T \cup Sbsrt$ the extension of T by introduction of the new sort, where $Sbsrt$ is the axiomatisation for subsort extension given in section 2.

We shall first show that we have a generalised eliminability: eliminability under connection. For this purpose we start by characterising the terms and atomic formulae of the larger language L'.

The following characterisation of terms of L' should be clear, since the subsort extension adds no new function (or constant) symbols having sort N as target, the other assertion following from the substitutivity property of equality.

Lemma 4.1 (Characterisation of terms of L') .

1. The terms of L' of sort N are the variables of sort N.

2. Each term $t'(\overline{v},\overline{y})$ of L' with sort in L is the result $t(\overline{v},\overline{j(y)})$ of replacing in term $t(\overline{v},\overline{u})$ of L every occurrence of a new variable u^k in $\overline{u}$ of sort S by $(j(y))^k$, and $\vdash \overline{u = j(y)} \rightarrow t'(\overline{v},\overline{y}) = t(\overline{v},\overline{u})$. (Notice that the case $|\overline{u}| = |\overline{y}| = 0$ accounts for the terms of L and that $\overline{u} \cap \overline{v} = \emptyset$.)

The following characterisation of atomic formulae of L' should also be clear, since the subsort extension adds no new predicate symbols, the other assertion again following from the substitutivity of equality.

Lemma 4.2 (Characterisation of atomic formulae of L') *Consider an atomic formula α' of L'. Then*

1. *either α' is an equality $y = y'$ between variables y and y' of N;*
2. *or else, $\alpha' = \alpha'(\overline{v},\overline{y})$ is the result $\alpha(\overline{v},\overline{j(y)})$ of replacing in atomic formula $\alpha(\overline{v},\overline{u})$ of L every occurrence of a variable u^k in $\overline{u}$ of sort S by $(j(y))^k$, and $\vdash \overline{u = j(y)} \rightarrow \alpha'(\overline{v},\overline{y}) = \alpha(\overline{v},\overline{u})$.*

We present a generalised version of the concept of eliminability when we extend our theory with a subsort. In this case, as in the case of the product, we have new variables over the subsort. The ideia is similar: we use a connection, which indicates how to translate variables ranging over the sort N as variables over sort S.

Consider sets of variables $\overline{y} = \{y^1, \ldots, y^n\}$ of sort N and $\overline{v}$ of sorts in L. By a (subsort) connection for $\overline{y}$ over $\overline{v}$ we mean a set $\overline{x} = \{x^1, \ldots, x^n\}$ of variables of sort S such that $\overline{v} \cap \overline{x} = \emptyset$. We also form the conjunction $K(\overline{y},\overline{x})$ of $j(y^k) = x^k$, for $k = 1, \ldots, n$. Call such conjunction $K(\overline{y},\overline{x})$ the **content** of the connection.

Such a (subsort) connection provides enough information for eliminating formulae of the larger language in favour of formulae of the smaller language, as stated in the next proposition, whose proof proceeds by induction on the structure of formulae.

Proposition 4.1 (Eliminability under connection) *Consider a (subsort) connection for $\overline{y}$ over $\overline{v}$ with content $K(\overline{y},\overline{x})$. For every formula $\rho(\overline{v},\overline{y})$ of L' having as its free variables those in $\overline{v} \cup \overline{y}$, there exists a formula $\varphi(\overline{v},\overline{x})$ of L with free variables in $\overline{v} \cup \overline{x}$, such that $T' \vdash K(\overline{y},\overline{x}) \rightarrow (\rho(\overline{v},\overline{y}) \leftrightarrow \varphi(\overline{v},\overline{x}))$. Call such a formula φ of L a* **reflection** *of ρ into L (via K), and denote it by ρ^K.*

Two immediate consequences of this result are the usual eliminability (without connection) for the case of formulae without free variables over the new sort and the invariance under elementary equivalence.

Corollary 4.1 (Eliminability of formulae without new free variables) *For each formula $\rho(\overline{v})$ of L' without free variables of sort N, there exists a formula $\varphi(\overline{v})$ of L with the same free variables, such that $T' \vdash \rho(\overline{v}) \leftrightarrow \varphi(\overline{v})$.*

Proposition 4.2 (Invariance under elementary equivalence) *Given $\mathcal{A}'_1, \mathcal{A}'_2 \models T'$ if their reducts $\mathcal{A}_1, \mathcal{A}_2$ to L are elementarily equivalent then so are $\mathcal{A}'_1$ and $\mathcal{A}'_2$.*

5 Extensions by Quotient Sorts behave like Definitions

Given theory T with language L we wish to add a new sort Q to behave as the quotient of sort S of L according to an equivalence relation q (i.e. given a relation symbol q, we must be able to prove in T that q is of equivalence).

Call $T' = T \cup Quot$ the extension of T by introduction of the new sort, where $Quot$ is the axiomatisation for quotient given in section 2.

We shall first show that we have a generalised eliminability: eliminability under connection. For this purpose we start by analysing the terms and atomic formulae of the larger language L'.

The following characterisation of terms of L' should be clear, since the quotient extension adds no new function (or constant) symbols having a sort in L as target.

Lemma 5.1 (Characterisation of terms of L') .

1. Each term of L is a term of L'.
2. Each term of L' of sort Q is
   - either a variable of sort Q
   - or $p(t)$ for some term t of L.

The following characterisation of atomic formulae of L' should also be clear, since the quotient extension adds no new predicate symbols.

Lemma 5.2 (Characterisation of atomic formulae of L') *Consider an atomic formula α' of L'.*

1. *Either α' is an equality of one of the forms:*
   - *$y = y'$ between variables y and y' of sort Q.*
   - *$y = p(t)$ or $p(t) = y$, with variable y over sort Q and term t of L with sort S.*
   - *$p(t) = p(t')$, for some terms t, t' of L with sort S.*
2. *Or else, α' is a atomic formula α of L, other than equality.*

We present a generalised version of the concept of eliminability similar to the case of subsort: a connection indicates how to translate variables ranging over the sort Q as variables over sort S.

Consider sets of variables $\overline{y} = \{y^1, \ldots, y^n\}$ of sort Q and $\overline{v}$ of sorts in L. A (quotient) connection for $\overline{y}$ over $\overline{v}$ consists of a set $\overline{x} = \{x^1, \ldots, x^n\}$ of variables of sort S such that $\overline{v} \cap \overline{x} = \emptyset$. We also form the conjunction $K(\overline{y}, \overline{x})$ of $j(y^k) = x^k$, for $k = 1, \ldots, n$. Call such conjunction $K(\overline{y}, \overline{x})$ the **content** of the connection.

As in the previous cases, such a (quotient) connection provides enough information for eliminating formulae of the larger language in favour of formulae of the smaller language. The (constructive) proof of the next proposition, by induction on the structure of formulae, provides such a elimination procedure.

Proposition 5.1 (Eliminability under connection) *Consider a (quotient) connection for $\overline{y}$ over $\overline{v}$ with content $K(\overline{y}, \overline{x})$.*

For every formula $\rho(\overline{v}, \overline{y})$ of L' having as its free variables those in $\overline{v} \cup \overline{y}$, there exists a formula $\varphi(\overline{v}, \overline{x})$ of L with free variables in $\overline{v} \cup \overline{x}$, such that $T' \vdash \forall \overline{x}(K(\overline{y}, \overline{x}) \rightarrow (\rho(\overline{v}, \overline{y}) \leftrightarrow \varphi(\overline{v}, \overline{x})))$. Call such a formula φ of L a **reflection** *of ρ into L (via K), and denote it by ρ^K.*

Immediate consequences of this result are the usual eliminability (without connection) for formulae without free variables over the new sort and the invariance under elementary equivalence.

Corollary 5.1 (Eliminability of formulae without new free variables) *For each formula $\rho(\overline{v})$ of L' without free variables of sort Q, there exists a formula $\varphi(\overline{v})$ of L with the same free variables, such that $T' \vdash \rho(\overline{v}) \leftrightarrow \varphi(\overline{v})$.*

Proposition 5.2 (Invariance under elementary equivalence) *Given $\mathcal{A}'_1, \mathcal{A}'_2 \models T'$ if their reducts $\mathcal{A}_1, \mathcal{A}_2$ to L are elementarily equivalent then so are $\mathcal{A}'_1$ and $\mathcal{A}'_2$.*

6 Extensions by Discriminated Union Sorts behave like Definitions

Given theory T with language L we wish to add a new sort U to behave as the discriminated union of sorts S_1 and S_2 of L.

Call $T' = T \cup D_Uni$ the extension of T by introduction of the new sort, where D_Uni is the axiomatisation for product extension given in section 2.

We shall first show that we have a generalised eliminability: eliminability under connection. For this purpose we start by analysing the terms and atomic formulae of the larger language L'.

The following characterisation of terms of L' should be clear, since the union extension adds no new function (or constant) symbols having a sort in L as target.

Lemma 6.1 (Characterisation of terms of L') .

1. Each term of L is a term of L'.
2. Each term of L' of sort U is
   - either a variable of sort U
   - or $i_l(t)$ for some term t of sort S_l of L where $l = 1, 2$.

The following characterisation of atomic formulae of L' should also be clear, since the union extension adds no new predicate symbols.

Lemma 6.2 (Characterisation of atomic formulae of L') *Consider an atomic formula α' of L'.*

1. *Either α' is an equality of one of the forms:*
   - *$y = y'$ between variables y and y' of sort U.*
   - *$y = i_l(t)$ or $i_l(t) = y$, with variable y over sort U and term t of L with sort S_l where $l = 1, 2$.*
   - *$i_l(t) = i_l(t')$, for some terms t, t' of L with sort S_l where $l = 1, 2$.*
2. *Or else, α' is a atomic formula α of L, other than equality.*

We present a generalised version of the concept of eliminability similar to the case of product: a connection indicates how to translate variables ranging over the sort U as, sometimes, variables over sort S_1 or, sometimes, over sort S_2.

Consider sets of variables $\overline{y} = \{y^1, \ldots, y^n\}$ of sort U and $\overline{v}$ of sorts in L. A (union) connection for $\overline{y}$ over $\overline{v}$ consists of sets $\overline{x_1} = \{x_1^1, \ldots, x_1^n\}$ of variables of sort S_1 and $\overline{x_2} = \{x_2^1, \ldots, x_2^n\}$ of variables of sort S_2 such that $\overline{v} \cap \overline{x_1} = \emptyset = \overline{v} \cap \overline{x_1}$. For each function $\delta : \{1, \ldots, n\} \to \{1, 2\}$ we form the conjuntion $K^\delta(\overline{y}, \overline{x_1}, \overline{x_2}) = \wedge_{j=1}^n (y_j = i^j_{\delta(j)}(x^j_{\delta(j)}))$ and the set $\overline{y}^\delta = \{x^1_{\delta(1)}, \ldots, x^n_{\delta(n)}\}$. Call such conjunction $K^\delta(\overline{y}, \overline{x_1}, \overline{x_2})$ and the set $\overline{y}^\delta$ the δ-**content** of the connection.

As in the previous cases, such a (union) connection provides enough information for eliminating formulae of the larger language in favour of formulae of the smaller language. The (constructive) proof of the next proposition, by induction on the structure of formulae, provides such a elimination procedure.

Proposition 6.1 (Eliminability under connection) *Given a (union) connection for $\overline{y}$ over $\overline{v}$ with δ-content $K^\delta(\overline{y}, \overline{x_1}, \overline{x_2})$ and $\overline{y}^\delta$.*

*Given a formula $\rho(\overline{v}, \overline{y})$ of L' having as its free variables those of $\overline{v} \cup \overline{y}$, there exist 2^n formulae $\varphi(\overline{v}, \overline{y}^\delta)$ of L with free variables of $\overline{v} \cup \overline{y}^\delta$, such that $T' \vdash K^\delta \to (\rho(\overline{v}, \overline{y}) \leftrightarrow \varphi(\overline{v}, \overline{y}^\delta))$. Call such a formula φ of L the δ-**reflection of** ρ into L (via K), and denote it by ρ^{K^δ}.*

Two immediate consequences of this result are the usual eliminability (without connection) for the case of formulae without free variables over the new sort and the invariance under elementary equivalence.

Corollary 6.1 (Eliminability of formulae without new free variables) *For each formula $\rho(\overline{v})$ of L' without free variables of sort U, there exists a formula $\varphi(\overline{v})$ of L with the same free variables, such that $T' \vdash \rho(\overline{v}) \leftrightarrow \varphi(\overline{v})$.*

Proposition 6.2 (Invariance under elementary equivalence) *Given $\mathcal{A}'_1, \mathcal{A}'_2 \models T'$ if their reducts $\mathcal{A}_1, \mathcal{A}_2$ to L are elementarily equivalent then so are $\mathcal{A}'_1$ and $\mathcal{A}'_2$.*

7 Conclusions

We have examined four sort introducing constructs: cartesian product, subsort, discriminated union and quotient.

Introduction of new sorts is of interest because it occurs often in implementing formal specifications. In constrast to the introduction of new function or predicate symbols [Sho67], the case of new sorts is not extensively studied.

The four constructs examined correspond to natural constructions of new sets. Three of them have found their way as data structuring constructs in programming languages [Hoa74]. The quotient construct is of importance in developing and verifying correct implementations [TM87, Vel87].

In section 2 we have remembered (from [MV92]) formal specifications for these four constructs by axiomatising their usual descriptions as extensions of specifications. In each case, in addition to the new sort, we also introduce appropriate conversion functions, connecting it to the old sorts.

In [MV92] we proved that, in each case, the new sort is unique, but only up to isomorphism. This contrasts with the case of introduction of new function or predicate symbols by definitions, when we have unique expandability of models. Also, these more familiar extensions by definitions exhibit eliminability: any formula of the extended language is equivalent to some formula in the original language. In our case of introduction of new sorts one should not expect full eliminability, because of the new variables ranging over the sort introduced. Nevertheless, a kind of generalised eliminability is still possible, as we have proved in the propositions called: eliminability under connection.

Actually, along this paper we have established, for each sort introduction construct, the following properties:

1. characterisation of terms and atomic formulae of the new sort;
2. eliminability under connection;
3. invariance under elementary equivalence;

The following properties will be presented in a forthcoming paper:

1. generalized interpretation to formulae with new free variables;
2. relationship between the relation defined by a formula in the new language and the relation defined by its ref l ection;
3. definability of a relation in the extension and its internalised version.

References

[End72] H.B. Enderton. *A Mathematical Introduction to Logic*. Academic Press, Inc., New York, 1972.

[Eri82] H.D. Erich. On the theory of specification, implementation and parametrization of abstract data types. *Journal of the Association for Computing Machinery*, 29(1):206–227, January 1982.

[Hoa74] C.A.R. Hoare. Notes on data structuring. In O.J. Dahl, E.W. Dijkstra, and C.A.R. Hoare, editors, *Structured Programming*, pages 83–174. Academic Press, 1974.

[MV92] M. C. Meré and P. A. Veloso. On Extensions by Sorts. *Monografias em Ciências da Computação, D.I., PUC-Rio*, 38/92, December 1992.

[MV94] M.C. Meré and P.A. Veloso. Definition-like extensions by sorts. *Anáis do Workshop on Logic, Language, Information and Computation '94*, Julho de 1994, Recife, Brasil.

[VM94] P.A. Veloso and M.C. Meré. Construtores Parametrizados de Sortes. *Monografias em Ciências da Computação, D.I., PUC-Rio*, 14/94, Julho 1992.

[Sho67] J.R. Shoenfield. *Mathematical Logic*. Addison-Wesley Publishing Company, 1967.

[TM87] W.M. Turski and T.S.E. Maibaum. *The Specification of Computer Programs*. Addison-Wesley Publishing Company, 1987.

[Vel87] P.A.S Veloso. *Estruturação e Verificação de Programas com Tipos de Dados*. Editora E. Blücher ltda., 1987.

A APPENDIX

Eliminability Results for Product Sort

In this appendix we prove the eliminability results for extensions by introduction of product sorts.

Lemma A.1 (Characterisation of terms of L') .

1. The terms of L' of sort P are the variables of sort P.
2. Each term $t'(\overline{v}, \overline{y})$ of L' with sort in L is the result $t(\overline{v}, \overline{p_i(y)})$ of replacing in term $t(\overline{v}, \overline{u})$ of L every occurrence of a new variable u^k in $\overline{u}$ of sort S_i by $(p_i(y))^k$, and $\vdash \overline{u = p_i(y)} \rightarrow t'(\overline{v}, \overline{y}) = t(\overline{v}, \overline{u})$. (Notice that the case $|\overline{u}| = |\overline{y}| = 0$ accounts for the terms of L and that $\overline{u} \cap \overline{v} = \emptyset$.)

Proof: That these are all the terms of L' should be clear, because the product extensin adds no new function (or constant) symbols having sort P as target. The second assertion in item 2 follows from the substitutivity property of equality: if $t(v, s)$ is the result of replacing in term $t(v, u)$ every occurrence of a new variable u by term s, then $\vdash u = s \rightarrow t(v, u) = t(v, s)$. More precisely, we can show by induction that the set D of terms described in items 1 and 2 contains all the variables and is closed under the term-forming operations of L'. Indeed, let Trm and Trm' be the set of terms of L and L', respetively.

1. Every variable of L' is in D. For, a variable $y : P$ is in D by item 1. And a variable $v : S$, with S in L, is in D by item 2.
2. Consider $t' \in$ Trm' of the form $f(d_1, \ldots, d_n)$, with $d_1, \ldots, d_n \in D$.
   - If some d_k is a variable $y : P$, then f is some projection p_l and $n = 1$. Thus, t' is $p_l(y)$, which is in D, by item 1 above and item 2 of the lemma.
   - If no d_k is a variable of sort N, then f cannot be a projection, so f must be a symbol of L. For each $k = 1, \ldots, n$, since $d_k \in D$, d_k is of the form $t_k(\overline{v_k}, \overline{p(y_k)})$ for some term $t_k(\overline{v_k}, \overline{u})$ of L. We can select the t_k's and $\overline{u_k}$'s so that different $\overline{u_k}$'s have no variable in common. Then, by taking $t(\overline{v}, \overline{u})$ as $f(t_1, \ldots, t_n)$, t' is $t(\overline{v}, \overline{p(y)})$.

QED

Lemma A.2 (Characterisation of atomic formulae of L') *Consider an atomic formula α' of L'. Then*

1. *either α' is an equality $y = y'$ between variables y and y' of P;*
2. *or else, $\alpha' = \alpha'(\overline{v}, \overline{y})$ is the result $\alpha(\overline{v}, \overline{p_i(y)})$ of replacing in atomic formula $\alpha(\overline{v}, \overline{u})$ of L every occurrence of a variable u^k in $\overline{u}$ of sort S_i by $(p_i(y))^k$, and $\vdash \overline{u = p_i(y)} \rightarrow \alpha'(\overline{v}, \overline{y}) = \alpha(\overline{v}, \overline{u})$.*

Proof: That these are exactly the atomic formulae of L' follows easily from lemma A.1, since the extension adds no new predicate symbols. The second assertion in item 2 follows from the substitutivity property of equality: if $R(v,s)$ is the result of replacing in atomic formula $R(v,u)$ every occurrence of a new variable u by term s, then $\vdash u = s \rightarrow R(v,u) = R(v,s)$.

Indeed, consider an atomic formula α' of L'. Then, α' is $R(t'_1,\ldots,t'_n)$, where $t'_1,\ldots,t'_n$ are terms of L' and R is either a predicate symbol of L' (so, of L) or equality $=$ or a predicate symbol of L' (so, of L).

- If no t'_k has a symbol of L, then by lemma A.1, they are variables of sort P. Then, R must be $=$. Thus, α' is $y = y'$, for variables y, y' of P.
- Otherwise, by lemma A.1, each term t'_k is of the form $t_k(\overline{v_k}, \overline{p_i(y_k)})$ obtained by substitution into term $t_k(\overline{v_k}, \overline{u_k})$ of L. As in the proof of lemma A.1 we can assume that different sets $\overline{u_k}$ have no variable in common. Take $\alpha(\overline{v},\overline{u})$ as $R(t_1,\ldots,t_n)$. Then, α' is the result $\alpha(\overline{v},\overline{p_i(y)})$ of the substitution into $\alpha(\overline{v},\overline{u})$.

QED

Recall the concept of (product) connection. Consider sets of variables $\overline{y} = \{y^1,\ldots,y^n\}$ over sort P and $\overline{v}$ of sorts in L. By a (product) connection for $\overline{y}$ over $\overline{v}$ we mean sets of variables $\overline{x_1} = \{x_1^1,\ldots,x_1^n\}$ with x_1^k of sort S_1 and $\overline{x_2} = \{x_2^1,\ldots,x_2^n\}$ with x_2^k of sort S_2 such that $\overline{v} \cap \overline{x_1} = \emptyset = \overline{v} \cap \overline{x_2}$. We also form the conjunction $K(\overline{y},\overline{x_1},\overline{x_2})$ of $p_1(y^k) = x_1^k \wedge p_2(y^k) = x_2^k$, for $k = 1,\ldots,n$. Call such conjunction $K(\overline{y},\overline{x_1},\overline{x_2})$ the **content** of the connection.

Proposition A.1 (Eliminability under connection) *Consider a (product) connection for $\overline{y}$ over $\overline{v}$ with content $K(\overline{y},\overline{x_1},\overline{x_2})$.*

For every formula $\rho(\overline{v},\overline{y})$ of L' having as its free variables those in $\overline{v}\cup\overline{y}$, there exists a formula $\varphi(\overline{v},\overline{x_1},\overline{x_2})$ of L with free variables in $\overline{v}\cup\overline{x_1}\cup\overline{x_2}$, such that $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (\rho(\overline{v},\overline{y}) \leftrightarrow \varphi(\overline{v},\overline{x_1},\overline{x_2}))$.

Call such a formula φ of L the **reflection** *of ρ into L (via K), and denote it by ρ^K.*

Proof: By induction on the structure of ρ.

- Basis: ρ is atomic. By lemma A.2, we have two cases to consider.

  1. ρ is $y = y'$. Then, take φ as $x_1 = x'_1 \wedge x_2 = x'_2$. Clearly, $T' \vdash y = y' \leftrightarrow p_1(y) = p_1(y') \wedge p_2(y) = p_2(y')$.
     Since $y \in \overline{y}$, we have $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (p_1(y) = p_1(y') \wedge p_2(y) = p_2(y'))$. And similarly for y'. Hence, $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (y = y' \leftrightarrow x_1 = x'_1 \wedge x_2 = x'_2)$.
     Consider the (atomic) formula $\varphi(\overline{v},\overline{x_i})$ of L obtained by replacing in $\alpha(\overline{v},\overline{u})$ each occurrence of variable u^k of sorte S_i in $\overline{u}$ by x_i^k (of sort S_i). By lemma A.1 we have $\vdash \overline{x_i = p_i(y)} \rightarrow (\rho(\overline{v},\overline{y}) \leftrightarrow \varphi(\overline{v},\overline{x_i}))$.
     We can rearrange the set of free variables of sorts S_1 and S_2 of $\varphi(\overline{v},\overline{x_i})$ as $\varphi(\overline{v},\overline{x_1},\overline{x_2})$. Clearly, we have $\vdash \overline{x_i = p_i(y)} \rightarrow (\rho(\overline{v},\overline{y}) \leftrightarrow \varphi(\overline{v},\overline{x_1},\overline{x_2}))$.

Observe that in $\rho(\overline{v},\overline{y})$ we use some (or possible all) of the projections of the variables in the set $\overline{y}$ then, we have that $\vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow \overline{x_i = p_i(y)}$. Hence, $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (\rho \leftrightarrow \varphi)$.

- $\rho(\overline{v},\overline{y})$ is a negation $\neg\rho'(\overline{v},\overline{y})$. The inductive hypothesis gives a formula $\varphi'(\overline{v},\overline{x_1},\overline{x_2})$ of L, such that $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (\rho' \leftrightarrow \varphi')$. So, we take $\varphi(\overline{v},\overline{x_1},\overline{x_2})$ as $\neg\varphi'(\overline{v},\overline{x_1},\overline{x_2})$.
- $\rho(\overline{v},\overline{y})$ is a conjunction $\rho'(\overline{v'},\overline{y'}) \wedge \rho''(\overline{v''},\overline{y''})$. The inductive hypothesis gives formulae $\varphi'(\overline{v'},\overline{x_1'},\overline{x_2'})$ and $\varphi''(\overline{v''},\overline{x_1''},\overline{x_2''})$ of L such that $T' \vdash K(\overline{y'},\overline{x_1'},\overline{x_2'}) \rightarrow (\rho' \leftrightarrow \varphi')$ and $T' \vdash K(\overline{y''},\overline{x_1''},\overline{x_2''}) \rightarrow (\rho'' \leftrightarrow \varphi'')$.
  Take $\overline{x_1} = \overline{x_1'} \cup \overline{x_1''}$, $\overline{x_2} = \overline{x_2'} \cup \overline{x_2''}$ and $\overline{y} = \overline{y'} \cup \overline{y''}$. Clearly, $\vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow K(\overline{y'},\overline{x_1'},\overline{x_2'}) \wedge K(\overline{y''},\overline{x_1''},\overline{x_2''})$. Hence, $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (\rho' \wedge \rho'' \leftrightarrow \varphi' \wedge \varphi''))$.
- ρ is a universal formula. We examine two subcases, depending on the sort of the variable quantified upon[2].
  - Quantification on a variable of sort in L: $\rho(\overline{v},\overline{y})$ is $(\forall v : S_i)\rho'(v,\overline{v},\overline{y})$. The inductive hypothesis gives a formula $\varphi'(v,\overline{v},\overline{x_1},\overline{x_2})$ of L, such that $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (\rho'(v,\overline{v},\overline{y}) \leftrightarrow \varphi'(v,\overline{v},\overline{x_1},\overline{x_2}))$. Since v does not occur (free) in $K(\overline{y},\overline{x_1},\overline{x_2})$ we have $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (\forall v : S_i)(\rho'(v,\overline{v},\overline{y}) \leftrightarrow \varphi'(v,\overline{v},\overline{x_1},\overline{x_2}))$. Hence $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow ((\forall v : S_i)\rho'(v,\overline{v},\overline{y}) \leftrightarrow (\forall v : S_i)\varphi'(v,\overline{v},\overline{x_1},\overline{x_2}))$ and we take $\varphi(\overline{v},\overline{x_1},\overline{x_2})$ as $(\forall v : S_i)\varphi'(v,\overline{v},\overline{x_1},\overline{x_2})$.
  - Quantification on a variable $y : P$: $\rho(\overline{v},\overline{y})$ is $(\forall y : P)\rho'(\overline{v},\overline{y},y)$. We can select two new variables x_1 and x_2 not occurring in $\overline{v} \cup \overline{x_1} \cup \overline{x_2}$ to construct a connection for ρ'. The inductive hypothesis gives a formula $\varphi'(\overline{v},\overline{x_1},\overline{x_2},x_1,x_2)$ of L, such that $T' \vdash K(\overline{y},y,\overline{x_1},\overline{x_2},x_1,x_2) \rightarrow (\rho'(\overline{v},\overline{y},y) \leftrightarrow \varphi'(\overline{v},\overline{x_1},\overline{x_2},x_1,x_2))$. Clearly, $K(\overline{y},\overline{x_1},\overline{x_2}) \wedge K(y,x_1,x_2) \rightarrow K(\overline{y},y,\overline{x_1},\overline{x_2},x_1,x_2)$. Thus, $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (K(y,x_1,x_2) \rightarrow (\rho' \leftrightarrow \varphi'))$. So, since y,x_1,x_2 do not occur (free) in $K(\overline{y},\overline{x_1},\overline{x_2})$, we have $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow (\forall y : P)(\forall x_1 : S_1)(\forall x_2 : S_2)(K(y,x_1,x_2) \rightarrow (\rho' \leftrightarrow \varphi'))$. But clearly, $T' \vdash (\forall y : P)(\exists x_1 : S_1)(\exists x_2 : S_2)K(y,x_1,x_2)$ and $T' \vdash (\forall x_1 : S_1)(\forall x_2 : S_2)(\exists y : P)K(y,x_1,x_2)$. Hence
    $T' \vdash K(\overline{y},\overline{x_1},\overline{x_2}) \rightarrow ((\forall y : P)\rho' \leftrightarrow (\forall x_1 : S_1)(\forall x_2 : S_2)\varphi')$.
    Indeed, this last step becomes clear in a Natural Deduction style.

$$\dfrac{\dfrac{\dfrac{T'}{\forall y \exists x_1 \exists x_2\ K(y,x_1,x_2)}\qquad \dfrac{\dfrac{T' \qquad K(\overline{y},\overline{x}_1,\overline{x}_2)}{K(y,x_1,x_2) \rightarrow (\varphi' \rightarrow \rho')} \qquad K(y,x_1,x_2)}{\varphi' \rightarrow \rho'} \qquad \dfrac{\forall x_1 \forall x_2\, \varphi'}{\varphi'}}{\exists x_1 \exists x_2\ K(y,x_1,x_2)} \qquad \rho'}{\rho'}}{\forall y \rho'}$$

[2]Notice that, by construction, the free variables in $K(\overline{y},\overline{x_1},\overline{x_2})$ are those in $\overline{y} \cup \overline{x_1} \cup \overline{x_2}$.

$$\dfrac{\dfrac{\dfrac{T'}{\forall x_1\,\forall x_2\,\exists y\,K(y,x_1,x_2)}\qquad \dfrac{\dfrac{T'\qquad K(\bar{y},\bar{x}_1,\bar{x}_2)}{K(y,x_1,x_2)\to(\rho'\to\varphi')}\qquad K(y,x_1,x_2)}{\rho'\to\varphi'}\qquad \dfrac{\forall y\rho'}{\rho'}}{\exists y\,K(y,x_1,x_2)}\qquad \dfrac{\varphi'}{\varphi'}}{\forall x_1\,\forall x_2\,\varphi'}$$

QED

Corollary A.1 (Eliminability of formulae without new free variables) *For each formula $\rho(\overline{v})$ of L' without free variables of sort P, there exists a formula $\varphi(\overline{v})$ of L with the same free variables, such that $T' \vdash \rho(\overline{v}) \leftrightarrow \varphi(\overline{v})$.*

Proof: This follows from eliminability under connection. Since $|\overline{y}| = 0$ we also have $|\overline{x_1}| = |\overline{x_2}| = 0$. So, we have a trivial connection with empty conjunction as its content.

QED

Proposition A.2 (Invariance under elementary equivalence) *Given $\mathcal{A}'_1, \mathcal{A}'_2 \models T'$ if their reducts $\mathcal{A}_1, \mathcal{A}_2$ to L are elementarily equivalent then so are $\mathcal{A}'_1$ and $\mathcal{A}'_2$.*

Proof: This follows from the preceding corollary. Indeed, for each sentence τ of L' we have a sentence σ of L such that $T' \vdash \tau \leftrightarrow \sigma$. Thus, for each model $\mathcal{A}'$ of T' with reduct $\mathcal{A} = \mathcal{A}'/L$, we have $\mathcal{A}' \models \tau$ iff $\mathcal{A}' \models \sigma$ iff $\mathcal{A} \models \sigma$. Hence, if $\mathcal{A}_1 \equiv \mathcal{A}_2$ then $\mathcal{A}'_1 \equiv \mathcal{A}'_2$.

QED

B APPENDIX

Eliminability Results for Subsort

In this appendix we prove the eliminability results for extensions by introduction of subsorts.

Lemma B.1 (Characterisation of terms of L') .

1. The terms of L' of sort N are the variables of sort N .
2. Each term $t'(\overline{v}, \overline{y})$ of L' with sort in L is the result $t(\overline{v}, \overline{j(y)})$ of replacing in term $t(\overline{v}, \overline{u})$ of L every occurrence of a new variable u^k in $\overline{u}$ of sort S by $(j(y))^k$, and $\vdash \overline{u = j(y)} \rightarrow t'(\overline{v}, \overline{y}) = t(\overline{v}, \overline{u})$. (Notice that the case $|\overline{u}| = |\overline{y}| = 0$ accounts for the terms of L and that $\overline{u} \cap \overline{v} = \emptyset$.)

Proof: That these are all terms of L' is clear because the extension adds no new function or constant symbol having N as target. More precisely, we can show that the set D of terms described above constains all the variables and is closed under the term forming operations of L'.

Indeed, let Trm and Trm' be the set of terms of L and L', respetively.

1. Every variable of L' is in D. For a variable $y : N$ is in D by item 1. And a variable $v : S$, with S in L, is in D by item 2.
2. Consider $t' \in$ Trm' of the form $f(d_1, \ldots, d_n)$, with $d_1, \ldots, d_n \in D$.
   - If some d_k is a variable $y : N$, then f is j and $n = 1$. Thus, t' is $j(y)$, which is in D, by the item 1 above and item 2 of the lemma.
   - If no d_k is a variable of sort N, then f cannot be j, so f must be a symbol of L. For each $k = 1, \ldots, n$, since $d_k \in D$, d_k is of the form $t_k(\overline{v_k}, \overline{j(y_k)})$ for some term $t_k(\overline{v_k}, \overline{u})$ of L. We can select the t_k's and $\overline{u_k}$'s so that different $\overline{u_k}$'s have no variable in common. Then, by taking $t(\overline{v}, \overline{u})$ as $f(t_1, \ldots, t_n)$, t' is $t(\overline{v}, \overline{j(y)})$.

The other assertion follows by induction on the structure of t:

- Basis: t is a variable u of sort S then, clearly $\vdash u = j(y) \rightarrow j(y) = u$
- Inductive step: t is a term $f(t_1, \ldots, t_n)$, so $t(\overline{v}, \overline{u})$ is $f(t_1(\overline{v_1}, \overline{u_1}), \ldots, t_n(\overline{v_n}, \overline{u_n}))$, where we can assume that different sets $\overline{u_k}$ have no variable in common. For each $k = 1, \ldots, n$, the inductive hypothesis gives $\vdash \overline{u_k = j(y_k)} \rightarrow t'_k = t_k$, and clearly $\vdash \overline{u = j(y)} \rightarrow \overline{u_k = j(y_k)}$ for each k. Hence $\vdash \overline{u = j(y)} \rightarrow f(t'_1, \ldots, t'_n) = f(t_1, \ldots, t_n)$

QED

Lemma B.2 (Characterisation of atomic formulae of L') *Consider an atomic formula α' of L'. Then*

1. *either α' is an equality $y = y'$ between variables y and y' of N;*
2. *or else, $\alpha' = \alpha'(\overline{v}, \overline{y})$ is the result $\alpha(\overline{v}, \overline{j(y)})$ of replacing in atomic formula $\alpha(\overline{v}, \overline{u})$ of L every occurrence of a variable u^k in $\overline{u}$ of sort S by $(j(y))^k$, and $\vdash \overline{u = j(y)} \rightarrow \alpha'(\overline{v}, \overline{y}) = \alpha(\overline{v}, \overline{u})$.*

Proof: That these are exactly the atomic formulae of L' follows easily from lemma B.1, since the extension adds no new predicate symbols.

Indeed, consider an atomic formula α' of L'. Then, α' is $R(t'_1, \ldots, t'_n)$, where $t'_1, \ldots, t'_n$ are terms of L' and R is either $=$ or a predicate symbol of L' (so, of L).

- If no t'_k has a symbol of L, then by lemma B.1, they are variables of sort N. Then, R must be $=$. Thus, α' is $y = y'$, for variables y, y' of N.
- Otherwise, by lemma B.1, each term t'_k is of the form $t_k(\overline{v_k}, \overline{j(y_k)})$ obtained by substitution into term $t_k(\overline{v_k}, \overline{u_k})$ of L. As in the proof of lemma B.1 we can assume that different sets $\overline{u_k}$ have no variable in common.

  Take $\alpha(\overline{v}, \overline{u})$ as $R(t_1, \ldots, t_n)$. Then, α' is the result $\alpha(\overline{v}, \overline{j(y)})$ of the substitution into $\alpha(\overline{v}, \overline{u})$.

  Now, the other assertion follows immediately from the corresponding one in lemma B.1. Indeed, for $k = 1, \ldots, n$ lemma B.1 gives $\vdash \overline{u_k = j(y_k)} \rightarrow t'_k = t_k$. Since clearly $\vdash \overline{u = j(y)} \rightarrow \overline{u_k = j(y_k)}$ we get $\vdash \overline{u = j(y)} \rightarrow R(t'_1, \ldots, t'_n) \leftrightarrow R(t_1, \ldots, t_n))$.

QED

Recall the concept of (subsort) connection. Consider sets of variables $\overline{y} = \{y^1, \ldots, y^n\}$ of sort N and $\overline{v}$ of sorts in L. By a (subsort) connection for $\overline{y}$ over $\overline{v}$ we mean a set $\overline{x} = \{x^1, \ldots, x^n\}$ of variables of sort S such that $\overline{v} \cap \overline{x} = \emptyset$. We also form the conjunction $K(\overline{y}, \overline{x})$ of $j(y^k) = x^k$, for $k = 1, \ldots, n$. Call such conjunction $K(\overline{y}, \overline{x})$ the **content** of the connection.

Proposition B.1 (Eliminability under connection) *Consider a (subsort) connection for $\overline{y}$ over $\overline{v}$ with content $K(\overline{y}, \overline{x})$. For every formula $\rho(\overline{v}, \overline{y})$ of L' having as its free variables those in $\overline{v} \cup \overline{y}$, there exists a formula $\varphi(\overline{v}, \overline{x})$ of L with free variables in $\overline{v} \cup \overline{x}$, such that $T' \vdash K(\overline{y}, \overline{x}) \rightarrow (\rho(\overline{v}, \overline{y}) \leftrightarrow \varphi(\overline{v}, \overline{x}))$. Call such a formula φ of L a* **reflection of** *ρ into L (via K), and denote it by ρ^K.*

Proof: by induction on the structure of ρ.

- Basis: ρ is atomic. By lemma B.2, we have two cases to consider.
  1. ρ is $y = y'$. Then, take φ as $x = x'$. Clearly, $T' \vdash y = y' \leftrightarrow j(y) = j(y')$. Hence, $T' \vdash K(y, y', x, x') \rightarrow (y = y' \leftrightarrow x = x')$ since $K(y, y', x, x')$ is $j(y) = x \wedge j(y') = x'$
  2. $\rho(\overline{v}, \overline{y})$ is $\alpha(\overline{v}, \overline{j(y)})$ for some atomic formula $\alpha(\overline{v}, \overline{u})$ of L. Consider the (atomic) formula $\varphi(\overline{v}, \overline{x})$ pf L, obtained by replacing in $\alpha(\overline{v}, \overline{u})$ each occurrence of variable u^k in $\overline{u}$ by x^k.

If we now replace in $\varphi(\overline{v},\overline{x})$ each occurrence of x^k by $j(y^k)$, we obtain $\alpha(\overline{v},\overline{j(y)})$ back. So, ρ is $\varphi(\overline{v},\overline{j(y)})$. By lemma B.1 we have $\vdash \overline{x = j(y)} \to (\varphi \leftrightarrow \rho)$. Clearly, $\vdash K(\overline{y},\overline{x}) \to x = j(y)$. hence, $T' \vdash K(\overline{y},\overline{x}) \to (\varphi \leftrightarrow \rho)$.

- $\rho(\overline{v},\overline{y})$ is a negation $\neg\rho'(\overline{v},\overline{y})$. The inductive hypothesis gives a formula $\varphi'(\overline{v},\overline{x})$ of L, such that $T' \vdash K(\overline{y},\overline{x}) \to (\rho' \leftrightarrow \varphi')$. So, we take $\varphi(\overline{v},\overline{x})$ as $\neg\varphi'(\overline{v},\overline{x})$.
- $\rho(\overline{v},\overline{y})$ is a conjunction $\rho'(\overline{v'},\overline{y'}) \wedge \rho''(\overline{v''},\overline{y''})$. The inductive hypothesis gives formulae $\varphi'(\overline{v'},\overline{x'})$ and $\varphi''(\overline{v''},\overline{x''})$ of L such that $T' \vdash K(\overline{y'},\overline{x'}) \to (\rho' \leftrightarrow \varphi')$ and $T' \vdash K(\overline{y''},\overline{x''}) \to (\rho'' \leftrightarrow \varphi'')$.

  Take $\overline{x} = \overline{x'} \cup \overline{x''}$ and $\overline{y} = \overline{y'} \cup \overline{y''}$. Clearly, $\vdash K(\overline{y},\overline{x}) \to K(\overline{y'},\overline{x'}) \wedge K(\overline{y''},\overline{x''})$. Hence, $T' \vdash K(\overline{y},\overline{x}) \to (\rho' \wedge \rho'' \leftrightarrow \varphi' \wedge \varphi''))$.
- ρ is a universal formula. We examine two subcases, depending on the sort of the variable quantified upon[3].
  - Quantification on a variable of sort in L, that is $\rho(\overline{v},\overline{y})$ is $(\forall v : S)\rho'(v,\overline{v},\overline{y})$. The inductive hypothesis gives a formula $\varphi'(v,\overline{v},\overline{x})$ of L, such that $T' \vdash K(\overline{y},\overline{x}) \to (\rho'(v,\overline{v},\overline{y}) \leftrightarrow \varphi'(v,\overline{v},\overline{x}))$. Since v does not occur (free) in $K(\overline{y},\overline{x})$ we have $T' \vdash K(\overline{y},\overline{x}) \to (\forall v : S)(\rho'(v,\overline{v},\overline{y}) \leftrightarrow \varphi'(v,\overline{v},\overline{x}))$. Hence $T' \vdash K(\overline{y},\overline{x}) \to ((\forall v : S)\rho'(v,\overline{v},\overline{y}) \leftrightarrow (\forall v : S)\varphi'(v,\overline{v},\overline{x}))$ and we take $\varphi(\overline{v},\overline{x})$ as $(\forall v : S)\varphi'(v,\overline{v},\overline{x})$.
  - Quantification on a variable $y : N$, that is $\rho(\overline{v},\overline{y})$ is $(\forall y : N)\rho'(\overline{v},\overline{y},y)$. The inductive hypothesis gives a formula $\varphi'(\overline{v},\overline{x},x)$ of L, such that $T' \vdash K(\overline{y},y,\overline{x},x) \to (\rho'(\overline{v},\overline{y},y) \leftrightarrow \varphi'(\overline{v},\overline{x},x))$. Clearly, $K(\overline{y},\overline{x}) \wedge K(y,x) \to K(\overline{y},y,\overline{x},x)$. Thus, $T' \vdash K(\overline{y},\overline{x}) \to (K(y,x) \to (\rho' \leftrightarrow \varphi'))$. So, since y,x do not occur (free) in $K(\overline{y},\overline{x})$, we have $T' \vdash K(\overline{y},\overline{x}) \to (\forall y : N)(\forall x : S)(K(y,x) \to (\rho' \leftrightarrow \varphi'))$. But clearly, $T' \vdash (\forall y : N)(\exists x : S)(r(x) \wedge K(y,x))$ and $T' \vdash (\forall x : S)(r(x) \to (\exists y : N)K(y,x))$. Hence
    $T' \vdash K(\overline{y},\overline{x}) \to ((\forall y : N)\rho' \leftrightarrow (\forall x : S)\varphi')$.

    Indeed, this last step becomes clear in a Natural Deduction style.

$$\dfrac{\dfrac{\dfrac{T'}{\forall x(r(x) \to \exists y K(y,x))}}{r(x) \to \exists y K(y,x)}\ r(x)}{\exists y K(y,x)} \qquad \dfrac{\dfrac{\dfrac{T' \quad K(\overline{y},\overline{x})}{K(y,x) \to (\rho' \to \varphi')}\ K(y,x)}{\rho' \to \varphi'}\qquad \dfrac{\forall y \rho'}{\rho'}}{\varphi'}}{r(x) \to \varphi'}$$

$$\dfrac{\exists y K(y,x) \qquad r(x) \to \varphi'}{r(x) \to \varphi'}$$

$$\dfrac{r(x) \to \varphi'}{\forall x(r(x) \to \varphi')}$$

[3]Notice that, by construction, the free variables in $K(\overline{y},\overline{x})$ are those in $\overline{y} \cup \overline{x}$.

$$\dfrac{\dfrac{\dfrac{T' \quad K(\overline{y},\overline{x})}{K(y,x)\to(\varphi'\to\rho')} \quad \dfrac{K(y,x)\wedge r(x)}{K(y,x)\quad r(x)} \quad \dfrac{\forall y(r(x)\to\varphi')}{r(x)\to\varphi'}}{\dfrac{T'}{\forall y\exists x(K(y,x)\wedge r(x))}\quad \dfrac{\varphi'\to\rho'}{\rho'}\quad \varphi'}}{\dfrac{\exists x(K(y,x)\wedge r(x))}{\rho'}}$$

$$\dfrac{\rho'}{\rho'}$$

$$\dfrac{\rho'}{\forall y\rho'}$$

QED

Corollary B.1 (Eliminability of formulae without new free variables) *For each formula $\rho(\overline{v})$ of L' without free variables of sort N, there exists a formula $\varphi(\overline{v})$ of L with the same free variables, such that $T' \vdash \rho(\overline{v}) \leftrightarrow \varphi(\overline{v})$.*

Proof: This follows from eliminability under connection. Since $|\overline{y}| = 0$ we also have $|\overline{x}| = 0$. So, we have a trivial connection with empty conjunction as its content.
QED

Proposition B.2 (Invariance under elementary equivalence) *Given $\mathcal{A}'_1, \mathcal{A}'_2 \models T'$ if their reducts $\mathcal{A}_1, \mathcal{A}_2$ to L are elementarily equivalent then so are $\mathcal{A}'_1$ and $\mathcal{A}'_2$.*

Proof: This follows from the preceding corollary. Indeed, for each sentence τ of L' we have a sentence σ of L such that $T' \vdash \tau \leftrightarrow \sigma$. Thus, for each model $\mathcal{A}'$ of T' with reduct $\mathcal{A} = \mathcal{A}'/L$, we have $\mathcal{A}' \models \tau$ iff $\mathcal{A}' \models \sigma$ iff $\mathcal{A} \models \sigma$. Hence, if $\mathcal{A}_1 \equiv \mathcal{A}_2$ then $\mathcal{A}'_1 \equiv \mathcal{A}'_2$.
QED

C APPENDIX

Eliminability Results for Quotient Sort

In this appendix we prove the eliminability results for extensions by introduction of quotient sorts.

Lemma C.1 (Characterisation of terms of L') .

1. Each term of L is a term of L'.
2. Each term of L' of sort Q is
   - either a variable of sort Q
   - or $p(t)$ for some term t of L.

Proof: The extension adds no new function (or constant) symbol having as target a sort in L, and the only new function symbol is $p : S \to Q$. So, if p occurs in a term, it can occur only as its outermost symbol.

QED

Lemma C.2 (Characterisation of atomic formulae of L') *Consider an atomic formula α' of L'.*

1. *Either α' is an equality of one of the forms:*
   - *$y = y'$ between variables y and y' of sort Q.*
   - *$y = p(t)$ or $p(t) = y$, with variable y over sort Q and term t of L with sort S.*
   - *$p(t) = p(t')$, for some terms t, t' of L with sort S.*
2. *Or else, α' is a atomic formula α of L, other than equality.*

Proof: Since the extension adds no new predicate symbol, these assertions follow immediately from the previous lemma C.1.

QED

Recall the concept of (quotient) connection. Consider sets of variables $\overline{y} = \{y^1, \ldots, y^n\}$ of sort Q and $\overline{v}$ of sorts in L. A (quotient) connection for $\overline{y}$ over $\overline{v}$ consists of a set $\overline{x} = \{x^1, \ldots, x^n\}$ of variables of sort S such that $\overline{v} \cap \overline{x} = \emptyset$. We also form the conjunction $K(\overline{y}, \overline{x})$ of $j(y^k) = x^k$, for $k = 1, \ldots, n$. Call such conjunction $K(\overline{y}, \overline{x})$ the **content** of the connection.

Proposition C.1 (Eliminability under connection) *Consider a (quotient) connection for $\overline{y}$ over $\overline{v}$ with content $K(\overline{y}, \overline{x})$.*

For every formula $\rho(\overline{v}, \overline{y})$ of L' having as its free variables those in $\overline{v} \cup \overline{y}$, there exists a formula $\varphi(\overline{v}, \overline{x})$ of L with free variables in $\overline{v} \cup \overline{x}$, such that $T' \vdash K(\overline{y}, \overline{x}) \to (\rho(\overline{v}, \overline{y}) \leftrightarrow \varphi(\overline{v}, \overline{x}))$. Call such a formula φ of L a **reflection** *of ρ into L (via K), and denote it by ρ^K.*

Proof: by induction on the structure of ρ.

- Basis: ρ is atomic. By lemma C.2, we have three cases to consider.

  1. ρ is $p(t) = p(t')$. Then, take φ as $q(t, t')$
     Clearly, $T' \vdash p(t) = p(t') \leftrightarrow q(t, t')$.
  2. ρ is $y = p(t)$ or $p(t) = y$, with t a term of L, of sort S. Since they are logically equivalent, we consider ρ as $y = p(t)$. Then, take φ as $q(x, t)$.
     Clearly, $T' \vdash K(y, x) \rightarrow (y = p(t) \leftrightarrow q(x, t))$.
  3. ρ is $y = y'$ with variable $y, y' : Q$. Take φ as $q(x, x')$.
     Clearly, $T' \vdash q(x, x') \leftrightarrow p(x) = p(x')$. Hence, $T' \vdash K(y, y', x, x') \rightarrow (y = y' \leftrightarrow q(x, x'))$, since $K(y, y', x, x')$ is $y = p(x) \wedge y' = p(x')$.

- $\rho(\overline{v}, \overline{y})$ is a negation $\neg\rho'(\overline{v}, \overline{y})$. The inductive hypothesis gives a formula $\varphi'(\overline{v}, \overline{x})$ of L, such that $T' \vdash K(\overline{y}, \overline{x}) \rightarrow (\rho' \leftrightarrow \varphi')$. So, we take $\varphi(\overline{v}, \overline{x})$ as $\neg\varphi'(\overline{v}, \overline{x})$.

- $\rho(\overline{v}, \overline{y})$ is a conjunction $\rho'(\overline{v'}, \overline{y'}) \wedge \rho''(\overline{v''}, \overline{y''})$. The inductive hypothesis gives formulae $\varphi'(\overline{v'}, \overline{x'})$ and $\varphi''(\overline{v''}, \overline{x''})$ of L such that $T' \vdash K(\overline{y'}, \overline{x'}) \rightarrow (\rho' \leftrightarrow \varphi')$ and $T' \vdash K(\overline{y''}, \overline{x''}) \rightarrow (\rho'' \leftrightarrow \varphi'')$.
  Then, we have that $T' \vdash K(\overline{y'}, \overline{x'}) \wedge K(\overline{y''}, \overline{x''}) \rightarrow (\rho' \wedge \rho'' \leftrightarrow \varphi' \wedge \varphi'')$.

- ρ is a universal formula. We examine two subcases, depending on the sort of the variable quantified upon[4].

  - Quantification on a variable of sort in L, that is $\rho(\overline{v}, \overline{y})$ is $(\forall v : S)\rho'(v, \overline{v}, \overline{y})$. The inductive hypothesis gives a formula $\varphi'(v, \overline{v}, \overline{x})$ of L, such that $T' \vdash K(\overline{y}, \overline{x}) \rightarrow (\rho'(v, \overline{v}, \overline{y}) \leftrightarrow \varphi'(v, \overline{v}, \overline{x}))$. Since v does not occur (free) in $K(\overline{y}, \overline{x})$ we have $T' \vdash K(\overline{y}, \overline{x}) \rightarrow (\forall v : S)(\rho'(v, \overline{v}, \overline{y}) \leftrightarrow \varphi'(v, \overline{v}, \overline{x}))$. Hence $T' \vdash K(\overline{y}, \overline{x}) \rightarrow ((\forall v : S)\rho'(v, \overline{v}, \overline{y}) \leftrightarrow (\forall v : S)\varphi'(v, \overline{v}, \overline{x}))$ and we take $\varphi(\overline{v}, \overline{x})$ as $(\forall v : S)\varphi'(v, \overline{v}, \overline{x})$.
  - Quantification on a variable $y : Q$, that is $\rho(\overline{v}, \overline{y})$ is $(\forall y : Q)\rho'(\overline{v}, \overline{y}, y)$. The inductive hypothesis gives a formula $\varphi'(\overline{v}, \overline{x}, x)$ of L, such that $T' \vdash K(\overline{y}, y, \overline{x}, x) \rightarrow (\rho'(\overline{v}, \overline{y}, y) \leftrightarrow \varphi'(\overline{v}, \overline{x}, x))$. Clearly, $K(\overline{y}, \overline{x}) \wedge K(y, x) \rightarrow K(\overline{y}, y, \overline{x}, x)$. Thus, $T' \vdash K(\overline{y}, \overline{x}) \rightarrow (K(y, x) \rightarrow (\rho' \leftrightarrow \varphi'))$. So, since y, x do not occur (free) in $K(\overline{y}, \overline{x})$, we have $T' \vdash K(\overline{y}, \overline{x}) \rightarrow (\forall y : Q)(\forall x : S)(K(y, x) \rightarrow (\rho' \leftrightarrow \varphi'))$. But, we also have, $T' \vdash (\forall y : Q)(\exists x : S)K(y, x)$ and $T' \vdash (\forall x : S)(\exists y : Q)K(y, x)$. Hence
    $T' \vdash K(\overline{y}, \overline{x}) \rightarrow ((\forall y : Q)\rho' \leftrightarrow (\forall x : S)\varphi')$.
    Indeed, this last step becomes clear in a Natural Deduction style.

[4]Notice that, by construction, the free variables in $K(\overline{y}, \overline{x})$ are those in $\overline{y} \cup \overline{x}$.

$$\dfrac{\dfrac{T' \quad K(\bar{y},\bar{x})}{K(y,x)\to(\varphi'\to\rho')}\;K(y,x) \qquad \dfrac{\forall x\,\varphi'}{\varphi'}}{\dfrac{\dfrac{T'}{\forall y\,\exists x\,K(y,x)}}{\dfrac{\exists x\,K(y,x)}{\dfrac{\rho'}{\forall y\,\rho'}}}\qquad \varphi'\to\rho' \qquad \rho'}$$

$$\dfrac{\dfrac{T' \quad K(\bar{y},\bar{x})}{K(y,x)\to(\rho'\to\varphi')}\;K(y,x) \qquad \dfrac{\forall y\,\rho'}{\rho'}}{\dfrac{\dfrac{T'}{\forall x\,\exists y\,K(y,x)}}{\dfrac{\exists y\,K(y,x)}{\dfrac{\varphi'}{\forall x\,\varphi'}}}\qquad \rho'\to\varphi' \qquad \varphi'}$$

QED

Corollary C.1 (Eliminability of formulae without new free variables) *For each formula $\rho(\overline{v})$ of L' without free variables of sort Q, there exists a formula $\varphi(\overline{v})$ of L with the same free variables, such that $T' \vdash \rho(\overline{v}) \leftrightarrow \varphi(\overline{v})$.*

Proof: This follows from eliminability under connection. Since $|\overline{y}| = 0$ we also have $|\overline{x}| = 0$. So, we have a trivial connection with empty conjunction as its content.
QED

Proposition C.2 (Invariance under elementary equivalence) *Given $\mathcal{A}'_1, \mathcal{A}'_2 \models T'$ if their reducts $\mathcal{A}_1, \mathcal{A}_2$ to L are elementarily equivalent then so are $\mathcal{A}'_1$ and $\mathcal{A}'_2$.*

Proof: This follows from the preceding corollary. Indeed, for each sentence τ of L' we have a sentence σ of L such that $T' \vdash \tau \leftrightarrow \sigma$. Thus, for each model $\mathcal{A}'$ of T' with reduct $\mathcal{A} = \mathcal{A}'/L$, we have $\mathcal{A}' \models \tau$ iff $\mathcal{A}' \models \sigma$ iff $\mathcal{A} \models \sigma$. Hence, if $\mathcal{A}_1 \equiv \mathcal{A}_2$ then $\mathcal{A}'_1 \equiv \mathcal{A}'_2$.
QED

D APPENDIX

Eliminability Results for Union Sort

In this appendix we prove the eliminability results for extensions by introduction of union sorts.

Lemma D.1 (Characterisation of terms of L') .

1. Each term of L is a term of L'.
2. Each term of L' of sort U is
   - either a variable of sort U
   - or $i_l(t)$ for some term t of L where $l = 1, 2$.

Proof: The extension adds no new function (or constant) symbol having as target a sort in L, and the only new function symbols are $-1 : S_1 \to U$ and $i_2 : S_2 \to U$. So, if i_l occurs in a term, it can occur only as its outermost symbol.

QED

Lemma D.2 (Characterisation of atomic formulae of L') *Consider an atomic formula α' of L'.*

1. *Either α' is an equality of one of the forms:*
   - *$y = y'$ between variables y and y' of sort U.*
   - *$y = i_l(t)$ or $i_l(t) = y$, with variable y over sort U and term t of L with sort S_l where $l = 1, 2$.*
   - *$i_l(t) = i_l(t')$, for some terms t, t' of L with sort S_l where $l = 1, 2$.*
2. *Or else, α' is a atomic formula α of L, other than equality.*

Proof: Since the extension adds no new predicate symbol, these assertions follow immediately from the previous lemma D.1.

QED

Recall the concept of (union) connection. Consider sets of variables $\overline{y} = \{y^1, \ldots, y^n\}$ of sort U and $\overline{v}$ of sorts in L. A (union) connection for $\overline{y}$ over $\overline{v}$ consists of sets $\overline{x_1} = \{x_1^1, \ldots, x_1^n\}$ of variables of sort S_1 and $\overline{x_2} = \{x_2^1, \ldots, x_2^n\}$ of variables of sort S_2 such that $\overline{v} \cap \overline{x_1} = \emptyset = \overline{v} \cap \overline{x_1}$. For each function $\delta : \{1, \ldots, n\} \to \{1, 2\}$ we form the conjuntion $K^\delta(\overline{y}, \overline{x_1}, \overline{x_2}) = \wedge_{j=1}^n (y_j = i^j_{\delta(j)}(x^j_{\delta(j)}))$ and the set $\overline{y}^\delta = \{x^1_{\delta(1)}, \ldots, x^n_{\delta(n)}\}$. Call such conjunction $K^\delta(\overline{y}, \overline{x_1}, \overline{x_2})$ and the set $\overline{y}^\delta$ the δ-**content** of the connection.

Proposition D.1 (Eliminability under connection) *Given a (union) connection for $\overline{y}$ over $\overline{v}$ with δ-content $K^\delta(\overline{y}, \overline{x_1}, \overline{x_2})$ and $\overline{y}^\delta$.*

*Given a formula $\rho(\overline{v}, \overline{y})$ of L' having as its free variables those of $\overline{v} \cup \overline{y}$, there exist 2^n formulae $\varphi(\overline{v}, \overline{y}^\delta)$ of L with free variables in $\overline{v} \cup \overline{y}^\delta$, such that $T' \vdash K^\delta \to (\rho(\overline{v}, \overline{y}) \leftrightarrow \varphi(\overline{v}, \overline{y}^\delta))$. Call such a formula φ of L the δ-**reflection of** ρ into L (via K), and denote it by $\rho^{K\delta}$.*

Proof: by induction on the structure of ρ.

- Basis: ρ is atomic. By lemma D.2, we have four cases to consider.
  - ρ is $y^1 = y^2$. Then, we have four functions δ to consider, and so we obtain four K^δ and formulae ρ^{K^δ}.
    - $\delta_1(l) = 1$ for $l = 1, 2$. Then, $K^{\delta_1} : y^1 = i_1(x_1^1) \wedge y^2 = i_1(x_1^2)$ and take $\rho^{K^{\delta_1}} : x_1^1 = x_1^2$.
    - $\delta_2(l) = 1$ for $l = 1, 2$. Then, $K^{\delta_2} : y^1 = i_2(x_2^1) \wedge y^2 = i_2(x_2^2)$ and take $\rho^{K^{\delta_2}} : x_2^1 = x_2^2$.
    - $\delta_3 = id$. Then, $K^{\delta_3} : y^1 = i_1(x_1^1) \wedge y^2 = i_2(x_2^2)$ and take $\rho^{K^{\delta_3}} : F$.
    - $\delta_4(1) = 2$ and $\delta_4(2) = 1$. Then,$K^{\delta_4} : y^1 = i_2(x_2^1) \wedge y^2 = i_1(x_1^2)$ and take $\rho^{K^{\delta_4}} : F$.

    It is easy to show that in each case we have:
    $T' \vdash K^{\delta_i} \rightarrow (\rho \leftrightarrow \rho^{K^{\delta_i}}$, where $i = 1, 2, 3, 4$.
  - ρ is $y^1 = i_1(t)$ or $i_1(t) = y^i$, with t a term of L, of sort S_1. Then, we have two cases to consider:
    - $\delta_1(1) = 1$, then $K^{\delta_1} : y^1 = i_1(x_1^1)$ and take $\rho^{K^{\delta_1}} : x_1^1 = t$.
    - $\delta_2(1) = 12$, then $K^{\delta_2} : y^1 = i_2(x_2^1)$ and take $\rho^{K^{\delta_2}} : F$.

    It is easy to show that in both cases we have:
    $T' \vdash K^{\delta_i} \rightarrow (\rho \leftrightarrow \rho^{K^{\delta_i}}$, where $i = 1, 2$.
  - ρ is $y^1 = i_2(t)$ or $i_2(t) = y^i$, with t a term of L, of sort S_2. Idem above.
  - ρ is $i_l(t) = i_l(t')$, with t, t' terms of L of sort S_l (where $l = 1, 2$). Then $\delta = \emptyset$, $K^\delta : V$ and $\rho^{K^\delta}; t = t'$.
- $\rho(\overline{v}, \overline{y})$ is a negation $\neg\rho'(\overline{v}, \overline{y})$ with n variables in $\overline{y}$. For each $\delta : \{1, \ldots, n\} \rightarrow \{1, 2\}$ the inductive hypothesis gives a formula $\rho'^{K^\delta}(\overline{v}, \overline{y}^\delta)$ of L, such that $T' \vdash K^\delta(\overline{y}, \overline{y}^\delta) \rightarrow (\rho' \leftrightarrow \rho'^{K^\delta})$. So, we take $\rho^{K^\delta}(\overline{v}, \overline{y}^\delta)$ as $\neg\rho'^{K^\delta}(\overline{v}, \overline{y}^\delta)$.
- $\rho(\overline{v}, \overline{y})$ is a conjunction $\rho'(\overline{v'}, \overline{y'}) \wedge \rho'(\overline{v''}, \overline{y''})$ with n variables in $\overline{y}$. For each $\delta : \{1, \ldots, n\} \rightarrow \{1, 2\}$ the inductive hypothesis gives formulae $\rho'^{K^\delta}(\overline{v'}, \overline{y'}^\delta)$ and $\rho''^{K^\delta}(\overline{v''}, \overline{y''}^\delta)$ of L such that $T' \vdash K^\delta(\overline{y'}, \overline{y'}^\delta) \rightarrow (\rho' \leftrightarrow \rho'^{K^\delta})$ and $T' \vdash K^\delta(\overline{y''}, \overline{y''}^\delta) \rightarrow (\rho'' \leftrightarrow \rho''^{K^\delta})$.

  Then, we have that $T' \vdash K^\delta(\overline{y'}, \overline{y'}^\delta) \wedge K(\overline{y''}, \overline{y''}^\delta) \rightarrow (\rho' \wedge \rho'' \leftrightarrow \rho'^{K^\delta} \wedge \rho''^{K^\delta})$.
- ρ is a universal formula. We examine two subcases, depending on the sort of the variable quantified upon[5].
  - Quantification on a variable of sort in L, that is $\rho(\overline{v}, \overline{y})$ is $(\forall v : S_l)\rho'(v, \overline{v}, \overline{y})$ with $l = 1, 2$ and n variables in $\overline{y}$. For each function $\delta : \{1, \ldots, n\} \rightarrow \{1, 2\}$ the inductive hypothesis gives a formula $\rho'^{K^\delta}(v, \overline{v}, \overline{y}^\delta)$ of L, such that $T' \vdash K^\delta(\overline{y}, \overline{y}^\delta) \rightarrow (\rho'(v, \overline{v}, \overline{y}) \leftrightarrow \rho'^{K^\delta}(v, \overline{v}, \overline{y}^\delta))$. Since v does not

[5]Notice that, by construction, the free variables in $K(\overline{y}, \overline{y}^\delta)$ are those of $\overline{y} \cup \overline{y}^\delta$.

occur (free) in $K^\delta(\overline{y},\overline{y}^\delta)$ we have $T' \vdash K^\delta(\overline{y},\overline{y}^\delta) \to (\forall v : S_l)(\rho'(v,\overline{v},\overline{y}) \leftrightarrow \rho'^{K^\delta}(v,\overline{v},\overline{y}^\delta))$. Hence $T' \vdash K^\delta(\overline{y},\overline{y}^\delta) \to ((\forall v : S_l)\rho'(v,\overline{v},\overline{y}) \leftrightarrow (\forall v : S_l)\rho'^{K^\delta}(v,\overline{v},\overline{y}^\delta))$ and we take $\rho^{K^\delta}(\overline{v},\overline{x})$ as $(\forall v : S_l)\rho'^{K^\delta}(v,\overline{v},\overline{y}^\delta)$.

– Quantification on a variable $y : U$, that is $\rho(\overline{v},\overline{y})$ is $(\forall y^{n+1} : U)\rho'(\overline{v},\overline{y},y)$ with n variables in $\overline{y}$. For each function $\delta' : \{1,\ldots,n,n+1\} \to \{1,2\}$ the inductive hypothesis gives a formula $\rho'^{K^{\delta'}}(\overline{v},\overline{y}^{\delta'},x^{n+1}_{\delta'(n+1)})$ of L, such that $T' \vdash K^{\delta'}(\overline{y},y^{n+1},\overline{y}^{\delta'},x^{n+1}_{\delta'(n+1)}) \to (\rho'(\overline{v},\overline{y},y^{n+1}) \leftrightarrow \rho'^{K^{\delta'}}(\overline{v},\overline{y}^{\delta'},x^{n+1}_{\delta'(n+1)}))$. Take the restriction function $\delta = \delta'/_{\{1,\ldots,n\}}$ and $K^\delta(\overline{y},\overline{y}^\delta = K^{\delta'}(\overline{y},\overline{y}^{\delta'})$. Then, $\rho^{K^\delta}(\overline{v},\overline{y}^\delta)$ is $\forall x^{n+1}_{\delta'(n+1)}\rho'^{K^{\delta'}}(\overline{v},\overline{y}^{\delta'})$.

We prove the equivalence of ρ and ρ^{K^δ} in a Natural Deduction style.

$$\dfrac{\dfrac{\dfrac{T'}{\forall x^{n+1}_{\delta'(n+1)}\,\exists y^{n+1}\,K^{\delta'}(y^{n+1},x^{n+1}_{\delta'(n+1)})}}{\exists y^{n+1}\,K^{\delta'}(y^{n+1},x^{n+1}_{\delta'(n+1)})} \qquad \dfrac{\dfrac{\dfrac{T' \quad K^{\delta'}(\overline{y},\overline{y}^{\delta'})}{K^{\delta'}(y^{n+1},x^{n+1}_{\delta'(n+1)}) \to (\rho' \to \varphi')} \quad K^{\delta'}(y^{n+1},x^{n+1}_{\delta'(n+1)})}{\rho' \to \varphi'} \qquad \dfrac{\forall y^{n+1}\rho'}{\rho'}}{\varphi'}}{\dfrac{\varphi'}{\forall x^{n+1}_{\delta'(n+1)}\,\varphi'}}$$

$$\dfrac{\dfrac{\dfrac{T'}{\forall y^{n+1}\,\exists x^{n+1}_{\delta'(n+1)}\,K^{\delta'}(y^{n+1},x^{n+1}_{\delta'(n+1)})}}{\exists x^{n+1}_{\delta'(n+1)}\,K^{\delta'}(y^{n+1},x^{n+1}_{\delta'(n+1)})} \qquad \dfrac{\dfrac{\dfrac{T' \quad K^{\delta'}(\overline{y},\overline{y}^{\delta'})}{K^{\delta'}(y^{n+1},x^{n+1}_{\delta'(n+1)}) \to (\varphi' \to \rho')} \quad K^{\delta'}(y^{n+1},x^{n+1}_{\delta'(n+1)})}{\varphi' \to \rho'} \qquad \dfrac{\forall x^{n+1}_{\delta'(n+1)}\,\varphi'}{\varphi'}}{\rho'}}{\dfrac{\rho'}{\forall y^{n+1}\rho'}}$$

QED

Corollary D.1 (Eliminability of formulae without new free variables) *For each formula $\rho(\overline{v})$ of L' without free variables of sort U, there exists a formula $\varphi(\overline{v})$ of L with the same free variables, such that $T' \vdash \rho(\overline{v}) \leftrightarrow \varphi(\overline{v})$.*

Proof: This follows from eliminability under connection. Since $|\overline{y}| = 0$ we also have $|\overline{x_1}| = |\overline{x_2}| = 0$. So, we have a trivial connection with empty conjunction as its content.

QED

Proposition D.2 (Invariance under elementary equivalence) *Given* $\mathcal{A}'_1, \mathcal{A}'_2 \models T'$ *if their reducts* $\mathcal{A}_1, \mathcal{A}_2$ *to* L *are elementarily equivalent then so are* $\mathcal{A}'_1$ *and* $\mathcal{A}'_2$.

Proof: This follows from the preceding corollary. Indeed, for each sentence τ of L' we have a sentence σ of L such that $T' \vdash \tau \leftrightarrow \sigma$. Thus, for each model $\mathcal{A}'$ of T' with reduct $\mathcal{A} = \mathcal{A}'/L$, we have $\mathcal{A}' \models \tau$ iff $\mathcal{A}' \models \sigma$ iff $\mathcal{A} \models \sigma$. Hence, if $\mathcal{A}_1 \equiv \mathcal{A}_2$ then $\mathcal{A}'_1 \equiv \mathcal{A}'_2$.

QED