Toward Isomorphism of Intersection and Union Types

(Article begins on next page)

14 October 2024

Toward Isomorphism of Intersection and Union Types[*]

Dedicated to Corrado Böhm on the occasion of his 90th Birthday

Mario Coppo   Mariangiola Dezani-Ciancaglini   Ines Margaria   Maddalena Zacchi

Dipartimento di Informatica Università di Torino, corso Svizzera 185, 10149 Torino, Italy

This paper investigates type isomorphism in a λ-calculus with intersection and union types. It is known that in λ-calculus, the isomorphism between two types is realised by a pair of terms inverse one each other. Notably, invertible terms are linear terms of a particular shape, called finite hereditary permutators. Typing properties of finite hereditary permutators are then studied in a relevant type inference system with intersection and union types for linear terms. In particular, an isomorphism preserving reduction between types is defined. Reduction of types is confluent and terminating, and induces a notion of normal form of types. The properties of normal types are a crucial step toward the complete characterisation of type isomorphism. The main results of this paper are, on one hand, the fact that two types with the same normal form are isomorphic, on the other hand, the characterisation of the isomorphism between types in normal form, modulo isomorphism of arrow types.

1 Introduction

In a calculus with types, two types σ and τ are *isomorphic* if there exist two terms P of type $\sigma \to \tau$ and P' of type $\tau \to \sigma$ such that both their compositions $P \circ P' = \lambda x.P(P'x)$ and $P' \circ P = \lambda x.P'(Px)$ give the identity (at the proper type). The study of type isomorphism started in the 1980s with the aim of finding all the type isomorphisms valid in every model of a given language [3]. If one looks at this problem choosing as language a λ-calculus with types, one can immediately note the close relation between type isomorphism and λ-term invertibility. Actually, in the untyped λ-calculus a λ-term P is *invertible* if there exists a λ-term P' such that $P \circ P' =_{\beta\eta} P' \circ P =_{\beta\eta} \mathbf{I}$ ($\mathbf{I} = \lambda x.x$). The problem of term invertibility has been extensively studied for the untyped λ-calculus since 1970 and the main result has been the complete characterisation of the invertible λ-terms in $\lambda\beta\eta$-calculus [6]: the invertible terms are all and only the *finite hereditary permutators*.

Definition 1.1 (Finite Hereditary Permutator)**.** *A finite hereditary permutator (FHP for short) is a λ-term of the form (modulo β-conversion)*

$$\lambda x y_1 \ldots y_n.x(P_1 y_{\pi(1)}) \ldots (P_n y_{\pi(n)}) \quad (n \geq 0)$$

where π is a permutation of $1, \ldots, n$, and $P_1, \ldots, P_n$ are FHPs.

Note that the identity is trivially an FHP (take $n = 0$). Another example of an FHP is
$$\lambda x y_1 y_2.x y_2 y_1 = \lambda x y_1 y_2.x((\lambda z.z) y_2)((\lambda z.z) y_1).$$
It is easy to show that FHPs are closed by composition.

Theorem 1.2. *A λ-term is invertible iff it is a finite hereditary permutator.*

This result, obtained in the framework of the untyped λ-calculus, has been the basis for studying type isomorphism in different type systems for the λ-calculus. Note that every FHP has, modulo $\beta\eta$-conversion,

---

a unique inverse P^{-1}. Even if in the type free λ-calculus FHPs are defined in [6] modulo $\beta\eta$-conversion, in this paper each FHP is considered only modulo β-conversion, because types are not invariant under η-expansion. Taking into account these properties, the definition of type isomorphism in a λ-calculus with types can be stated as follows:

Definition 1.3 (Type isomorphism). *Given a λ-calculus with types, two types σ and τ are isomorphic ($\sigma \approx \tau$) if there exists a pair $< P, P^{-1} >$ of FHPs, inverse of each other, such that $\vdash P : \sigma \to \tau$ and $\vdash P^{-1} : \tau \to \sigma$. The pair $< P, P^{-1} >$ proves the isomorphism.*

When $P = P^{-1}$ one can simply write "P proves the isomorphism".

The main approach used to characterise type isomorphism in a given system has been to provide a suitable set of equations and to prove that these equations induce the type isomorphism w.r.t. $\beta\eta$-conversion, i.e. that the types of the FHPs are all and only those induced by the set of equations.

The typed λ-calculus studied first has been the simply typed λ-calculus. For this calculus Bruce and Longo proved in [3] that only one equation is needed, namely the **swap** equation:

$$\sigma \to \tau \to \rho \;\approx\; \tau \to \sigma \to \rho$$

Later, the study has been directed toward richer λ-calculi, obtained from the simply typed λ-calculus in an incremental way, by adding some other type constructors (like product and unit types [15, 2, 14]) or by allowing higher-order types (System F [3, 8]). Di Cosmo summarised in [9] the equations characterising type isomorphism; the set of equations grows incrementally in the sense that the set of equations for a typed λ-calculus, obtained by adding a primitive to a given λ-calculus, is an extension of the set of equations of the λ-calculus without that primitive.

In the presence of intersection, this incremental approach does not work, as pointed out in [7]; in particular with intersection types, the isomorphism is no longer a congruence and type equality in the standard models of intersection types does not entail type isomorphism. These quite unexpected facts required the introduction of a syntactical notion of type similarity in order to fully characterise the isomorphic types [7].

The study of isomorphism looks even harder for type systems with intersection and union types because for these systems, in general, the Subject Reduction property does not hold [1]. As in the case of intersection types, the isomorphism of union types is not a congruence and it is not complete for type equality in standard models. For example $\sigma \vee \tau \to \rho$ and $\tau \vee \sigma \to \rho$ are isomorphic, while $(\sigma \vee \tau \to \rho) \vee \varphi$ and $(\tau \vee \sigma \to \rho) \vee \varphi$ are not isomorphic, whenever φ is an atomic type.

This paper gives essential results for the characterisation of isomorphism of intersection and union types. To this aim a relevant type system, defined as a slight modification of the standard one in [11], has been introduced. In this system, in particular, Subject Conversion holds for linear terms.

A main difficulty in studying types for FHPs is that intersection/union introduction and elimination rules allow to write types in different, although isomorphic, ways. Since the standard distributive laws of union and intersection correspond to provable isomorphisms, types can be considered both in disjunctive and in conjunctive normal forms. This, besides providing a very useful technical tool, allows one to define, together with other basic isomorphisms involving the $\to$ type constructor, a general notion of normal form of types. A main result proved in this paper is that if $\sigma \to \tau$ is a type of an FHP P, then:

- for all μ in the disjunctive normal form of σ, there is ν in the disjunctive normal form of τ such that $\mu \to \nu$ is a type of P;

- for all κ in the conjunctive normal form of τ, there is χ in the conjunctive normal form of σ such that $\chi \to \kappa$ is a type of P.

Another crucial contribution is the introduction of normalisation rules which allow to split arrows over intersections/unions and to erase "useless" types by preserving isomorphism. The proof of soundness of these rules is done by building the pairs of FHPs witnessing isomorphism. Termination and confluence of the normalisation rules are also shown. Two types with the same normal form are isomorphic. Normal types are intersections of unions of atomic and arrow types. A key property is that two isomorphic normal types have the same number of intersections and unions and that the arrows and atoms are pairwise isomorphic. The last step toward a complete characterisation of type isomorphism is that of combining normal forms with the **swap** equation, and this is done in [4].

2   Type assignment system

$$(Ax) \qquad x\!:\!\sigma \vdash x\!:\!\sigma$$

$$(\rightarrow I) \quad \frac{\Gamma, x\!:\!\sigma \vdash M\!:\!\tau}{\Gamma \vdash \lambda x.M\!:\!\sigma \rightarrow \tau} \qquad (\rightarrow E) \quad \frac{\Gamma_1 \vdash M\!:\!\sigma \rightarrow \tau \quad \Gamma_2 \vdash N\!:\!\sigma}{\Gamma_1, \Gamma_2 \vdash MN\!:\!\tau}$$

$$(\wedge I) \quad \frac{\Gamma \vdash M\!:\!\sigma \quad \Gamma \vdash M\!:\!\tau}{\Gamma \vdash M\!:\!\sigma \wedge \tau} \qquad (\wedge E) \quad \frac{\Gamma \vdash M\!:\!\sigma \wedge \tau}{\Gamma \vdash M\!:\!\sigma} \qquad \frac{\Gamma \vdash M\!:\!\sigma \wedge \tau}{\Gamma \vdash M\!:\!\tau}$$

$$(\vee I) \quad \frac{\Gamma \vdash M\!:\!\sigma}{\Gamma \vdash M\!:\!\sigma \vee \tau} \qquad \frac{\Gamma \vdash M\!:\!\sigma}{\Gamma \vdash M\!:\!\tau \vee \sigma}$$

$$(\vee E) \quad \frac{\Gamma_1, x\!:\!\sigma \wedge \theta \vdash M\!:\!\rho \quad \Gamma_1, x\!:\!\tau \wedge \theta \vdash M\!:\!\rho \quad \Gamma_2 \vdash N\!:\!(\sigma \vee \tau) \wedge \theta}{\Gamma_1, \Gamma_2 \vdash M[N/x]\!:\!\rho}$$

Figure 1: Typing rules.

The syntax of intersection and union types is given by:
$$\sigma \quad ::= \quad \varphi \quad | \quad \sigma \rightarrow \sigma \quad | \quad \sigma \wedge \sigma \quad | \quad \sigma \vee \sigma$$
where φ denotes an atomic type. It is useful to distinguish between different kinds of types. So in the following:

- $\sigma, \tau, \rho, \theta, \vartheta, \zeta, \varsigma$ range over arbitrary types;
- $\alpha, \beta, \gamma, \delta, \eth$ range over atomic and arrow types, defined as $\alpha \ ::= \ \varphi \ | \ \sigma \rightarrow \sigma$;
- $\mu, \nu, \lambda, \xi, \eta$ range over intersections of atomic and arrow types (*basic intersections*), defined as $\mu \ ::= \ \alpha \ | \ \mu \wedge \mu$;
- $\chi, \kappa, \iota, \omega$ range over unions of atomic and arrow types (*basic unions*), defined as $\chi \ ::= \ \alpha \ | \ \chi \vee \chi$.

Note that no structural equivalence is assumed between types, for instance $\sigma \vee \tau$ is different from $\tau \vee \sigma$. As usual, parentheses are omitted according to the precedence rule "$\vee$ and $\wedge$ over $\rightarrow$" and $\rightarrow$ associates to the right.

The union/intersection type system considered in this paper is a modified version of the basic one introduced in the seminal paper [11], restricted to linear λ-terms. A λ-term is *linear* if each free or bound variable occurs exactly once in it.

Figure 1 gives the typing rules. As usual, *environments* associate variables to types and contain at most one type for each variable. The environments are relevant, i.e. they contain only the used premises. When writing Γ_1, Γ_2 one convenes that the sets of variables in Γ_1 and Γ_2 are disjoint.
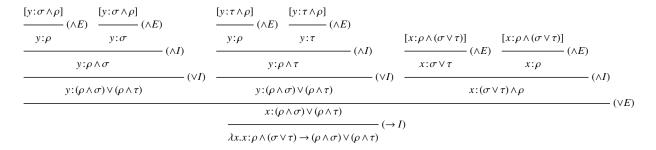
$$\dfrac{\dfrac{\dfrac{[y:\sigma\wedge\rho]}{y:\rho}(\wedge E)\quad\dfrac{[y:\sigma\wedge\rho]}{y:\sigma}(\wedge E)}{y:\rho\wedge\sigma}(\wedge I)}{y:(\rho\wedge\sigma)\vee(\rho\wedge\tau)}(\vee I)\qquad\dfrac{\dfrac{\dfrac{[y:\tau\wedge\rho]}{y:\rho}(\wedge E)\quad\dfrac{[y:\tau\wedge\rho]}{y:\tau}(\wedge E)}{y:\rho\wedge\tau}(\wedge I)}{y:(\rho\wedge\sigma)\vee(\rho\wedge\tau)}(\vee I)\qquad\dfrac{\dfrac{[x:\rho\wedge(\sigma\vee\tau)]}{x:\sigma\vee\tau}(\wedge E)\quad\dfrac{[x:\rho\wedge(\sigma\vee\tau)]}{x:\rho}(\wedge E)}{x:(\sigma\vee\tau)\wedge\rho}(\wedge I)}{\dfrac{\dfrac{x:(\rho\wedge\sigma)\vee(\rho\wedge\tau)}{\lambda x.x:\rho\wedge(\sigma\vee\tau)\to(\rho\wedge\sigma)\vee(\rho\wedge\tau)}(\to I)}{}}(\vee E)$$

Figure 2: A deduction of $\vdash\lambda x.x:\rho\wedge(\sigma\vee\tau)\to(\rho\wedge\sigma)\vee(\rho\wedge\tau)$.

The only non-standard rule is $(\vee E)$. This rule takes into account the fact that, as it seems natural in a system with intersection types, one variable can be used in a deduction with different types in different occurrences (by applications of the $(\wedge E)$ rule). It should then be possible, in general, to apply the union elimination only to the type of one of these occurrences. A paradigmatic example is the one in Figure 2 where one occurrence of the variable y is used (after an application of $(\wedge E)$) with type σ in one branch of the $(\vee E)$ rule and with type τ in the other branch. Other occurrences of y are used instead with type ρ in both branches. Rule $(\vee E)$ is then the right way to formulate union elimination in a type system in which union and intersection interact. It is indeed a generalisation of the $(\vee E')$ rule given in [11]. A last observation is that, being M linear, in an application of the $(\vee E)$ rule, exactly one occurrence of x is replaced inside M.

Some useful admissible rules are:

$$(L)\quad\dfrac{x:\sigma\vdash x:\tau\quad\Gamma,x:\tau\vdash M:\rho}{\Gamma,x:\sigma\vdash M:\rho}\qquad\qquad(C)\quad\dfrac{\Gamma_1,x:\sigma\vdash M:\tau\quad\Gamma_2\vdash N:\sigma}{\Gamma_1,\Gamma_2\vdash M[N/x]:\tau}$$

$$(\vee I')\quad\dfrac{\Gamma,x:\sigma\vdash M:\rho\quad\Gamma,x:\tau\vdash M:\rho}{\Gamma,x:\sigma\vee\tau\vdash M:\rho}\qquad(\vee E')\quad\dfrac{\Gamma_1,x:\sigma\vdash M:\rho\quad\Gamma_1,x:\tau\vdash M:\rho\quad\Gamma_2\vdash N:\sigma\vee\tau}{\Gamma_1,\Gamma_2\vdash M[N/x]:\rho}$$

To show $(\vee I')$ it is enough to apply rule $(\vee E)$ with $x:\sigma\vee\tau\vdash x:(\sigma\vee\tau)\wedge(\sigma\vee\tau)$ as third premise.

The system of Figure 1 can be extended to non-linear terms simply by erasing the condition that, in rules $(\to E)$ and $(\vee E)$, the environments need to be disjoint. It is easy to check that this extended system is conservative over the present one. Therefore the types that can be derived for FHPs are the same in the two systems, so the present study of type isomorphism holds for the extended system too.

In order to show Subject Reduction one can follow the classical approach of [12] by considering a sequent formulation of the type assignment system and showing cut elimination. This is done in [1] for a system which differs from the present one for being not relevant, having the universal type and rule $(\vee E')$ instead of $(\vee E)$. It is just routine to modify that proof by taking as left and right rules for the $\vee$ constructor:

$$(\vee L)\quad\dfrac{\Gamma,x:\sigma\wedge\theta\vdash M:\rho\quad\Gamma,x:\tau\wedge\theta\vdash M:\rho}{\Gamma,x:(\sigma\vee\tau)\wedge\theta\vdash M:\rho}\qquad(\vee R)\quad\dfrac{\Gamma\vdash M:\sigma\wedge\theta}{\Gamma\vdash M:(\sigma\vee\tau)\wedge\theta}\quad\dfrac{\Gamma\vdash M:\tau\wedge\theta}{\Gamma\vdash M:(\sigma\vee\tau)\wedge\theta}$$

Remark that, considering only linear terms, cut elimination corresponds to standard β-reduction, while for arbitrary terms parallel reductions are needed; for details see [1]. Therefore one can conclude:

Theorem 2.1 (SR). *If $\Gamma\vdash M:\sigma$ and $M\longrightarrow_\beta^* N$, then $\Gamma\vdash N:\sigma$.*

The Subject Reduction Theorem allows one to show the following corollary.

Corollary 2.2.   *1. If $\Gamma \vdash \lambda x.M : \sigma \to \tau$, then $\Gamma, x : \sigma \vdash M : \tau$.*

  2. If $\Gamma, x : \sigma \vee \rho \vdash M : \tau$, then $\Gamma, x : \sigma \vdash M : \tau$ and $\Gamma, x : \rho \vdash M : \tau$.

  3. If $\Gamma \vdash \lambda x.M : \sigma \to \rho$ and $\Gamma \vdash \lambda x.M : \tau \to \theta$, then $\Gamma \vdash \lambda x.M : \sigma \wedge \tau \to \rho \wedge \theta$ and $\Gamma \vdash \lambda x.M : \sigma \vee \tau \to \rho \vee \theta$.

  4. If $\Gamma \vdash \lambda x.M : \sigma \to \tau$, then $\Gamma \vdash \lambda x.M : \sigma \wedge \rho \to \tau \vee \theta$ for all ρ, θ.

Proof. (1). One gets $\Gamma, x : \sigma \vdash (\lambda y.M[y/x])x : \tau$ by rule $(\to E)$ and α-renaming. So the Subject Reduction Theorem (Theorem 2.1) concludes the proof.
(2). One gets $\Gamma \vdash \lambda y.M[y/x] : \sigma \vee \rho \to \tau$ by rule $(\to I)$ and α-renaming, and $\Gamma, x : \sigma \vdash (\lambda y.M[y/x])x : \tau$ by rules $(\vee I)$ and $(\to E)$. So Theorem 2.1 concludes the proof.
(3). By Point (1) $\Gamma, x : \sigma \vdash M : \rho$ and $\Gamma, x : \tau \vdash M : \theta$, so by rules $(\vee I)$ and $(\vee I')$ one gets $\Gamma, x : \sigma \vee \tau \vdash M : \rho \vee \theta$, which implies the result by rule $(\to I)$. The proof of $\Gamma \vdash \lambda x.M : \sigma \wedge \tau \to \rho \wedge \theta$ is similar.
(4). Obvious because, by Point (1), $\Gamma, x : \sigma \vdash M : \tau$. □

Also subject expansion holds.

Theorem 2.3 (SE). *If $M \longrightarrow_{\beta}^{*} N$ and M is a linear λ-term and $\Gamma \vdash N : \sigma$, then $\Gamma \vdash M : \sigma$.*

Proof. It is enough to show: $\Gamma \vdash M[N/x] : \sigma$ implies $\Gamma \vdash (\lambda x.M)N : \sigma$. The proof is by induction on the derivation of $\Gamma \vdash M[N/x] : \sigma$. The only interesting case is when the last applied rule is

$$(\vee E) \quad \frac{\Gamma_1, x : \rho \wedge \theta \vdash M : \sigma \quad \Gamma_1, x : \tau \wedge \theta \vdash M : \sigma \quad \Gamma_2 \vdash N : (\rho \vee \tau) \wedge \theta}{\Gamma_1, \Gamma_2 \vdash M[N/x] : \sigma}$$

It is easy to derive $x : (\rho \vee \tau) \wedge \theta \vdash x : (\rho \wedge \theta) \vee (\tau \wedge \theta)$. Rule $(\vee I')$ applied to the first two premises gives $\Gamma_1, x : (\rho \wedge \theta) \vee (\tau \wedge \theta) \vdash M : \sigma$. So rule (L) derives $\Gamma_1, x : (\rho \vee \tau) \wedge \theta \vdash M : \sigma$, and rule $(\to I)$ derives $\Gamma_1 \vdash \lambda x.M : (\rho \vee \tau) \wedge \theta \to \sigma$. Rule $(\to E)$ gives the conclusion. □

The following basic isomorphisms are directly related to standard properties of functional types and to set theoretic properties of union and intersection. It is interesting to remark that all these isomorphisms are provable equalities in the system $\mathbf{B}_+$ of relevant logic [13].

Lemma 2.4. *The following isomorphisms hold:*

idem.	$\sigma \wedge \sigma \approx \sigma,\ \sigma \vee \sigma \approx \sigma$
comm.	$\sigma \wedge \tau \approx \tau \wedge \sigma,\ \sigma \vee \tau \approx \tau \vee \sigma$
assoc.	$(\sigma \wedge \tau) \wedge \rho \approx \sigma \wedge (\tau \wedge \rho),\ (\sigma \vee \tau) \vee \rho \approx \sigma \vee (\tau \vee \rho)$
dist→ ∧.	$\sigma \to \tau \wedge \rho \approx (\sigma \to \tau) \wedge (\sigma \to \rho)$
dist→ ∨.	$\sigma \vee \tau \to \rho \approx (\sigma \to \rho) \wedge (\tau \to \rho)$
swap.	$\sigma \to \tau \to \rho \approx \tau \to \sigma \to \rho$
dist∧∨.	$(\sigma \vee \tau) \wedge \rho \approx (\sigma \wedge \rho) \vee (\tau \wedge \rho)$
dist∨∧.	$(\sigma \wedge \tau) \vee \rho \approx (\sigma \vee \rho) \wedge (\tau \vee \rho)$

Proof. The η-expansion of the identity $\lambda xy.xy$ proves the fourth and the fifth isomorphisms, $\lambda xy_1y_2.xy_2y_1$ proves the sixth one and the identity proves all the remaining ones. □

The isomorphisms **idem**, **comm** and **assoc** allow one to consider types, at top level, modulo idempotence, commutativity and associativity of $\wedge$ and $\vee$. Then types, at top level, can be written as $\bigwedge_{i \in I} \sigma_i$ and $\bigvee_{i \in I} \sigma_i$ with finite I, where a single arrow or atomic type is seen both as an intersection and as a union (in this case I is a singleton). However, as noted in the introduction, these isomorphisms are not preserved

by arbitrary contexts since, for example, $\sigma \vee \tau \to \rho \approx \tau \vee \sigma \to \rho$ but $(\sigma \vee \tau \to \rho) \wedge \varphi$ and $(\tau \vee \sigma \to \rho) \wedge \varphi$ are not isomorphic.

The isomorphisms of Lemma 2.4 naturally induce the notions of disjunctive and conjunctive forms. In particular:

- the *disjunctive weak normal form* of a type σ (notation $dw(\sigma)$) is obtained by using (**dist**$\wedge\vee$) from left to right at top level;

- the *conjunctive weak normal form* of a type σ (notation $cw(\sigma)$) is obtained by using (**dist**$\to \wedge$), (**dist**$\to \vee$), and (**dist**$\vee\wedge$) from left to right at top level.

Notice that the isomorphisms (**dist**$\to \wedge$) and (**dist**$\to \vee$) are useful only to get conjunctive normal forms, since they only generate intersections.

This section ends with some lemmas on derivability properties. Lemmas 2.5 characterises the types derivable for variables using disjunctive weak normal form, Corollary 2.6 considers three useful particular cases of previous lemma, while Lemma 2.7 gives typing properties of the application of a variable to n λ-terms.

Lemma 2.5. *If $dw(\sigma) = \bigvee_{i \in I}(\bigwedge_{h \in H_i} \alpha_h^{(i)})$, $dw(\tau) = \bigvee_{j \in J}(\bigwedge_{k \in K_j} \beta_k^{(j)})$ and $x : \sigma \vdash x : \tau$, then for all $i \in I$ there is $j_i \in J$ such that $\{\beta_k^{(j_i)} \mid k \in K_{j_i}\} \subseteq \{\alpha_h^{(i)} \mid h \in H_i\}$, which implies $x : \bigwedge_{h \in H_i} \alpha_h^{(i)} \vdash x : \bigwedge_{k \in K_{j_i}} \beta_k^{(j_i)}$.*

Proof. By induction on derivations. Assume $dw(\rho) = \bigvee_{l \in L}(\bigwedge_{w \in W_l} \gamma_w^{(l)})$ and $dw(\theta) = \bigvee_{r \in R}(\bigwedge_{s \in S_r} \delta_s^{(r)})$ and $dw(\vartheta) = \bigvee_{u \in U}(\bigwedge_{v \in V_u} \eth_v^{(u)})$. If the last applied rule is (Ax) or $(\vee I)$ it is easy.

If the deduction ends with $(\wedge I)$:

$$(\wedge I) \quad \frac{x : \sigma \vdash x : \rho \quad x : \sigma \vdash x : \theta}{x : \sigma \vdash x : \rho \wedge \theta}$$

by definition $dw(\rho \wedge \theta) = \bigvee_{l \in L} \bigvee_{r \in R}((\bigwedge_{w \in W_l} \gamma_w^{(l)}) \wedge (\bigwedge_{s \in S_r} \delta_s^{(r)}))$. By induction for all $i \in I$ there is $l_i \in L$ such that $\{\gamma_w^{(l_i)} \mid w \in W_{l_i}\} \subseteq \{\alpha_h^{(i)} \mid h \in H_i\}$ and for all $i \in I$ there is $r_i \in R$ such that $\{\delta_s^{(r_i)} \mid s \in S_{r_i}\} \subseteq \{\alpha_h^{(i)} \mid h \in H_i\}$, then for all $i \in I$ there are $l_i \in L$ and $r_i \in R$ such that $\{\gamma_w^{(l_i)} \mid w \in W_{l_i}\} \cup \{\delta_s^{(r_i)} \mid s \in S_{r_i}\} \subseteq \{\alpha_h^{(i)} \mid h \in H_i\}$.

If the deduction ends with $(\wedge E)$:

$$(\wedge E) \quad \frac{x : \sigma \vdash x : \tau \wedge \rho}{x : \sigma \vdash x : \tau}$$

by definition $dw(\tau \wedge \rho) = \bigvee_{j \in J} \bigvee_{l \in L}((\bigwedge_{k \in K_j} \beta_k^{(j)}) \wedge (\bigwedge_{w \in W_l} \gamma_w^{(l)}))$. By induction for all $i \in I$ there are $j_i \in J$ and $l_i \in L$ such that $\{\beta_k^{(j_i)} \mid k \in K_{j_i}\} \cup \{\gamma_w^{(l_i)} \mid w \in W_{l_i}\} \subseteq \{\alpha_h^{(i)} \mid h \in H_i\}$.

If the deduction ends with $(\vee E)$:

$$(\vee E) \quad \frac{y : \rho \wedge \theta \vdash y : \tau \quad y : \vartheta \wedge \theta \vdash y : \tau \quad x : \sigma \vdash x : (\rho \vee \vartheta) \wedge \theta}{x : \sigma \vdash x : \tau}$$

By definition $dw(\rho \wedge \theta) = \bigvee_{l \in L} \bigvee_{r \in R}((\bigwedge_{w \in W_l} \gamma_w^{(l)}) \wedge (\bigwedge_{s \in S_r} \delta_s^{(r)}))$ and $dw(\vartheta \wedge \theta) = \bigvee_{u \in U} \bigvee_{r \in R}((\bigwedge_{v \in V_u} \eth_v^{(u)}) \wedge (\bigwedge_{s \in S_r} \delta_s^{(r)}))$ and $dw((\rho \vee \vartheta) \wedge \theta) = dw((\rho \wedge \theta) \vee (\vartheta \wedge \theta)) = (\bigvee_{l \in L} \bigvee_{r \in R}((\bigwedge_{w \in W_l} \gamma_w^{(l)}) \wedge (\bigwedge_{s \in S_r} \delta_s^{(r)}))) \vee (\bigvee_{u \in U} \bigvee_{r \in R}((\bigwedge_{v \in V_u} \eth_v^{(u)}) \wedge (\bigwedge_{s \in S_r} \delta_s^{(r)})))$. By induction:

- on the first premise for all $l \in L$ and $r \in R$ there is $j_{l,r} \in J$ such that $\{\beta_k^{(j_{l,r})} \mid k \in K_{j_{l,r}}\} \subseteq \{\gamma_w^{(l)} \mid w \in W_l\} \cup \{\delta_s^{(r)} \mid s \in S_r\}$ and

- on the second premise for all $u \in U$ and $r \in R$ there is $j_{u,r} \in J$ such that
  $\{\beta_k^{(j_{u,r})} \mid k \in K_{j_{u,r}}\} \subseteq \{\eth_v^{(u)} \mid v \in V_u\} \cup \{\delta_s^{(r)} \mid s \in S_r\}$ and

- on the third premise for all $i \in I$ either there are $l_i \in L$ and $r_i \in R$ such that
  $\{\gamma_w^{(l_i)} \mid w \in W_{l_i}\} \cup \{\delta_s^{(r_i)} \mid s \in S_{r_i}\} \subseteq \{\alpha_h^{(i)} \mid h \in H_i\}$ or there are $u_i \in U$ and $r_i \in R$ such that
  $\{\eth_v^{(u_i)} \mid v \in V_{u_i}\} \cup \{\delta_s^{(r_i)} \mid s \in S_{r_i}\} \subseteq \{\alpha_h^{(i)} \mid h \in H_i\}$.

So for all $i \in I$ there is $j_i \in J$ such that $\{\beta_k^{(j_i)} \mid k \in K_{j_i}\} \subseteq \{\alpha_h^{(i)} \mid h \in H_i\}$. $\square$

Corollary 2.6.   *1. If $x:\sigma \to \tau \vdash x:\rho \to \theta$, then $\sigma \to \tau = \rho \to \theta$.*

   2. If $x:\sigma \to \tau \vdash x:(\rho \vee \theta) \wedge \vartheta$, then either $x:\sigma \to \tau \vdash x:\rho \wedge \vartheta$ or $x:\sigma \to \tau \vdash x:\theta \wedge \vartheta$.

   3. Let χ be a union of atomic and arrow types pairwise different. Then $x:\chi \vdash x:\kappa$ implies either $\kappa = \chi$ or $\kappa = \chi \vee \iota$ for some type ι.

Proof. (1),(3). By Lemma 2.5.
(2). By rule $(\wedge E)$, Lemma 2.5 and rule $(\wedge I)$. $\square$

Note that Point (3) of previous corollary holds only under the given condition on type χ, since for example $x:(\varphi \to \varphi) \vee (\varphi \to \varphi) \vdash x:\varphi \to \varphi$.

In the following, as usual, $\Gamma \upharpoonright FV(M)$ denotes the set of premises in Γ whose subjects are the free variable of M.

Lemma 2.7. *Let $\Gamma_x = \Gamma, x:\tau_1 \to \ldots \to \tau_n \to \sigma$ and $\Gamma_x \vdash xM_1 \ldots M_n:\rho$. Then:*

   1. $\Gamma_x \vdash xM_1 \ldots M_n:\sigma$ and $\Gamma \upharpoonright FV(M_i) \vdash M_i:\tau_i$ for $1 \leq i \leq n$;

   2. $y:\sigma \vdash y:\rho$.

Proof. A stronger statement is proved, i.e. that for all types ς:
$$x:\tau_1 \to \ldots \to \tau_n \to \sigma \vdash x:\varsigma \quad \text{and} \quad \Gamma, x:\varsigma \vdash xM_1 \ldots M_n:\rho$$
imply Points (1) and (2) above.
If $m = 0$ Point (1) is immediate, Point (2) follows by rule (L).
For $m > 0$ the proof is by induction on derivations. First note that the last applied rule can be neither (Ax) nor $(\to I)$. If the last applied rule is $(\wedge I)$, $(\vee I)$ or $(\wedge E)$ Points (1) and (2) easily follow.
If the deduction ends with $(\to E)$:

$$(\to E) \quad \frac{\Gamma \upharpoonright FV(xM_1 \ldots M_{n-1}), x:\varsigma \vdash xM_1 \ldots M_{n-1}:\theta \to \rho \quad \Gamma \upharpoonright FV(M_n) \vdash M_n:\theta}{\Gamma, x:\varsigma \vdash xM_1 \ldots M_{n-1}M_n:\rho}$$

By induction, Point (2) implies $y:\tau_n \to \sigma \vdash y:\theta \to \rho$, which gives $\theta = \tau_n$ and $\rho = \sigma$ by Corollary 2.6(1). This shows Point (2) and $\Gamma \upharpoonright FV(M_n) \vdash M_n:\tau_n$. By induction $\Gamma \upharpoonright FV(M_i) \vdash M_i:\tau_i$ for $1 \leq i \leq n-1$ and $\Gamma_x \upharpoonright FV(xM_1 \ldots M_{n-1}) \vdash xM_1 \ldots M_{n-1}:\tau_n \to \sigma$, so rule $(\to E)$ gives $\Gamma_x \vdash xM_1 \ldots M_n:\sigma$ and this concludes the proof of Point (1).
If the deduction ends with $(\vee E)$ two different cases are considered according to the subterms which are the subjects of the premises. In the first case:

$$(\vee E) \quad \frac{\Gamma_1, z:\theta_1 \wedge \zeta \vdash zM_{s+1} \ldots M_n:\rho \quad \Gamma_1, z:\theta_2 \wedge \zeta \vdash zM_{s+1} \ldots M_n:\rho}{\Gamma \upharpoonright FV(M_1, \ldots, M_s), x:\varsigma \vdash xM_1 \ldots M_s:(\theta_1 \vee \theta_2) \wedge \zeta}}{\Gamma, x:\varsigma \vdash xM_1 \ldots M_n:\rho}$$

where $\Gamma_1 = \Gamma \upharpoonright FV(M_{s+1}, \ldots, M_n)$ and $0 \leq s \leq n$. By induction the third premise implies

$$\Gamma_x \upharpoonright FV(M_1, \ldots, M_s) \vdash xM_1 \ldots M_s : \tau_{s+1} \to \ldots \to \tau_n \to \sigma \tag{1}$$

and $\Gamma \upharpoonright FV(M_i) \vdash M_i : \tau_i$ for $1 \le i \le s$ and

$$z : \tau_{s+1} \to \ldots \to \tau_n \to \sigma \vdash z : (\theta_1 \vee \theta_2) \wedge \zeta. \tag{2}$$

Corollary 2.6(2) applied to (2) gives

$$z : \tau_{s+1} \to \ldots \to \tau_n \to \sigma \vdash z : \theta_i \wedge \zeta \tag{3}$$

where either $i = 1$ or $i = 2$. Let $i = 1$, then induction applied to (3) and to the first premise gives

$$\Gamma \upharpoonright FV(M_{s+1} \ldots M_n), z : \tau_{s+1} \to \ldots \to \tau_n \to \sigma \vdash zM_{s+1} \ldots M_n : \sigma \tag{4}$$

and $\Gamma \upharpoonright FV(M_i) \vdash M_i : \tau_i$ for $s + 1 \le i \le n$ and $y : \sigma \vdash y : \rho$ (i.e. Point (2)).
Rule (C) applied to (4) and (1) derives

$$\Gamma_x \vdash xM_1 \ldots M_n : \tau_{m+1} \to \ldots \to \tau_n \to \sigma$$

and this completes the proof of Point (1).

   In the second case:

$$(\vee E) \quad \frac{\Gamma_1, z : \theta_1 \wedge \zeta \vdash xM_1 \ldots M_{s-1} MM_{s+1} \ldots M_n : \rho \qquad \Gamma_1, z : \theta_2 \wedge \zeta \vdash xM_1 \ldots M_{s-1} MM_{s+1} \ldots M_n : \rho \qquad \Gamma \upharpoonright FV(N) \vdash N : (\theta_1 \vee \theta_2) \wedge \zeta}{\Gamma \vdash xM_1 \ldots M_n : \rho}$$

where $\Gamma_1 = \Gamma \upharpoonright FV(M_1, \ldots, M_{s-1} MM_{s+1} \ldots M_n), x : \varsigma$ and $M_s = M[N/z]$. Induction on the first two premises gives:

$$\Gamma \upharpoonright FV(M), z : \theta_1 \wedge \zeta \vdash M : \tau_s \qquad \Gamma \upharpoonright FV(M), z : \theta_2 \wedge \zeta \vdash M : \tau_s \tag{5}$$

so the application of rule $(\vee E)$ to (5) and to the third premise derives $\Gamma \upharpoonright FV(M_s) \vdash M_s : \tau_s$. The other
Points follow by induction. $\qquad\qquad\square$

3   Types of finite hereditary permutators

Aim of this section is to characterise the types derivable for the FHPs. In particular, for an arbitrary FHP
P such that $\vdash P : \sigma \to \tau$, two crucial properties of the disjunctive and conjunctive weak normal forms of
σ and τ are proved:

   P1   if $dw(\sigma) = \bigvee_{i \in I} \mu_i$ and $dw(\tau) = \bigvee_{j \in J} \nu_j$, then for all $i \in I$ there is $j_i \in J$ such that $\vdash P : \mu_i \to \nu_{j_i}$;

   P2   if $cw(\sigma) = \bigwedge_{i \in I} \chi_i$ and $cw(\tau) = \bigwedge_{j \in J} \kappa_j$, then for all $j \in J$ there is $i_j \in I$ such that $\vdash P : \chi_{i_j} \to \kappa_j$.

The content of this section can be summarised as follows:

- Lemma 3.1 gives all possible ways of getting an FHP, possibly with some missing abstractions, as
  result of a substitution (this is useful to deal with rule $(\vee E)$);

- Theorem 3.2 characterises the types derivable for FHPs using disjunctive weak normal form: it
  gives **P1**;

- Lemma 3.3 characterises the types derivable for FHPs, possibly with some missing abstractions,
  using conjunctive weak normal form: it implies **P2**, i.e. Theorem 3.4.

The proof of **P1** is much simpler than that of **P2**. The reason is that Theorem 3.2 uses the property of
union stated in Corollary 2.2(2), while there is no similar property for intersection.

Lemma 3.1. *If* $\lambda xy_1 \ldots y_n.xQ_1 \ldots Q_n$ $(n \ge 0)$ *is an* FHP *and* $\lambda y_{m+1} \ldots y_n.xQ_1 \ldots Q_n = M[N/z]$ *with* $0 \le m \le$
n^1, *then the possible cases are:*

---

[1] for $m = n$ $M[N/z] = xQ_1 \ldots Q_n$.

1. $M = \lambda y_{m+1} \ldots y_n.zQ_{l+1} \ldots Q_n$ *with* $l \leq m$ *and* $N = xQ_1 \ldots Q_l$ *and* $FV(N) \subseteq \{x, y_1, \ldots, y_m\}$;

2. $M = z$ *and* $N = \lambda y_{m+1} \ldots y_n.xQ_1 \ldots Q_n$;

3. $M = \lambda y_{m+1} \ldots y_n.xQ_1 \ldots Q_{j-1}zQ_{j+1} \ldots Q_n$ *and* $N = Q_j$ *and the head variable of* Q_j *belongs to* $\{y_1, \ldots, y_m\}$;

4. $M = \lambda y_{m+1} \ldots y_n.xQ_1 \ldots Q_{j-1}Q'_j Q_{j+1} \ldots Q_n$ *and* $Q'_j = Q_j[z/y_l]$ *and* $N = y_l$, *where* $y_l \in \{y_1, \ldots, y_m\}$ *is the head variable of* Q_j.

Proof. Easy observing that the variables $y_1, ..., y_n$ must be the head variables of $Q_1, ..., Q_n$. □

Theorem 3.2 (Property **P1**). *Let* $dw(\sigma) = \bigvee_{i \in I} \mu_i$, $dw(\tau) = \bigvee_{j \in J} \nu_j$ *and* P *be an* FHP. *Then* $\vdash P : \sigma \to \tau$ *implies that for all* $i \in I$ *there is* $j_i \in J$ *such that* $\vdash P : \mu_i \to \nu_{j_i}$.

Proof. If $P = \lambda x.x$ the proof follows immediately from Lemma 2.5.

Otherwise let $P = \lambda x y_1 \ldots y_n.xQ_1 \ldots Q_n$. By Corollary 2.2(1) $x : \sigma \vdash \lambda y_1 \ldots y_n.xQ_1 \ldots Q_n : \tau$. The proof is by induction on the derivation of $x : \sigma \vdash \lambda y_1 \ldots y_n.xQ_1 \ldots Q_n : \tau$. Assume $dw(\rho) = \bigvee_{h \in H} \lambda_h$ and $dw(\theta) = \bigvee_{k \in K} \xi_k$ and $dw(\vartheta) = \bigvee_{l \in L} \eta_l$. Let $Q = \lambda y_1 \ldots y_n.xQ_1 \ldots Q_n$.

If the last applied rule is $(\vee I)$ the proof is easy.

If the last applied rule is $(\to I)$, then τ is an arrow type. Corollary 2.2(2) gives $x : \mu_i \vdash Q : \tau$ for all $i \in I$.

Let the last applied rule be $(\wedge E)$:

$$(\wedge E) \quad \frac{x : \sigma \vdash Q : \tau \wedge \rho}{x : \sigma \vdash Q : \tau}$$

By definition $dw(\tau \wedge \rho) = \bigvee_{j \in J} \bigvee_{h \in H}(\nu_j \wedge \lambda_h)$. By induction, for all $i \in I$ there are $j_i \in J$ and $h_i \in H$ such that $x : \mu_i \vdash Q : \nu_{j_i} \wedge \lambda_{h_i}$, which gives $x : \mu_i \vdash Q : \nu_{j_i}$ for all $i \in I$ using rule $(\wedge E)$.

Let the last applied rule be $(\wedge I)$:

$$(\wedge I) \quad \frac{x : \sigma \vdash Q : \rho \quad x : \sigma \vdash Q : \theta}{x : \sigma \vdash Q : \rho \wedge \theta}$$

By definition $dw(\rho \wedge \theta) = \bigvee_{h \in H} \bigvee_{k \in K}(\lambda_h \wedge \xi_k)$.

By induction, for all $i \in I$ there is $h_i \in H$ such that $x : \mu_i \vdash Q : \lambda_{h_i}$ and for all $i \in I$ there is $k_i \in K$ such that $x : \mu_i \vdash Q : \xi_{k_i}$. Then rule $(\wedge I)$ derives $x : \mu_i \vdash Q : \lambda_{h_i} \wedge \xi_{k_i}$ for all $i \in I$.

If the last applied rule is $(\vee E)$ by Lemma 3.1 there are two cases to consider.

By definition, $dw(\rho \wedge \theta) = \bigvee_{h \in H} \bigvee_{k \in K}(\lambda_h \wedge \xi_k)$ and $dw(\vartheta \wedge \theta) = \bigvee_{l \in L} \bigvee_{k \in K}(\eta_l \wedge \xi_k)$ and $dw((\rho \vee \vartheta) \wedge \theta) = (\bigvee_{h \in H} \bigvee_{k \in K}(\lambda_h \wedge \xi_k)) \vee (\bigvee_{l \in L} \bigvee_{k \in K}(\eta_l \wedge \xi_k))$. In the first case:

$$(\vee E) \quad \frac{z : \rho \wedge \theta \vdash Q' : \tau \quad z : \vartheta \wedge \theta \vdash Q' : \tau \quad x : \sigma \vdash x : (\rho \vee \vartheta) \wedge \theta}{x : \sigma \vdash Q : \tau}$$

where $Q' = \lambda y_1 \ldots y_n.zQ_1 \ldots Q_n$. By induction:

- on the first premise for all $h \in H$ and $k \in K$ there is $j_{h,k} \in J$ such that $z : \lambda_h \wedge \xi_k \vdash Q' : \nu_{j_{h,k}}$;

- on the second premise for all $l \in L$ and $k \in K$ there is $j_{l,k} \in J$ such that $z : \eta_l \wedge \xi_k \vdash Q' : \nu_{j_{l,k}}$;

- on the third premise for all $i \in I$ either there are $h_i \in H$ and $k_i \in K$ such that $x : \mu_i \vdash x : \lambda_{h_i} \wedge \xi_{k_i}$ or there are $l_i \in L$ and $k_i \in K$ such that $x : \mu_i \vdash x : \eta_{l_i} \wedge \xi_{k_i}$.

So rule (L) implies that for all $i \in I$ either there is $j_{h_i,k_i} \in J$ such that $x:\mu_i \vdash Q:\nu_{j_{h_i,k_i}}$ or there is $j_{l_i,k_i} \in J$ such that $x:\mu_i \vdash Q:\nu_{j_{l_i,k_i}}$. The other possible case is given by:

$$(\vee E) \quad \frac{z:\rho \wedge \theta \vdash z:\tau \quad z:\vartheta \wedge \theta \vdash z:\tau \quad x:\sigma \vdash Q:(\rho \vee \vartheta) \wedge \theta}{x:\sigma \vdash Q:\tau}$$

By induction:

- on the first premise for all $h \in H$ and $k \in K$ there is $j_{h,k} \in J$ such that $z:\lambda_h \wedge \xi_k \vdash z:\nu_{j_{h,k}}$;

- on the second premise for all $l \in L$ and $k \in K$ there is $j_{l,k} \in J$ such that $z:\eta_l \wedge \xi_k \vdash z:\nu_{j_{l,k}}$;

- on the third premise for all $i \in I$ either there are $h_i \in H$ and $k_i \in K$ such that $x:\mu_i \vdash Q:\lambda_{h_i} \wedge \xi_{k_i}$ or there are $l_i \in L$ and $k_i \in K$ such that $x:\mu_i \vdash Q:\eta_{l_i} \wedge \xi_{k_i}$.

So rule (C) implies that for all $i \in I$ either there is $j_{h_i,k_i} \in J$ such that $x:\mu_i \vdash Q:\nu_{j_{h_i,k_i}}$ or there is $j_{l_i,k_i} \in J$ such that $x:\mu_i \vdash Q:\nu_{j_{l_i,k_i}}$. $\qquad\square$

Lemma 3.3. *Let $cw(\sigma) = \bigwedge_{i\in I}\chi_i$, $cw(\tau) = \bigwedge_{j\in J}\kappa_j$ and $\lambda xy_1 \ldots y_n.xQ_1 \ldots Q_n$ $(n \geq 0)$ be an* FHP. *Then $x:\sigma, y_1:\rho_1,\ldots,y_m:\rho_m \vdash \lambda y_{m+1} \ldots y_n.xQ_1 \ldots Q_n:\tau$ and $dw(\rho_h) = \bigvee_{k\in K_h}\mu_k^{(h)}$ $(1 \leq h \leq m \leq n)$ imply that for all $j \in J$ and for all $k_h \in K_h$ $(1 \leq h \leq m)$ there is $i_{j,k_1,\ldots,k_m} \in I$ such that*

$$x:\chi_{i_{j,k_1,\ldots,k_m}}, y_1:\mu_{k_1}^{(1)},\ldots,y_m:\mu_{k_m}^{(m)} \vdash \lambda y_{m+1} \ldots y_n.xQ_1 \ldots Q_n:\kappa_j.$$

Proof. By induction on derivations. If the last applied rule is (Ax) or $(\wedge I)$ the proof is easy.

Assume $cw(\theta) = \bigwedge_{l\in L}\iota_l$ and $cw(\vartheta) = \bigwedge_{s\in S}\upsilon_s$ and $cw(\zeta) = \bigwedge_{t\in T}\omega_t$. Then $cw(\theta \wedge \vartheta) = \bigwedge_{l\in L}\iota_l \wedge \bigwedge_{s\in S}\upsilon_s$ and $cw(\zeta \wedge \vartheta) = \bigwedge_{t\in T}\omega_t \wedge \bigwedge_{s\in S}\upsilon_s$ and $cw((\theta \vee \zeta) \wedge \vartheta) = \bigwedge_{l\in L}\bigwedge_{t\in T}(\iota_l \vee \omega_t) \wedge \bigwedge_{s\in S}\upsilon_s$.

Assume $dw(\theta) = \bigvee_{l\in L}\nu_l$ and $dw(\vartheta) = \bigvee_{s\in S}\lambda_s$ and $dw(\zeta) = \bigvee_{t\in T}\xi_t$. Then $dw(\theta \wedge \vartheta) = \bigvee_{l\in L}\bigvee_{s\in S}(\nu_l \wedge \lambda_s)$ and $dw(\zeta \wedge \vartheta) = \bigvee_{t\in T}\bigvee_{s\in S}(\xi_t \wedge \lambda_s)$ and $dw((\theta \vee \zeta) \wedge \vartheta) = \bigvee_{l\in L}\bigvee_{s\in S}(\nu_l \wedge \lambda_s) \vee \bigvee_{t\in T}\bigvee_{s\in S}(\xi_t \wedge \lambda_s)$.

Let the last applied rule be $(\to I)$:

$$(\to I) \quad \frac{\Gamma, y_{m+1}:\rho_{m+1} \vdash R:\theta}{\Gamma \vdash \lambda y_{m+1}.R:\rho_{m+1} \to \theta}$$

where $\Gamma = x:\sigma, y_1:\rho_1,\ldots,y_m:\rho_m$ and $R = \lambda y_{m+2} \ldots y_n.xQ_1 \ldots Q_n$.
By definition $cw(\rho_{m+1} \to \theta) = \bigwedge_{l\in L}\bigwedge_{k\in K_{m+1}}(\mu_k^{(m+1)} \to \iota_l)$. By induction, for all $l \in L$ and for all $k_h \in K_h$ $(1 \leq h \leq m+1)$ there is $i_{l,k_1,\ldots,k_{m+1}} \in I$ such that $x:\chi_{i_{l,k_1,\ldots,k_{m+1}}}, y_1:\mu_{k_1}^{(1)},\ldots,y_{m+1}:\mu_{k_{m+1}}^{(m+1)} \vdash R:\iota_l$, so the application of rule $(\to I)$ concludes the proof.

Let the last applied rule be $(\to E)$:

$$(\to E) \quad \frac{\Gamma \vdash R:\theta \to \tau \quad y_{\pi(r)}:\rho_{\pi(r)} \vdash Q_r:\theta}{\Gamma \vdash RQ_r:\tau}$$

where $\Gamma = x:\sigma, y_{\pi(1)}:\rho_{\pi(1)},\ldots,y_{\pi(r-1)}:\rho_{\pi(r-1)}$ and $R = xQ_1 \ldots Q_{r-1}$. If $dw(\theta) = \bigvee_{u\in U}\nu_u$, then $cw(\theta \to \tau) = \bigwedge_{u\in U}\bigwedge_{j\in J}(\nu_u \to \kappa_j)$. By induction for all $u \in U$, $j \in J$ and for all $k_{\pi(s)} \in K_{\pi(s)}$ $(1 \leq s \leq r-1)$ there is $i_{(u,j),k_{\pi(1)},\ldots,k_{\pi(r-1)}} \in I$ such that

$$x:\chi_{i_{(u,j),k_{\pi(1)},\ldots,k_{\pi(r-1)}}}, y_1:\mu_{k_{\pi(1)}}^{(\pi(1))},\ldots,y_{\pi(r-1)}:\mu_{k_{\pi(r)}}^{(\pi(r-1))} \vdash R:\nu_u \to \kappa_j. \tag{6}$$

By Theorem 3.2 for all $k_{\pi(r)} \in K_{\pi(r)}$ there is $u_{k_{\pi(r)}} \in U$ such that

$$y_{\pi(r)}:\mu_{k_{\pi(r)}}^{(\pi(r))} \vdash Q_r:\nu_{u_{k_{\pi(r)}}}. \tag{7}$$

Choosing $u = u_{k_{\pi(r)}}$ in (6) the application of rule $(\to E)$ to (6) and (7) gives the result.

Let the last applied rule be ($\wedge E$):

$$(\wedge E) \quad \frac{\Gamma \vdash R : \tau \wedge \theta}{\Gamma \vdash R : \tau}$$

where $\Gamma = x : \sigma, y_1 : \rho_1, \dots, y_m : \rho_m$ and $R = \lambda y_{m+1} \dots y_n . x Q_1 \dots Q_n$. Since $cw(\tau \wedge \theta) = cw(\tau) \wedge cw(\theta)$, this case easily follows by induction.

Let the last applied rule be ($\vee I$):

$$(\vee I) \quad \frac{\Gamma \vdash R : \theta}{\Gamma \vdash R : \theta \vee \vartheta}$$

where $\Gamma = x : \sigma, y_1 : \rho_1, \dots, y_m : \rho_m$ and $R = \lambda y_{m+1} \dots y_n . x Q_1 \dots Q_n$. Since $cw(\theta \vee \vartheta) = \bigwedge_{l \in L} \bigwedge_{s \in S} (\iota_l \vee \upsilon_s)$, this case easily follows by induction.

If the last applied rule is ($\vee E$) there are four different cases as prescribed by Lemma 3.1. In the first case:

$$(\vee E) \quad \frac{\Gamma_1, z : \theta \wedge \vartheta \vdash R : \tau \quad , \Gamma_1, z : \zeta \wedge \vartheta \vdash R : \tau \quad \Gamma_2 \vdash x Q_1 \dots Q_u : (\theta \vee \zeta) \wedge \vartheta}{\Gamma_1, \Gamma_2 \vdash \lambda y_{m+1} \dots y_n . x Q_1 \dots Q_n : \tau}$$

where $R = \lambda y_{m+1} \dots y_n . z Q_{u+1} \dots Q_n$, $\Gamma_1 = \{y_1 : \rho_1, \dots, y_m : \rho_m\} \upharpoonright FV(R)$,
$\Gamma_2 = x : \sigma, y_{\pi(1)} : \rho_{\pi(1)}, \dots, y_{\pi(u)} : \rho_{\pi(u)}$ and $(0 \le u \le m)$. Let $\Gamma_1 = y_{w_1} : \rho_{w_1}, \dots, y_{w_{m-u}} : \rho_{w_{m-u}}$.

.

For all $j \in J$ and for all $k_{w_v} \in K_{w_v}$ $(1 \le v \le m - u)$ by induction :

- on the first premise either there is $l_{j, k_{w_1}, \dots, k_{w_{m-u}}} \in L$ such that
  $z : \iota_{l_{j,k_{w_1},\dots,k_{w_{m-u}}}}, y_{w_1} : \mu_{k_{w_1}}^{(w_1)}, \dots, y_{w_{m-u}} : \mu_{k_{w_{m-u}}}^{(w_{m-u})} \vdash R : \kappa_j$ or there is
  $s_{j, k_{w_1}, \dots, k_{w_{m-u}}} \in S$ such that $z : \upsilon_{s_{j,k_{w_1},\dots,k_{w_{m-u}}}}, y_{w_1} : \mu_{k_{w_1}}^{(w_1)}, \dots, y_{w_{m-u}} : \mu_{k_{w_{m-u}}}^{(w_{m-u})} \vdash R : \kappa_j$;

- on the second premise either there is $t_{j, k_{u+1}, \dots, k_m} \in T$ such that
  $z : \omega_{t_{j,k_{w_1},\dots,k_{w_{m-u}}}}, y_{w_1} : \mu_{k_{w_1}}^{(w_1)}, \dots, y_{w_{m-u}} : \mu_{k_{w_{m-u}}}^{(w_{m-u})} \vdash R : \kappa_j$ or there is
  $s_{j, k_{w_1}, \dots, k_{w_{m-u}}} \in S$ such that $z : \upsilon_{s_{j,k_{w_1},\dots,k_{w_{m-u}}}}, y_{w_1} : \mu_{k_{w_1}}^{(w_1)}, \dots, y_{w_{m-u}} : \mu_{k_{w_{m-u}}}^{(w_{m-u})} \vdash R : \kappa_j$.

Therefore for all $j \in J$ and for all $k_{w_v} \in K_{w_v}$ $(1 \le v \le m - u)$:

- either there is $l_{j, k_{w_1}, \dots, k_{w_{m-u}}} \in L$ such that

$$z : \iota_{l_{j,k_{w_1},\dots,k_{w_{m-u}}}}, y_{w_1} : \mu_{k_{w_1}}^{(w_1)}, \dots, y_{w_{m-u}} : \mu_{k_{w_{m-u}}}^{(w_{m-u})} \vdash R : \kappa_j \tag{8}$$

  and there is $t_{j, k_{w_1}, \dots, k_{w_{m-u}}} \in T$ such that

$$z : \omega_{t_{j,k_{w_1},\dots,k_{w_{m-u}}}}, y_{w_1} : \mu_{k_{w_1}}^{(w_1)}, \dots, y_{w_{m-u}} : \mu_{k_{w_{m-u}}}^{(w_{m-u})} \vdash R : \kappa_j; \tag{9}$$

- or there is $s_{j, k_{u+1}, \dots, k_m} \in S$ such that

$$z : \upsilon_{s_{j,k_{w_1},\dots,k_{w_{m-u}}}}, y_{w_1} : \mu_{k_{w_1}}^{(w_1)}, \dots, y_{w_{m-u}} : \mu_{k_{w_{m-u}}}^{(w_{m-u})} \vdash R : \kappa_j. \tag{10}$$

By induction the third premise implies that for all $l \in L$, $t \in T$ and for all $k_{\pi(h)} \in K_{\pi(h)}$ $(1 \le h \le u)$ there is $i_{l,t,k_{\pi(1)},\dots,k_{\pi(u)}} \in I$ such that

$$x : \chi_{i_{l,t,k_{\pi(1)},\dots,k_{\pi(u)}}}, y_{\pi(1)} : \mu_{k_{\pi(1)}}^{(\pi(1))}, \dots, y_{\pi(u)} : \mu_{k_{\pi(u)}}^{(\pi(u))} \vdash x Q_1 \dots Q_u : \iota_l \vee \omega_t \tag{11}$$

and for all $s \in S$ and for all $k_{\pi(h)} \in K_{\pi(h)}$ $(1 \le h \le u)$ there is $i_{s,k_{\pi(1)},\dots,k_{\pi(u)}} \in I$ such that

$$x : \chi_{i_{s,k_{\pi(1)},\dots,k_{\pi(u)}}}, y_{\pi(1)} : \mu_{k_{\pi(1)}}^{(\pi(1))}, \dots, y_{\pi(u)} : \mu_{k_{\pi(u)}}^{(\pi(u))} \vdash x Q_1 \dots Q_u : \upsilon_s. \tag{12}$$

If (8) and (9) hold, then the conclusion follows from the application of rule ($\vee E$) to (8), (9) and (11) by choosing $l = l_{j, k_{w_1}, \dots, k_{w_{m-u}}}$ and $t = t_{j, k_{w_1}, \dots, k_{w_{m-u}}}$. Otherwise (10) must hold, and the conclusion follows from the application of rule (C) to (10) and (12) by choosing $s = s_{j, k_{u+1}, \dots, k_m}$.

In the second case:

$$(\vee E) \quad \frac{z:\theta\wedge\vartheta\vdash z:\tau \quad z:\zeta\wedge\vartheta\vdash z:\tau \quad \Gamma\vdash \lambda y_{m+1}\ldots y_n.xQ_1\ldots Q_n:(\theta\vee\zeta)\wedge\vartheta}{\Gamma\vdash \lambda y_{m+1}\ldots y_n.xQ_1\ldots Q_n:\tau}$$

where $\Gamma = x:\sigma, y_1:\rho_1,\ldots,y_m:\rho_m$.

For all $j \in J$ by induction:

- the first premise implies that either there is $l_j \in L$ such that $z:\iota_{l_j}\vdash z:\kappa_j$ or there is $s_j \in S$ such that $z:\upsilon_{s_j}\vdash z:\kappa_j$;

- the second premise implies that either there is $t_j \in T$ such that $z:\omega_{t_j}\vdash z:\kappa_j$ or there is $s_j \in S$ such that $z:\upsilon_{s_j}\vdash z:\kappa_j$.

Therefore for all $j \in J$:

- either there are $l_j \in L$ and $t_j \in T$ such that

- or there is $s_j \in S$ such that $\qquad\quad z:\iota_{l_j}\vdash z:\kappa_j \qquad z:\omega_{t_j}\vdash z:\kappa_j;$ $\hfill(13)$
$$z:\upsilon_{s_j}\vdash z:\kappa_j. \hfill(14)$$

By induction on the third premise for all $l \in L$, $t \in T$ and for all $k_h \in K_h$ ($1 \le h \le m$) there is $i_{l,k_1,\ldots,k_m} \in I$ such that

$$x:\chi_{i_{l,k_1,\ldots,k_m}}, y_1:\mu_{k_1}^{(1)},\ldots,y_m:\mu_{k_m}^{(m)}\vdash \lambda y_{m+1}\ldots y_n.xQ_1\ldots Q_n:\iota_l\vee\omega_t \hfill(15)$$

and for all $s \in S$ and $k_h \in K_h$ ($1 \le h \le m$) there is $i_{s,k_1,\ldots,k_m} \in I$ such that

$$x:\chi_{s_{l,k_1,\ldots,k_m}}, y_1:\mu_{k_1}^{(1)},\ldots,y_m:\mu_{k_m}^{(m)}\vdash \lambda y_{m+1}\ldots y_n.xQ_1\ldots Q_n:\upsilon_s. \hfill(16)$$

If (13) holds, then the conclusion follows from the application of rule $(\vee E)$ to (13) and (15) by choosing $l = l_j$ and $t = t_j$. Otherwise (14) must hold, and the conclusion follows from the application of rule (C) to (14) and (16) by choosing $s = s_j$.

As for the third case:

$$(\vee E) \quad \frac{\Gamma, z:\theta\wedge\vartheta\vdash R:\tau \quad \Gamma, z:\zeta\wedge\vartheta\vdash R:\tau \quad y_u:\rho_u\vdash Q_v:(\theta\vee\zeta)\wedge\vartheta}{\Gamma\vdash \lambda y_{m+1}\ldots y_n.xQ_1\ldots Q_n:\tau}$$

where $\Gamma = x:\sigma, y_1:\rho_1,\ldots,y_{u-1}:\rho_{u-1}, y_{u+1}:\rho_{u+1},\ldots,y_m:\rho_m$, $R = \lambda y_{m+1}\ldots y_n.xQ_1\ldots Q_{v-1}zQ_{v+1}\ldots Q_n$ and $u = \pi(v)$.

By induction on the first premise for all $j \in J$, $l \in L$, $s \in S$, and for all $k_h \in K_h$ ($1 \le h \le m, h \ne u$), there is $i_{(j,l,s),k_1,\ldots,k_{u-1},k_{u+1},\ldots,k_m} \in I$ such that

$$x:\chi_{i_{(j,l,s),k_1,\ldots,k_{u-1},k_{u+1},\ldots,k_m}}, y_1:\mu_{k_1}^{(1)},\ldots,y_{u-1}:\mu_{k_{u-1}}^{(u-1)}, y_{u+1}:\mu_{k_{u+1}}^{(u+1)},\ldots,y_m:\mu_{k_m}^{(m)}, z:\nu_l\wedge\lambda_s\vdash R:\kappa_j. \hfill(17)$$

By induction on the second premise for all $j \in J$, $t \in T$, $s \in S$, and for all $k_h \in K_h$ ($1 \le h \le m+1, h \ne u$), there is $i_{(j,t,s),k_1,\ldots,k_{u-1},k_{u+1},\ldots,k_m} \in I$ such that

$$x:\chi_{i_{(j,t,s),k_1,\ldots,k_{u-1},k_{u+1},\ldots,k_m}}, y_1:\mu_{k_1}^{(1)},\ldots,y_{u-1}:\rho_{u-1}, y_{u+1}:\rho_{u+1},\ldots,y_m:\mu_{k_m}^{(m)}, z:\xi_t\wedge\lambda_s\vdash R:\kappa_j. \hfill(18)$$

By Theorem 3.2 applied to the third premise for all $k_u \in K_u$:

either there are $l_{k_u} \in L$, $s_{k_u} \in S$ such that $y_u:\mu_{k_u}^{(u)}\vdash Q_v:\nu_{l_{k_u}}\wedge\lambda_{s_{k_u}};$ $\hfill(19)$

or there are $t_{k_u} \in T$, $s_{k_u} \in S$ such that $y_u:\mu_{k_u}^{(u)}\vdash Q_v:\xi_{t_{k_u}}\wedge\lambda_{s_{k_u}}.$ $\hfill(20)$

If (19) holds, then the conclusion follows from the application of rule (C) to (17) and (19) by choosing $l = l_{k_u}$ and $s = s_{k_u}$. Otherwise (18) must hold, and the conclusion follows from the application of rule (C) to (18) and (20) by choosing $t = t_{k_u}$ and $s = s_{k_u}$.

The proof for the last case is similar and simpler than that one of the third case. $\hfill\square$

Theorem 3.4 (Property **P2**). *If P is an* FHP, $cw(\sigma) = \bigwedge_{i\in I}\chi_i$, $cw(\tau) = \bigwedge_{j\in J}\kappa_j$, *and* $\vdash P:\sigma\to\tau$, *then for all $j \in J$ there is $i_j \in I$ such that* $\vdash P:\chi_{i_j}\to\kappa_j$.

4 Normalisation of types

To investigate type isomorphism it is necessary to consider the basic laws introduced in Lemma 2.4, for finding conditions allowing to apply them also at the level of subtypes and to exploit some provable properties of type inclusion. To this aim, following a common approach [2, 7], a notion of *normal form* of types is introduced. *Normal type* is short for type in normal form.

Normal types are obtained by applying as far as possible a set of isomorphism preserving transformations, that are all realised by suitable η-expansions of the identity. The transformations applied to obtain normal types are essentially:

- the distribution of intersections over unions or vice versa, in such a way that all types to the right of an arrow are in conjunctive normal form and all types to the left of an arrow are in disjunctive normal form. This is obtained by using (**dist**$\vee\wedge$) and (**dist**$\wedge\vee$) (*distribution*);

- the elimination of intersections to the right of arrows and of unions to the left of arrows using the isomorphisms (**dist**$\rightarrow\wedge$) and (**dist**$\rightarrow\vee$) from left to right (*splitting*);

- the elimination of redundant intersections and unions, corresponding roughly to intersections and unions performed on types provably included in one another, as $(\sigma \rightarrow \tau) \wedge (\sigma \vee \rho \rightarrow \tau)$, that can be reduced to $\sigma \vee \rho \rightarrow \tau$; similarly $(\sigma \rightarrow \rho \vee \tau) \wedge (\sigma \rightarrow \tau)$ can be reduced to $\sigma \rightarrow \tau$ (*erasure*);

- the transformation of types at top level in conjunctive normal form.

For example the type $((\varphi_1 \wedge \varphi_2 \rightarrow \varphi_2 \vee \varphi_3) \vee (\varphi_2 \rightarrow \varphi_5)) \wedge ((\varphi_2 \wedge \varphi_3 \rightarrow \varphi_5) \vee (\varphi_4 \rightarrow \varphi_3 \vee \varphi_5))$ is normal.

The normalisation process, although rather intuitive, needs some care when performed inside a type context since the used transformations must be isomorphism preserving.

The following subsection defines the normalisation rules. In Subsection 4.2 the soundness and termination of the normalisation rules and the unicity of normal forms are proved. The notion of normal form is effective since an algorithm to find the normal form of an arbitrary type can be given. Lastly Subsection 4.3 presents interesting properties of normal types, in particular Theorem 4.15 characterises the isomorphic normal types.

4.1 Normalisation rules

Since the normalisation rules have to be applied (whenever possible) also to subtypes, the (standard) notion of type *context* is introduced.

$$C[\,] ::= [\,] \mid C[\,] \rightarrow \sigma \mid \sigma \rightarrow C[\,] \mid \sigma \wedge C[\,] \mid C[\,] \wedge \sigma \mid \sigma \vee C[\,] \mid C[\,] \vee \sigma.$$

The possibility of applying transformations to subtypes strongly depends on the context in which they occur. An example of this problem was already given at page 2. Also the types $(\sigma \vee \tau \rightarrow \rho) \wedge (\sigma \rightarrow \rho)$ and $\sigma \vee \tau \rightarrow \rho$, are isomorphic in the context [ ], with $\lambda xy.xy$ showing the isomorphism. But the same types are not isomorphic in the context $[\,] \wedge \varphi$, because no η-expansion of the identity can map an atomic type into itself.

To formalise this notion, *paths of type contexts* are useful (Definition 4.2). The path of a context describes which arrows need to be traversed in order to reach the hole, if it is possible, i.e. when there are no atoms on the way. It is handy to have a notion of *agreement of a type with a path* (Definition 4.1(3)), in order to assure that the types which are composed by intersection or union with the type context do not block the transformation. An intersection or a union agrees with a path only if all types belonging to the intersection or to the union agree with that path.

In paths the symbol $\nearrow$ represents going down to the left of an arrow and the symbol $\searrow$ represents going down to the right of an arrow. For distribution rules it is enough to reach the hole, while for splitting rules one more arrow needs to be traversed. So two kinds of paths are useful. They are dubbed d-paths and s-paths, being used in distribution and splitting rules, respectively. An s-path is a d-path terminated by the symbol $\square$.

The agreement of a type with a set of d-paths (Definition 4.1(4)) and the concatenation of d-paths (Definition 4.1(5)) are useful for defining the erasure rules (Definition 4.5).

Definition 4.1.   *1. A* d-path p *is a possibly empty string on the alphabet* $\{\nearrow, \searrow\}$.

   2. An s-path p *is a d-path followed by* $\square$.

   3. The agreement *of a type* σ *with a d-path or an s-path* p *(notation* $\sigma \propto$ p*) is the smallest relation between types and d-paths (s-paths) such that:*

$\sigma \propto \epsilon$ *for all* σ; $\tau \to \rho \propto \square$ *for all* τ, ρ;

$\tau \propto$ p *implies* $\tau \to \rho \propto_{\nearrow}$ p; $\rho \propto$ p *implies* $\tau \to \rho \propto_{\searrow}$ p;

$\tau \propto$ p *and* $\rho \propto$ p *imply* $\tau \wedge \rho \propto$ p; $\tau \propto$ p *and* $\rho \propto$ p *imply* $\tau \vee \rho \propto$ p.

   4. A type σ agrees *with a set of d-paths* $\mathcal{P}$ *(notation* $\sigma \propto \mathcal{P}$*) if it agrees with all the d-paths in* $\mathcal{P}$, *i.e.* $\sigma \propto$ p *for all* p $\in \mathcal{P}$.

   5. If p *and* p$'$ *are d-paths,* p$\cdot$p$'$ *denotes their* concatenation; *if* $\mathcal{P}$ *is a set of d-paths,* p$\cdot\mathcal{P}$ *denotes the set* $\{$p$\cdot$p$' \mid$ p$' \in \mathcal{P}\} \cup \{p\}$.

For example the type $\sigma_1 \to (\sigma_2 \to \rho_1 \wedge \rho_2) \wedge (\sigma_3 \vee \sigma_1 \to \tau_1) \to \tau_2$ agrees with the d-path $\searrow\nearrow\nearrow$ and with the s-path $\searrow\nearrow \square$, while the type $\sigma_1 \to (\sigma_2 \to \rho_1 \wedge \rho_2) \wedge (\sigma_3 \vee \sigma_1 \to \tau_1) \wedge \varphi \to \tau_2$ agrees with the d-path $\searrow\nearrow$ and with the s-path $\searrow \square$, but it does not agree with the d-path $\searrow\nearrow\nearrow$ nor with the s-path $\searrow\nearrow \square$, since φ does not agree with $\nearrow$ nor with $\square$.

The d-paths and s-paths of contexts can be formalised using the agreement between types and paths.

Definition 4.2. *The d-path and the s-path of a type context* $C[\ ]$ *(notations* $d(C[\ ])$ *and* $s(C[\ ])$, *respectively) are defined by induction on* $C[\ ]$:

$$d(C[\ ]) = \epsilon \text{ if } C[\ ] = [\ ]; \qquad s(C[\ ]) = \square \text{ if } C[\ ] = [\ ];$$
$$*(C'[\ ]) = p \text{ implies } *(C[\ ]) = \nearrow p \text{ if } C[\ ] = C'[\ ] \to \sigma \text{ and } *(C[\ ]) = \searrow p \text{ if } C[\ ] = \sigma \to C'[\ ];$$
$$\sigma \propto *(C'[\ ]) \text{ implies } *(C[\ ]) = *(C'[\ ]) \text{ if } C[\ ] = C'[\ ] \wedge \sigma \text{ or } C[\ ] = \sigma \wedge C'[\ ] \text{ or}$$
$$C[\ ] = C'[\ ] \vee \sigma \text{ or } C[\ ] = \sigma \vee C'[\ ].$$

where $*$ *holds for* d *and* s.

For example the d-path and the s-path of the context $\sigma_1 \to [\ ] \wedge (\sigma_2 \to \tau_1 \vee \tau_2) \to \tau_2$ are $\searrow\nearrow$ and $\searrow\nearrow \square$, respectively, while the d-path and the s-path of the context $\sigma_1 \to ([\ ] \wedge \sigma_2 \to \tau_1) \vee \varphi \to \tau_2$ are undefined, since $\varphi \not\propto \nearrow$ and $\varphi \not\propto \square$.

In giving the normalisation rules one can consider types in holes modulo idempotence, commutativity and associativity, when the d-paths of contexts are defined. This is assured by the following lemma, that can be easily proved by induction on d-paths.

Lemma 4.3. *If* $\sigma \approx \tau$ *holds the isomorphisms* (**idem**), (**comm**), (**assoc**), *and* $d(C[\ ])$ *is defined, then* $C[\sigma] \approx C[\tau]$.

Distribution and splitting rules can now be defined.

Definition 4.4 (Distribution and Splitting).   *1. The two distribution rules are:*

$$C[(\sigma \wedge \tau) \vee \rho] \Longrightarrow C[(\sigma \vee \rho) \wedge (\tau \vee \rho)] \qquad \text{if } d(C[\ ]) = \epsilon \text{ or } d(C[\ ]) = \mathsf{p} \cdot \searrow \text{ for some path } \mathsf{p};$$
$$C[(\sigma \vee \tau) \wedge \rho] \Longrightarrow C[(\sigma \wedge \rho) \vee (\tau \wedge \rho)] \qquad \text{if } d(C[\ ]) = \mathsf{p} \cdot \nearrow \text{ for some path } \mathsf{p}.$$

2. *The two splitting rules are:*

$$C[\sigma \to \tau \wedge \rho] \Longrightarrow C[(\sigma \to \tau) \wedge (\sigma \to \rho)] \qquad \text{if } s(C[\ ]) \text{ is defined;}$$
$$C[\sigma \vee \tau \to \rho] \Longrightarrow C[(\sigma \to \rho) \wedge (\tau \to \rho)] \qquad \text{if } s(C[\ ]) \text{ is defined.}$$

The conditions for erasure rules use two preorders on types, defined in Figure 3 between basic intersections and between basic unions (see page 3), respectively. This is enough since the distribution and splitting rules (when applicable) give arrow types with basic intersections as left-hand-sides and basic unions as right-hand-sides. The symbol $\leq^\diamond$ stands for either $\leq^\wedge$ or $\leq^\vee$. It easy to verify that $\alpha \leq^\wedge \beta$ if and only if $\alpha \leq^\vee \beta$, so comparing two arrows or two atomic types one can write $\alpha \leq^\diamond \beta$. For example $\mu \wedge \nu \leq^\wedge \mu$ and $\chi \leq^\vee \chi \vee \kappa$ imply $\mu \to \chi \leq^\diamond \mu \wedge \nu \to \chi \vee \kappa$ and $(\mu \wedge \nu \to \chi \vee \kappa) \to \iota \leq^\diamond (\mu \to \chi) \to \iota$.

It is easy to show that $\leq^\wedge$ and $\leq^\vee$ are preorders since transitivity holds. The presence, at top level, of an atomic type on both sides of $\leq^\diamond$ forces atomic and arrow types to be only erased or added. In relating types one can exploit also idempotence. For instance two copies of $(\mu \to \chi)$ are needed in deriving $\mu \to \chi \leq^\wedge (\mu \wedge \nu \to \chi) \wedge (\mu \to \chi \vee \kappa)$ to show $\mu \to \chi \leq^\diamond \mu \wedge \nu \to \chi$ and $\mu \to \chi \leq^\diamond \mu \to \chi \vee \kappa$.

$$\mu \leq^\wedge \mu \qquad \chi \leq^\vee \chi \qquad \varphi \wedge \mu \leq^\wedge \varphi \qquad \varphi \leq^\vee \varphi \vee \chi$$
$$\varphi \wedge \mu \wedge \lambda \leq^\wedge \varphi \wedge \mu \qquad \varphi \vee \chi \leq^\vee \varphi \vee \chi \vee \iota$$
$$\nu_i \leq^\wedge \mu_i, \chi_i \leq^\vee \kappa_i \text{ for all } i \in I \;\Rightarrow\; \bigwedge_{i \in I}(\mu_i \to \chi_i)[\wedge\lambda] \leq^\wedge \bigwedge_{i \in I}(\nu_i \to \kappa_i)$$
$$\nu_i \leq^\wedge \mu_i, \chi_i \leq^\vee \kappa_i \text{ for all } i \in I \;\Rightarrow\; \bigvee_{i \in I}(\mu_i \to \chi_i) \leq^\vee \bigvee_{i \in I}(\nu_i \to \kappa_i)[\vee\iota]$$

where the notation $[\wedge\lambda]$ ($[\vee\iota]$) means that $\wedge\lambda$ ($\vee\iota$) can either occur or not.

Figure 3: Preorders on types.

$$e(\mu \leq^\wedge \mu) = e(\chi \leq^\vee \chi) = \{\}$$

$$e(\varphi \wedge \mu \leq^\wedge \varphi) = e(\varphi \leq^\vee \varphi \vee \chi) = e(\varphi \wedge \mu \wedge \lambda \leq^\wedge \varphi \wedge \mu) = e(\varphi \vee \chi \leq^\vee \varphi \vee \chi \vee \iota) = \{\epsilon\}$$

$$\begin{aligned} &e(\bigwedge_{i \in I}(\mu_i \to \chi_i)[\wedge\lambda] \leq^\wedge \bigwedge_{i \in I}(\nu_i \to \kappa_i)) \\ &e(\bigvee_{i \in I}(\mu_i \to \chi_i) \leq^\vee \bigvee_{i \in I}(\nu_i \to \kappa_i)[\vee\iota]) \end{aligned} = \begin{cases} \{\epsilon\} & \text{if } \lambda \text{ or } \iota \text{ is present and} \\ & \quad e(\nu_i \leq^\wedge \mu_i) = e(\chi_i \leq^\vee \kappa_i) = \{\} \text{ for all } i \in I, \\ \bigcup_{i \in I}(\nearrow \cdot e(\nu_i \leq^\wedge \mu_i) \cup \searrow \cdot e(\chi_i \leq^\vee \kappa_i)) & \text{otherwise} \end{cases}$$

$$\text{if } \nu_i \leq^\wedge \mu_i, \chi_i \leq^\vee \kappa_i \text{ for all } i \in I$$

Figure 4: Set of d-paths of a preorder derivation.

These preorders are crucial for the definition of the erasure rules. In fact some types in an intersection can be erased only if the remaining types are smaller or equal to the erased ones. Dually some types in a union can be erased only if the remaining types are bigger or equal to the erased ones. Another necessary condition for erasing types is that the FHPs can reach the subtypes in which the types related by the preorder differ. In order to formalise this, one d-path is not enough, since there can be many subtypes in which the types differ, so sets of d-paths are needed. Sets of d-paths are then associated with derivations of preorders between types, so that one can check when a type can be erased in a type context. The set of d-paths of $\sigma \leq^\diamond \tau$ (notation $e(\sigma \leq^\diamond \tau)$) represents the set of paths that make accessible the points in which

σ and τ differ. For this reason, $e(\mu \leq^\wedge \mu)$ and $e(\chi \leq^\vee \chi)$ are defined as the empty set and the sets:

$$e(\varphi \wedge \mu \leq^\wedge \varphi), \; e(\varphi \leq^\vee \varphi \vee \chi), \; e(\varphi \wedge \mu \wedge \lambda \leq^\wedge \varphi \wedge \mu), \; \text{and} \; e(\varphi \vee \chi \leq^\vee \varphi \vee \chi \vee \iota)$$

contain only ϵ; in the other cases this set must be built from the sets of paths associated with the subtypes using $\nearrow$ and $\searrow$. This definition is given in Figure 4. Notice that the condition $e(\sigma \leq^\diamond \tau) = \{\}$ implies $\sigma = \tau$.

For example $e(\mu \to \chi \leq^\diamond \mu \wedge \nu \to \chi \vee \kappa) = \{\nearrow, \searrow\}$ and $e((\mu \wedge \nu \to \chi \vee \kappa) \to \iota \leq^\diamond (\mu \to \chi) \to \iota) = \{\nearrow\nearrow, \nearrow\searrow\}$.

Finally one can define erasure rules.

Definition 4.5 (Erasure). *The three* erasure rules *are:*

$$\bigwedge_{i \in I} \chi_i \Longrightarrow \bigwedge_{j \in J} \chi_j \qquad \textit{if } J \subset I \textit{ and } \forall i \in I \; \exists j_i \in J. \chi_{j_i} \leq^\vee \chi_i \textit{ and } \forall j \in J. \chi_j \propto \mathcal{P},$$
$$\textit{where } \mathcal{P} = \bigcup_{i \in I} e(\chi_{j_i} \leq^\vee \chi_i);$$

$$C[\bigwedge_{i \in I} \alpha_i] \Longrightarrow C[\bigwedge_{j \in J} \alpha_j] \qquad \textit{if } J \subset I \textit{ and } \forall i \in I \; \exists j_i \in J. \alpha_{j_i} \leq^\diamond \alpha_i \textit{ and } \forall j \in J. C[\alpha_j] \propto \mathcal{P},$$
$$\textit{where } \mathcal{P} = d(C[]) \cdot \bigcup_{i \in I} e(\alpha_{j_i} \leq^\diamond \alpha_i);$$

$$C[\bigvee_{i \in I} \alpha_i] \Longrightarrow C[\bigvee_{j \in J} \alpha_j] \qquad \textit{if } J \subset I \textit{ and } \forall i \in I \; \exists j_i \in J. \alpha_i \leq^\diamond \alpha_{j_i} \textit{ and } \forall j \in J. C[\alpha_j] \propto \mathcal{P},$$
$$\textit{where } \mathcal{P} = d(C[]) \cdot \bigcup_{i \in I} e(\alpha_i \leq^\diamond \alpha_{j_i}).$$

In the first erasure rule the absence of the context indicates that it can be applied only at top level, i.e. in the empty context.

By applying the erasure rules, it is essential to allow to remove more than one type in a single step. For example $(\mu \to \varphi \to \chi) \wedge (\mu \to (\varphi \wedge \nu \to \chi) \vee \psi_1) \wedge (\mu \to (\varphi \wedge \nu \to \chi) \vee \psi_2) \Longrightarrow \mu \to \varphi \to \chi$, but this type does not reduce to $(\mu \to \varphi \to \chi) \wedge (\mu \to (\varphi \wedge \nu \to \chi) \vee \psi_i)$ for $i = 1$ or $i = 2$. The problem is that $\mu \to (\varphi \wedge \nu \to \chi) \vee \psi_1$ does not agree with $e(\mu \to \varphi \to \chi \leq^\diamond \mu \to (\varphi \wedge \nu \to \chi) \vee \psi_2) = \{\searrow\nearrow\}$ and dually exchanging ψ_1 with ψ_2.

Normalisation can create redexes, for example the first distribution rule applied to $\sigma \to (\tau \wedge \rho) \vee \theta$ gives $\sigma \to (\tau \vee \theta) \wedge (\rho \vee \theta)$, which can be reduced to $(\sigma \to \tau \vee \theta) \wedge (\sigma \to \rho \vee \theta)$ by the first splitting rule. The second splitting rule applied to $(\sigma \vee \varphi \to \varphi) \wedge (\varphi \wedge \psi \to \varphi)$ gives $(\sigma \to \varphi) \wedge (\varphi \to \varphi) \wedge (\varphi \wedge \psi \to \varphi)$, which can be reduced to $(\sigma \to \varphi) \wedge (\varphi \to \varphi)$ by the first or second erasure rule. A more interesting example is $(\varphi \wedge (\psi \to \psi) \to \psi) \wedge ((\psi \to \psi) \to \psi) \wedge (((\sigma \vee \tau) \wedge \rho \to \rho) \to \rho)$: this type can only be reduced to $((\psi \to \psi) \to \psi) \wedge (((\sigma \vee \tau) \wedge \rho \to \rho) \to \rho)$ by the first or second erasure rule and then the second distribution rule becomes applicable.

4.2   Soundness, confluence and termination of type normalisation

The soundness of the normalisation rules, i.e. that $\sigma \Longrightarrow \tau$ implies $\sigma \approx \tau$, uses η-expansions of the identity, called *finite hereditarily identities (*FHI*s)*. More precisely for each rule $\sigma \Longrightarrow \tau$ two FHIs $\mathsf{Id}, \mathsf{Id}'$ such that $\vdash \mathsf{Id}: \sigma \to \tau$ and $\vdash \mathsf{Id}': \tau \to \sigma$ are built. FHIs can be associated with d-paths, s-paths and sets of d-paths.

Definition 4.6.   *1. The* FHI *induced by the s-path* p *(notation* $\mathsf{Id_p}$*) is defined by induction on* p*:*

$$\mathsf{Id_\square} = \lambda xy.xy \qquad \mathsf{Id}_{\nearrow \mathsf{p} \; \beta} \longleftarrow \lambda xy.x(\mathsf{Id_p}y) \qquad \mathsf{Id}_{\searrow \mathsf{p} \; \beta} \longleftarrow \lambda xy.\mathsf{Id_p}(xy)$$

2. The FHI *induced by the set of d-paths* $\mathcal{P}$ *(notation* $\mathsf{Id}_\mathcal{P}$*) is defined by induction on the d-paths in* $\mathcal{P}$*:*

$$\mathsf{Id}_{\{\}} = \mathsf{Id}_{\{\epsilon\}} = \lambda x.x \qquad \mathsf{Id}_\mathcal{P} \; \beta \longleftarrow \lambda xy.\mathsf{Id}_{\mathcal{R}(\mathcal{P})}(x(\mathsf{Id}_{\mathcal{L}(\mathcal{P})}y)) \; \textit{if } \mathcal{P} \neq \{\}, \{\epsilon\}$$

where $\mathcal{L}(\mathcal{P}) = \{\mathsf{p} \mid \nearrow \mathsf{p} \in \mathcal{P}\}$ *and* $\mathcal{R}(\mathcal{P}) = \{\mathsf{p} \mid \searrow \mathsf{p} \in \mathcal{P}\}$*.*

3. The FHI *induced by the d-path* p *(notation* $\mathsf{Id_p}$*) is* $\mathsf{Id_p} = \mathsf{Id}_{\{\mathsf{p}\}}$*.*

For example $\mathsf{Id}_{\searrow\nearrow\square \; \beta} \longleftarrow \lambda x_1 y_1.\mathsf{Id}_{\nearrow\square}(x_1 y_1) \; \beta \longleftarrow \lambda x_1 y_1.(\lambda x_2 y_2.x_2(\mathsf{Id_\square}y_2))(x_1 y_1) \; \beta \longleftarrow$
$\lambda x_1 y_1.(\lambda x_2 y_2.x_2((\lambda x_3 y_3.x_3 y_3)y_2))(x_1 y_1)$, so $\mathsf{Id}_{\searrow\nearrow\square} = \lambda x_1 y_1 y_2.x_1 y_1(\lambda y_3.y_2 y_3)$.

The following lemma shows that the FHI associated with a d-path, an s-path or a set of d-paths maps to itself each type that agrees with it.

Lemma 4.7.   *1. Let* p *be a d-path or an s-path, then* $\sigma \propto \mathsf{p}$ *implies* $\vdash \mathsf{Id}_\mathsf{p}:\sigma \to \sigma$.

2. *Let* $\mathcal{P}$ *be a set of d-paths, then* $\sigma \propto \mathcal{P}$ *implies* $\vdash \mathsf{Id}_\mathcal{P}:\sigma \to \sigma$.

Proof. Only Point (2) is proved, being the proof of Point (1) similar and simpler. The proof is by induction on σ and $\mathcal{P}$. If $\sigma = \tau \to \rho$, then by definition $\tau \propto \mathcal{L}(\mathcal{P})$ and $\rho \propto \mathcal{R}(\mathcal{P})$. By induction $\vdash \mathsf{Id}_{\mathcal{L}(\mathcal{P})}:\tau \to \tau$ and $\vdash \mathsf{Id}_{\mathcal{R}(\mathcal{P})}:\rho \to \rho$, which imply $\vdash \lambda xy.\mathsf{Id}_{\mathcal{R}(\mathcal{P})}(x(\mathsf{Id}_{\mathcal{L}(\mathcal{P})}y)):\sigma \to \sigma$. If $\sigma = \tau \vee \rho$ or $\sigma = \tau \wedge \rho$, then by definition $\tau \propto \mathcal{P}$ and $\rho \propto \mathcal{P}$. These cases easily follow by induction using Corollary 2.2(3). $\qquad\square$

To prove the soundness of erasure one needs to show that the FHI associated with a set of d-paths "respects" the preorder relation, in the sense that, if the set of d-paths of the derivation $\sigma \leq^\diamond \tau$ is contained in a set $\mathcal{P}$ and either σ or τ agrees with $\mathcal{P}$, then the FHI $\mathsf{Id}_\mathcal{P}$ maps σ to τ.

Lemma 4.8. *If* $e(\sigma \leq^\diamond \tau) \subseteq \mathcal{P}$ *and either* $\sigma \propto \mathcal{P}$ *or* $\tau \propto \mathcal{P}$, *then* $\vdash \mathsf{Id}_\mathcal{P}:\sigma \to \tau$.

Proof. By induction on the proof of $\sigma \leq^\diamond \tau$. The cases $\mu \leq^\wedge \mu$ and $\chi \leq^\vee \chi$ follow immediately by Lemma 4.7. In cases $\varphi \wedge \mu \leq^\wedge \varphi, \varphi \leq^\vee \varphi \vee \chi, \varphi \wedge \mu \wedge \lambda \leq^\wedge \varphi \wedge \mu$ and $\varphi \vee \chi \leq^\vee \varphi \vee \chi \vee \iota$ one has $e(\sigma \leq^\diamond \tau) = \{\epsilon\}$. Since $\sigma \propto \mathcal{P}$ or $\tau \propto \mathcal{P}$ implies $\varphi \propto \mathcal{P}$, $\mathcal{P} = \{\epsilon\}$ and by Definition 4.6(2), one gets $\mathsf{Id}_\mathcal{P} = \lambda x.x$.
Consider the case: $\nu_i \leq^\wedge \mu_i, \chi_i \leq^\vee \kappa_i$ for all $i \in I \implies \bigwedge_{i \in I}(\mu_i \to \chi_i)[\wedge \lambda] \leq^\wedge \bigwedge_{i \in I}(\nu_i \to \kappa_i)$. By definition:

- $e(\bigwedge_{i \in I}(\mu_i \to \chi_i)[\wedge \lambda] \leq^\wedge \bigwedge_{i \in I}(\nu_i \to \kappa_i)) \subseteq \mathcal{P}$ implies $e(\nu_i \leq^\wedge \mu_i) \subseteq \mathcal{L}(\mathcal{P})$ and $e(\chi_i \leq^\vee \kappa_i) \subseteq \mathcal{R}(\mathcal{P})$ for all $i \in I$;

- either $\bigwedge_{i \in I}(\mu_i \to \chi_i)[\wedge \lambda] \propto \mathcal{P}$ or $\bigwedge_{i \in I}(\nu_i \to \kappa_i) \propto \mathcal{P}$ implies either $\nu_i \propto \mathcal{L}(\mathcal{P})$ or $\mu_i \propto \mathcal{L}(\mathcal{P})$ and either $\chi_i \propto \mathcal{R}(\mathcal{P})$ or $\kappa_i \propto \mathcal{R}(\mathcal{P})$ for all $i \in I$.

This gives by induction $\vdash \mathsf{Id}_{\mathcal{L}(\mathcal{P})}:\nu_i \to \mu_i$ and $\vdash \mathsf{Id}_{\mathcal{R}(\mathcal{P})}:\chi_i \to \kappa_i$ for all $i \in I$. By definition
$\mathsf{Id}_\mathcal{P} \;_\beta\!\longleftarrow \lambda xy.\mathsf{Id}_{\mathcal{R}(\mathcal{P})}(x(\mathsf{Id}_{\mathcal{L}(\mathcal{P})}y))$. It can be easily shown that:
$$\vdash \lambda xy.\mathsf{Id}_{\mathcal{R}(\mathcal{P})}(x(\mathsf{Id}_{\mathcal{L}(\mathcal{P})}y)):(\mu_i \to \chi_i) \to \nu_i \to \kappa_i \text{ for all } i \in I.$$
The Subject Reduction (Theorem 2.1) implies $\vdash \mathsf{Id}_\mathcal{P}:(\mu_i \to \chi_i) \to \nu_i \to \kappa_i$ for all $i \in I$, and so by Corollary 2.2(3) and (4)
$$\vdash \mathsf{Id}_\mathcal{P}:\bigwedge_{i \in I}(\mu_i \to \chi_i)[\wedge \lambda] \to \bigwedge_{i \in I}(\nu_i \to \kappa_i).$$
For the case: $\nu_i \leq^\wedge \mu_i, \chi_i \leq^\vee \kappa_i$ for all $i \in I \implies \bigvee_{i \in I}(\mu_i \to \chi_i) \leq^\vee \bigvee_{i \in I}(\nu_i \to \kappa_i)[\vee \iota]$ a similar argument gives $\vdash \mathsf{Id}_\mathcal{P}:\bigvee_{i \in I}(\mu_i \to \chi_i) \to \bigvee_{i \in I}(\nu_i \to \kappa_i)[\vee \iota]$. $\qquad\square$

The soundness of the normalisation rules can now be proved.

Theorem 4.9.   *1. If* $d(C[\ ])$ *is defined, then for arbitrary* $\sigma, \tau,$ *and* $\rho,$ *the* FHI $\mathsf{Id}_{d(C[\ ])}$ *proves the isomorphisms:* $C[(\sigma \vee \tau) \wedge \rho] \approx C[(\sigma \wedge \rho) \vee (\tau \wedge \rho)]$ *and* $C[(\sigma \wedge \tau) \vee \rho] \approx C[(\sigma \vee \rho) \wedge (\tau \vee \rho)]$.

2. *If* $s(C[\ ])$ *is defined, then for arbitrary* $\sigma, \tau,$ *and* $\rho,$ *the* FHI $\mathsf{Id}_{s(C[\ ])}$ *proves the isomorphisms:* $C[\sigma \to \tau \wedge \rho] \approx C[(\sigma \to \tau) \wedge (\sigma \to \rho)]$ *and* $C[\sigma \vee \tau \to \rho] \approx C[(\sigma \to \rho) \wedge (\tau \to \rho)]$.

3. *Let* $\bigwedge_{i \in I}\chi_i \implies \bigwedge_{j \in J}\chi_j,$ *i.e.* $J \subset I$ *and* $\forall i \in I \exists j_i \in J. \chi_{j_i} \leq^\vee \chi_i$ *and* $\forall j \in J. \chi_j \propto \mathcal{P},$ *where* $\mathcal{P} = \bigcup_{i \in I}e(\chi_{j_i} \leq^\vee \chi_i)$. *Then* $\mathsf{Id}_\mathcal{P}$ *proves* $\bigwedge_{j \in J}\chi_j \approx \bigwedge_{i \in I}\chi_i$.

4. *Let* $C[\bigwedge_{i \in I}\alpha_i] \implies C[\bigwedge_{j \in J}\alpha_j],$ *i.e.* $J \subset I$ *and* $\forall i \in I \exists j_i \in J. \alpha_{j_i} \leq^\diamond \alpha_i$ *and* $\forall j \in J. C[\alpha_j] \propto \mathcal{P},$ *where* $\mathcal{P} = d(C[]) \cdot \bigcup_{i \in I}e(\alpha_{j_i} \leq^\diamond \alpha_i)$. *Then* $\mathsf{Id}_\mathcal{P}$ *proves* $C[\bigwedge_{j \in J}\alpha_j] \approx C[\bigwedge_{i \in I}\alpha_i]$.

5. *Let* $C[\bigvee_{i \in I}\alpha_i] \implies C[\bigvee_{j \in J}\alpha_j]$ *i.e.* $J \subset I$ *and* $\forall i \in I. \exists j_i \in J. \alpha_i \leq^\diamond \alpha_{j_i}$ *and* $\forall j \in J. C[\alpha_j] \propto \mathcal{P},$ *where* $\mathcal{P} = d(C[]) \cdot \bigcup_{i \in I}e(\alpha_i \leq^\diamond \alpha_{j_i})$. *Then* $\mathsf{Id}_\mathcal{P}$ *proves* $C[\bigvee_{i \in I}\alpha_i] \approx C[\bigvee_{j \in J}\alpha_j]$.

Proof. (1). By induction on $C[\ ]$. If $C[\ ] = [\ ]$ by Definition 4.2 $d([\ ]) = \epsilon$ and $\mathsf{Id}_\epsilon = \lambda x.x$.

If $C[\ ] = C'[\ ] \to \theta$, then by induction

$$\vdash \mathsf{Id}_{d(C'[\ ])} : C'[(\sigma \vee \tau) \wedge \rho] \to C'[(\sigma \wedge \rho) \vee (\tau \wedge \rho)]$$
$$\vdash \mathsf{Id}_{d(C'[\ ])} : C'[(\sigma \wedge \rho) \vee (\tau \wedge \rho)] \to C'[(\sigma \vee \tau) \wedge \rho].$$

Since by definition $\mathsf{Id}_{d(C[\ ])} \overset{\beta}{\longleftarrow} \lambda xy.x(\mathsf{Id}_{d(C'[\ ])}y)$ the result follows.

If $C[\ ] = \theta \to C'[\ ]$ the proof is similar to that one of previous case.

If $C[\ ] = C'[\ ] \wedge \theta$, then by induction

$$\vdash \mathsf{Id}_{d(C'[\ ])} : C'[(\sigma \vee \tau) \wedge \rho] \to C'[(\sigma \wedge \rho) \vee (\tau \wedge \rho)]$$
$$\vdash \mathsf{Id}_{d(C'[\ ])} : C'[(\sigma \wedge \rho) \vee (\tau \wedge \rho)] \to C'[(\sigma \vee \tau) \wedge \rho].$$

Moreover $d(C[\ ]) = d(C'[\ ])$ and $\theta \propto d(C'[\ ])$, so from Lemma 4.7(1) $\vdash \mathsf{Id}_{d(C'[\ ])} : \theta \to \theta$ and by Corollary 2.2(3) the proof is done.

If $C[\ ] = C'[\ ] \vee \theta$, the proof is similar.

(2). Similar to the proof of Point (1). The only difference is case $C[\ ] = [\ ]$, in which by Definition 4.2 $s([\ ]) = \square$ and $\mathsf{Id}_\square = \lambda xy.xy$.

(3). Lemma 4.8 implies $\vdash \mathsf{Id}_{\mathcal{P}} : \chi_{j_i} \to \chi_i$, since $\chi_{j_i} \propto \mathcal{P}$ for all $i \in I$, and $\mathcal{P} = \bigcup_{i \in I} e(\chi_{j_i} \leq^\vee \chi_i)$. Lemma 4.7(2) gives $\vdash \mathsf{Id}_{\mathcal{P}} : \chi_j \to \chi_j$ for all $j \in J$, since $\chi_j \propto \mathcal{P}$ for all $j \in J$. So, by Corollary 2.2(3), $\mathsf{Id}_{\mathcal{P}}$ has both the types $\bigwedge_{i \in I} \chi_{j_i} \to \bigwedge_{i \in I} \chi_i$ and $\bigwedge_{j \in J} \chi_j \to \bigwedge_{j \in J} \chi_j$. Finally, Corollary 2.2(4) implies $\vdash \mathsf{Id}_{\mathcal{P}} : \bigwedge_{j \in J} \chi_j \to \bigwedge_{i \in I} \chi_i$.

(4). By induction on $C[\ ]$. If $C[\ ] = [\ ]$, the proof is immediate from Point (3).

Let $\mathcal{P}' = d(C'[]) \cdot \bigcup_{i \in I} e(\alpha_{j_i} \leq^\diamond \alpha_i)$.

If $C[\ ] = C'[\ ] \to \sigma$, then $d(C[\ ]) = \diagup \cdot d(C'[\ ])$. By induction

$$\vdash \mathsf{Id}_{\mathcal{P}'} : C'[\bigwedge_{j \in J} \alpha_j] \to C'[\bigwedge_{i \in I} \alpha_i] \quad \text{and} \quad \vdash \mathsf{Id}_{\mathcal{P}'} : C'[\bigwedge_{i \in I} \alpha_i] \to C'[\bigwedge_{j \in J} \alpha_j] \qquad .$$

Since by definition $\mathsf{Id}_{\mathcal{P}} \overset{\beta}{\longleftarrow} \lambda xy.x(\mathsf{Id}_{\mathcal{P}'}y)$, the result follows.

If $C[\ ] = \sigma \to C'[\ ]$ the proof is similar to that one of previous case.

If $C[\ ] = C'[\ ] \wedge \sigma$, then by induction

$$\vdash \mathsf{Id}_{\mathcal{P}'} : C'[\bigwedge_{j \in J} \alpha_j] \to C'[\bigwedge_{i \in I} \alpha_i] \quad \text{and} \quad \vdash \mathsf{Id}_{\mathcal{P}'} : C'[\bigwedge_{i \in I} \alpha_i] \to C'[\bigwedge_{j \in J} \alpha_j] \qquad .$$

In this case $\mathcal{P} = \mathcal{P}'$ and $\sigma \propto \mathcal{P}$. Lemma 4.7(2) gives $\vdash \mathsf{Id}_{\mathcal{P}} : \sigma \to \sigma$ and Corollary 2.2(3) concludes the proof.

If $C[\ ] = C'[\ ] \vee \sigma$, the proof is similar.

(5). Similar to the proof of Point (4).   □

This subsection ends with the proof of the existence and unicity of normal forms, i.e. that the normalisation rules are terminating and confluent.

Theorem 4.10 (Normal Forms). *The rewriting system of Definitions 4.4 and 4.5 is terminating and confluent.*

Proof. The *termination* follows from an easy adaptation of the recursive path ordering method [5]. The partial order on operators is defined by: $\to\ >\ \vee\ >\ \wedge$ for holes at top level or in the right-hand-sites of arrow types and $\to\ >\ \wedge\ >\ \vee$ for holes in the left-hand-sites of arrow types. Notice that the induced recursive path ordering $>^*$ has the subterm property. This solves the case of erasure rules. For the first distributive rule, since $\vee\ >\ \wedge$ for holes at top level or in the right-hand-sites of arrow types, it is enough to observe that $(\sigma \wedge \tau) \vee \rho >^* \sigma \vee \rho$ and $(\sigma \wedge \tau) \vee \rho >^* \tau \vee \rho$. For the first splitting rule, since $\to\ >\ \wedge$, it is enough to observe that $\sigma \to \tau \wedge \rho >^* \sigma \to \tau$ and $\sigma \to \tau \wedge \rho >^* \sigma \to \rho$. The proof for the remaining rules are similar.

For *confluence*, following the Knuth-Bendix algorithm [10] it is sufficient to prove the convergence of the critical pairs, that are generated (modulo commutativity and associativity of union and intersection) by:

$$(\bigwedge_{i \in I} \alpha_i) \vee \sigma, \qquad \sigma \to (\bigwedge_{i \in I} \alpha_i) \vee \tau, \qquad (\bigvee_{i \in I} \alpha_i) \wedge \sigma \to \tau,$$

$$\sigma \to (\bigvee_{i\in I}\alpha_i)\wedge\tau, \qquad \sigma\vee(\bigwedge_{i\in I}\alpha_i)\to\tau, \;\; \sigma\to\bigwedge_{i\in I}\alpha_i\wedge\tau, \qquad \sigma\vee\bigvee_{i\in I}\alpha_i\to\tau,$$
$$(\sigma\wedge\tau)\vee\rho\vee(\theta\wedge\vartheta), \qquad (\sigma\vee\tau)\wedge\rho\wedge(\theta\vee\vartheta)\to\zeta,$$
$$\sigma\vee\tau\vee\rho\to\theta, \qquad \sigma\to\tau\wedge\rho\wedge\theta,$$
$$\bigwedge_{i\in I}\chi_i, \qquad \bigwedge_{i\in I}\alpha_i, \qquad \bigvee_{i\in I}\alpha_i,$$

when $\bigwedge_{j\in J}\alpha_j \leq^\wedge \bigwedge_{i\in I}\alpha_i$, $\bigvee_{i\in I}\alpha_i \leq^\vee \bigvee_{j\in J}\alpha_j$, $J\subset I$ and all the conditions required by the normalisation rules are satisfied. The types of the last line generate critical pairs when, for some $L\subset I$, $L\neq J$, one has, for the first type $\forall i\in I \,\exists j_i\in J \,l_i\in L$ such that $\chi_{j_i}\leq^\vee\chi_i$ and $\chi_{l_i}\leq^\vee\chi_i$ and, for the second and the third types, in addition to the above conditions, $\bigwedge_{l\in L}\alpha_j\leq^\wedge\bigwedge_{i\in I}\alpha_i$ and $\bigvee_{l\in L}\alpha_i\leq^\vee\bigvee_{j\in J}\alpha_j$, respectively.

The types in the first line can be reduced by distribution and erasure rules, the types in the second line can be reduced by splitting and erasure rules, the types in the third line can be reduced by distribution rules, the types in the forth line can be reduced by splitting rules, the types in the last line can be reduced by erasure rules. Notice that distribution and splitting rules do not generate critical pairs, since they require respectively unions and intersections for holes in the right-hand-sites of arrow types and intersections and unions for holes in the left-hand-sites of arrow types.

The proof is given only for a top level occurrence of the type $(\bigwedge_{i\in I}\alpha_i)\vee\sigma$, the proof for the other cases being similar. In this case $(\bigwedge_{i\in I}\alpha_i)\vee\sigma \Longrightarrow (\bigwedge_{j\in J}\alpha_j)\vee\sigma$ by erasure since $J\subset I$ and $\forall i\in I \,\exists j_i\in J$ such that $\alpha_{j_i}\leq^\diamond\alpha_i$ and $(\bigwedge_{i\in I}\alpha_i)\vee\sigma \propto \bigcup_{i\in I}e(\alpha_{j_i}\leq^\diamond\alpha_i)$. Moreover
$$(\bigwedge_{i\in I}\alpha_i)\vee\sigma \Longrightarrow (\bigwedge_{i\in I_1}\alpha_i\vee\sigma)\wedge(\bigwedge_{i\in I_2}\alpha_i\vee\sigma)$$
with $I_1\cup I_2 = I$ and $I_1\cap I_2 = \emptyset$ by distribution. Let $\bigwedge_{l\in L}\chi_l$ be the conjunctive normal form of σ, then $\sigma \Longrightarrow^* \bigwedge_{l\in L}\chi_l$. This implies $(\bigwedge_{j\in J}\alpha_j)\vee\sigma \Longrightarrow^* \bigwedge_{l\in L}\bigwedge_{j\in J}(\alpha_j\vee\chi_l)$ and
$$(\bigwedge_{i\in I_1}\alpha_i\vee\sigma)\wedge(\bigwedge_{i\in I_2}\alpha_i\vee\sigma) \Longrightarrow^* \bigwedge_{l\in L}(\bigwedge_{i\in I_1}\alpha_i\vee\chi_l)\wedge(\bigwedge_{i\in I_2}\alpha_i\vee\chi_l) = \bigwedge_{l\in L}\bigwedge_{i\in I}(\alpha_i\vee\chi_l)$$
by the first distribution rule. Since $\alpha_{j_i}\leq^\diamond\alpha_i$ implies $\alpha_{j_i}\vee\chi_l\leq^\vee\alpha_i\vee\chi_l$, and $(\bigwedge_{j\in J}\alpha_j)\vee\sigma \propto \bigcup_{i\in I}e(\alpha_{j_i}\leq^\diamond\alpha_i)$ implies $\bigwedge_{l\in L}\bigwedge_{j\in J}(\alpha_j\vee\chi_l)\propto\bigcup_{l\in L}\bigcup_{i\in I}e(\alpha_{j_i}\vee\chi_l\leq^\vee\alpha_i\vee\chi_l)$, the first erasure rule gives
$$\bigwedge_{l\in L}\bigwedge_{i\in I}(\alpha_i\vee\chi_l) \Longrightarrow \bigwedge_{l\in L}\bigwedge_{j\in J}(\alpha_j\vee\chi_l). \qquad \square$$

4.3 Properties of normal types

It is interesting to show that normal types do not contain "superfluous" subtypes, in particular that:

1. if $\chi\wedge\kappa$ is a normal type, then there is no Id such that $\vdash \mathsf{Id}:\chi\to\kappa$;

2. if $\alpha\vee\beta$ is a normal type, then there is no Id such that $\vdash \mathsf{Id}:\alpha\to\beta$;

Theorem 4.14 shows Points (1) and (2).

The more interesting result is Theorem 4.15, which assures that isomorphic normal types have the same number of intersections and unions and that the atomic and arrow types are pairwise isomorphic.

It is easy to verify that each FHI Id different from the identity is such that $\mathsf{Id} \mathop{{}_\beta\!\longleftarrow} \lambda xy.\mathsf{Id}_1(x(\mathsf{Id}_2 y))$ for unique $\mathsf{Id}_1, \mathsf{Id}_2$. A key result is a relation between the arrow types that can be derived for $\mathsf{Id},\mathsf{Id}_1,\mathsf{Id}_2$. Lemma 4.11(4) gives this relation, by exploiting the constraints on typings of variables and FHIS, shown in the first three points of the same lemma.

Building on Definition 4.6(2) a non-empty set of d-paths is associated with each FHI (Definition 4.12). This association is based on the natural correspondence between lambda abstractions and arrow types.

A last Lemma (Lemma 4.13) relates basic intersections and unions in normal form (when they can be mapped by FHIS) to preorders and to sets of d-paths.

Lemma 4.11. *1. If $\Gamma, x:\sigma\to\tau, y:\rho \vdash x(My):\theta$, then $\Gamma \vdash \lambda y.My:\rho\to\sigma$.*

2. *Let M be either a FHI or a free variable. Then $\Gamma, x:\sigma\to\tau, y:\rho \vdash M(xy):\theta$ implies $\Gamma \vdash \lambda z.Mz:\tau\to\theta$ and $z:\rho \vdash z:\sigma$.*

3. *Let M_1, M_2 be either* FHI*s or free variables and $FV(M_i)$ be the set of variables in Γ_i for $i = 1, 2$. Then $\Gamma_1, \Gamma_2, x:\sigma \to \tau, y:\rho \vdash M_1(x(M_2y)):\theta$ implies $\Gamma_1 \vdash \lambda z.M_1z:\tau \to \theta$ and $\Gamma_2 \vdash \lambda z.M_2z:\rho \to \sigma$.*

4. *If $\vdash \mathsf{Id}:(\mu \to \chi) \to \nu \to \kappa$ and $\mathsf{Id} \; {}_\beta\!\!\longleftarrow \lambda xy.\mathsf{Id}_1(x(\mathsf{Id}_2y))$, then $\vdash \mathsf{Id}_1:\chi \to \kappa$ and $\vdash \mathsf{Id}_2:\nu \to \mu$.*

Proof. (1). Lemma 2.7(1) implies $\Gamma, y:\rho \vdash My:\sigma$, and rule $(\to I)$ derives $\Gamma \vdash \lambda y.My:\rho \to \sigma$.

(2). The proof is similar and simpler than that of (3).

(3). A stronger statement, i.e.

$$x:\sigma \to \tau \vdash x:\varsigma \text{ and } \Gamma_1, \Gamma_2, x:\varsigma, y:\rho \vdash M_1(x(M_2y)):\theta \text{ imply } \Gamma_1 \vdash \lambda z.M_1z:\tau \to \theta \text{ and } \Gamma_2 \vdash \lambda z.M_2z:\rho \to \sigma,$$

is proved by induction on the derivation of $\Gamma_1, \Gamma_2, x:\varsigma, y:\rho \vdash M_1(x(M_2y)):\theta$.

Let the last applied rule be $(\to E)$:

$$(\to E) \quad \frac{\Gamma_1 \vdash M_1:\vartheta \to \theta \quad \Gamma_2, x:\varsigma, y:\rho \vdash x(M_2y):\vartheta}{\Gamma_1, \Gamma_2, x:\varsigma, y:\rho \vdash M_1(x(M_2y)):\theta}$$

The second premise and $x:\sigma \to \tau \vdash x:\varsigma$ imply $\Gamma_2, x:\sigma \to \tau, y:\rho \vdash x(M_2y):\vartheta$ by rule (L). Point (1) gives $\Gamma_2 \vdash \lambda y.M_2y:\rho \to \sigma$ and Lemma 2.7(2) gives $z:\tau \vdash z:\vartheta$. The application of $(\to E)$ to the first premise and to $z:\tau \vdash z:\vartheta$ derives $\Gamma_1, z:\tau \vdash M_1z:\theta$, and then $\Gamma_1 \vdash \lambda z.M_1z:\tau \to \theta$ by using $(\to I)$.

If the last applied rule is $(\wedge I)$, $(\wedge E)$, or $(\vee I)$ the proof easily follows by induction.

For rule $(\vee E)$ there are seven cases, which differ for the subjects of the premises. I.e. if t is the replaced variable the subjects of the first two premises can be: $t(x(M_2y))$, $M_1(t(M_2y))$, $M_1(x(ty))$, $M_1(x(M_2t))$, M_1t, $M_1(xt)$ and t. The proof is given for all the cases but the last one, which easily follows by induction. Notice that the proof of the sixth case needs Point (2).

In the first case:

$$(\vee E) \quad \frac{\begin{array}{cc}\Gamma_2, x:\varsigma, y:\rho, t:\vartheta_1 \wedge \zeta \vdash t(x(M_2y)):\theta \quad \Gamma_2, x:\varsigma, y:\rho, t:\vartheta_2 \wedge \zeta \vdash t(x(M_2y)):\theta \\ \Gamma_1 \vdash M_1:(\vartheta_1 \vee \vartheta_2) \wedge \zeta\end{array}}{\Gamma_1, \Gamma_2, x:\varsigma, y:\rho \vdash M_1(x(M_2y)):\theta}$$

By induction $t:\vartheta_1 \wedge \zeta \vdash \lambda z.tz:\tau \to \theta$ and $t:\vartheta_2 \wedge \zeta \vdash \lambda z.tz:\tau \to \theta$ and $\Gamma_2 \vdash \lambda y.M_2y:\rho \to \sigma$. By rule $(\vee L)$ $t:(\vartheta_1 \vee \vartheta_2) \wedge \zeta \vdash \lambda z.tz:\tau \to \theta$, so the application of rule (C) to the third premise derives $\Gamma_1 \vdash \lambda z.M_1z:\tau \to \theta$.

In the second case:

$$(\vee E) \quad \frac{\begin{array}{cc}\Gamma_1, \Gamma_2, y:\rho, t:\vartheta_1 \wedge \zeta \vdash M_1(t(M_2y)):\theta \quad \Gamma_1, \Gamma_2, y:\rho, t:\vartheta_2 \wedge \zeta \vdash M_1(t(M_2y)):\theta \\ x:\varsigma \vdash x:(\vartheta_1 \vee \vartheta_2) \wedge \zeta\end{array}}{\Gamma_1, \Gamma_2, x:\varsigma, y:\rho \vdash M_1(x(M_2y)):\theta}$$

Rule (L) applied to $x:\sigma \to \tau \vdash x:\varsigma$ and to the third premise derives $x:\sigma \to \tau \vdash x:(\vartheta_1 \vee \vartheta_2) \wedge \zeta$. Corollary 2.6(2) gives either $x:\sigma \to \tau \vdash x:\vartheta_1 \wedge \zeta$ or $x:\sigma \to \tau \vdash x:\vartheta_2 \wedge \zeta$. This implies either $t:\sigma \to \tau \vdash t:\vartheta_1 \wedge \zeta$ or $t:\sigma \to \tau \vdash t:\vartheta_2 \wedge \zeta$. By induction on the first premise in the first case and on the second premise in the second case $\Gamma_1 \vdash \lambda z.M_1z:\tau \to \theta$ and $\Gamma_2 \vdash \lambda z.M_2z:\rho \to \sigma$.

In the third case:

$$(\vee E) \quad \frac{\begin{array}{cc}\Gamma_1, x:\varsigma, y:\rho, t:\vartheta_1 \wedge \zeta \vdash M_1(x(ty)):\theta \quad \Gamma_1, x:\varsigma, y:\rho, t:\vartheta_2 \wedge \zeta \vdash M_1(x(ty)):\theta \\ \Gamma_2 \vdash M_2:(\vartheta_1 \vee \vartheta_2) \wedge \zeta\end{array}}{\Gamma_1, \Gamma_2, x:\varsigma, y:\rho \vdash M_1(x(M_2y)):\theta}$$

By induction $\Gamma_1 \vdash \lambda z.M_1z:\tau \to \theta$ and $t:\vartheta_1 \wedge \zeta \vdash \lambda z.tz:\rho \to \sigma$ and $t:\vartheta_2 \wedge \zeta \vdash \lambda z.tz:\rho \to \sigma$. Rule $(\vee L)$ derives $t:(\vartheta_1 \vee \vartheta_2) \wedge \zeta \vdash \lambda z.tz:\rho \to \sigma$, so the application of rule (C) to the third premise gives $\Gamma_2 \vdash \lambda y.M_2y:\rho \to \sigma$.

In the fourth case:

$$(\vee E) \quad \frac{\Gamma_1,\Gamma_2,x{:}\varsigma,t{:}\vartheta_1\wedge\zeta \vdash M_1(x(M_2t)){:}\theta \quad \Gamma_1,\Gamma_2,x{:}\varsigma,t{:}\vartheta_2\wedge\zeta \vdash M_1(x(M_2t)){:}\theta}{\Gamma_1,\Gamma_2,x{:}\varsigma,y{:}\rho \vdash M_1(x(M_2y)){:}\theta}$$
$$y{:}\rho \vdash y{:}(\vartheta_1\vee\vartheta_2)\wedge\zeta$$

By induction on one of the first two premises $\Gamma_1 \vdash \lambda z.M_1z{:}\tau \to \theta$ and $\Gamma_2 \vdash \lambda z.M_2z{:}\rho \to \sigma$.

In the fifth case:

$$(\vee E) \quad \frac{\Gamma_1,t{:}\vartheta_1\wedge\zeta \vdash M_1t{:}\theta \quad \Gamma_1,t{:}\vartheta_2\wedge\zeta \vdash M_1t{:}\theta \quad \Gamma_2,x{:}\varsigma,y{:}\rho \vdash x(M_2y){:}(\vartheta_1\vee\vartheta_2)\wedge\zeta}{\Gamma_1,\Gamma_2,x{:}\varsigma,y{:}\rho \vdash M_1(x(M_2y)){:}\theta}$$

The third premise with $x{:}\sigma \to \tau \vdash x{:}\varsigma$ give $\Gamma_2,x{:}\sigma \to \tau,y{:}\rho \vdash x(M_2y){:}(\vartheta_1\vee\vartheta_2)\wedge\zeta$, so by Point (1) $\Gamma_2 \vdash \lambda z.M_2z{:}\rho \to \sigma$. By Lemma 2.7(2) $\Gamma_2,x{:}\sigma \to \tau,y{:}\rho \vdash x(M_2y){:}(\vartheta_1\vee\vartheta_2)\wedge\zeta$ implies $z{:}\tau \vdash z{:}(\vartheta_1\vee\vartheta_2)\wedge\zeta$. The application of rule $(\vee E)$ to the first two premises and to $z{:}\tau \vdash z{:}(\vartheta_1\vee\vartheta_2)\wedge\zeta$ derives $\Gamma_1,z{:}\tau \vdash M_1z{:}\theta$, which implies $\Gamma_1 \vdash \lambda z.M_1z{:}\tau \to \theta$ by rule $(\to I)$.

In the sixth case:

$$(\vee E) \quad \frac{\Gamma_1,x{:}\varsigma,t{:}\vartheta_1\wedge\zeta \vdash M_1(xt){:}\theta \quad \Gamma_1,x{:}\varsigma,t{:}\vartheta_2\wedge\zeta \vdash M_1(xt){:}\theta}{\Gamma_1,\Gamma_2,x{:}\varsigma,y{:}\rho \vdash M_1(x(M_2y)){:}\theta}$$
$$\Gamma_2,y{:}\rho \vdash M_2y{:}(\vartheta_1\vee\vartheta_2)\wedge\zeta$$

The first and the second premise with $x{:}\sigma \to \tau \vdash x{:}\varsigma$ give $\Gamma_1,x{:}\sigma \to \tau,t{:}\vartheta_1\wedge\zeta \vdash M_1(xt){:}\theta$ and $\Gamma_1,x{:}\sigma \to \tau,t{:}\vartheta_2\wedge\zeta \vdash M_1(xt){:}\theta$. So Point (2) implies $\Gamma_1 \vdash \lambda z.M_1z{:}\tau \to \theta, t{:}\vartheta_1\wedge\zeta \vdash t{:}\sigma$ and $t{:}\vartheta_2\wedge\zeta \vdash t{:}\sigma$. The application of rule $(\vee E)$ to the last two statements and to the third premise derives $\Gamma_2,y{:}\rho \vdash M_2y{:}\sigma$, so rule $(\to I)$ concludes the proof.

(4). The Subject Expansion (Theorem 2.3) gives $\vdash \lambda xy.\mathsf{Id}_1(x(\mathsf{Id}_2y)){:}(\mu \to \chi) \to \nu \to \kappa$. Corollary 2.2(1) implies

$$x{:}\mu \to \chi, y{:}\nu \vdash \mathsf{Id}_1(x(\mathsf{Id}_2y)){:}\kappa.$$

Point (3) and Subject Reduction conclude the proof. □

Definition 4.12 (Set of d-paths of an FHI). *The set of d-paths of the FHI* Id *(notation* $\#(\mathsf{Id})$*) is defined by:*

$$\#(\lambda x.x) = \{\epsilon\} \quad \#(\lambda xy.\mathsf{Id}_1(x(\mathsf{Id}_2y))) = \{\nearrow \mathsf{p} \mid \mathsf{p} \in \#(\mathsf{Id}_2)\} \cup \{\searrow \mathsf{p} \mid \mathsf{p} \in \#(\mathsf{Id}_1)\}.$$

Lemma 4.13. *1. If* $\vdash \mathsf{Id}{:}\sigma \to \tau$*, where* σ, τ *are both basic intersections or basic unions in normal form, then* $\sigma \leq^\diamond \tau$ *and* $\#(\mathsf{Id}) \supseteq e(\sigma \leq^\diamond \tau)$*, with* $\diamond = \wedge$ *if* σ, τ *are intersections and* $\diamond = \vee$ *if* σ, τ *are unions.*

2. If $\vdash \mathsf{Id}{:}\sigma \to \sigma$ *where* σ *is either a basic intersection or a basic union in normal form, then* $\sigma \propto \#(\mathsf{Id})$*.*

Proof. (1). By induction on Id. If $\mathsf{Id} = \lambda x.x$, then $x{:}\sigma \vdash x{:}\tau$ by Corollary 2.2(1). This implies either $\sigma = \tau$ or $\sigma = \mu \wedge \nu$ and $\tau = \mu$ or $\sigma = \chi$ and $\tau = \chi \vee \kappa$ by Lemma 2.5. By definition either $e(\sigma \leq^\diamond \tau) = \{\,\}$ or $e(\sigma \leq^\diamond \tau) = \{\epsilon\}$. If $\mathsf{Id} \neq \lambda x.x$, the following stronger statement is proved:

If $\vdash \mathsf{Id}{:}\sigma_l \to \tau_l$*, where* σ_l, τ_l *are normal types and either intersections of arrows or unions of arrows for all* $l \in L$*, then* $\sigma_l \leq^\diamond \tau_l$ *and* $\#(\mathsf{Id}) \supseteq \bigcup_{l\in L} e(\sigma_l \leq^\diamond \tau_l)$ *for all* $l \in L$*, with* $\diamond = \wedge$ *if* σ_l, τ_l *are intersections and* $\diamond = \vee$ *if* σ_l, τ_l *are unions.*

If $\mathsf{Id} \;_\beta\!\!\longleftarrow \lambda xy.\mathsf{Id}_1(x(\mathsf{Id}_2y))$, let $\sigma_l = \bigwedge_{i\in I_l}(\mu_i^{(l)} \to \chi_i^{(l)})$, $\tau_l = \bigwedge_{j\in J_l}(\nu_j^{(l)} \to \kappa_j^{(l)})$ (the proof for the case the types are basic unions is similar). Theorem 3.4 and Corollary 2.2(1) imply for all $l \in L$ and $j \in J_l$ there is $i_j \in I_l$ such that $x{:}\mu_{i_j}^{(l)} \to \chi_{i_j}^{(l)}, y{:}\nu_j^{(l)} \vdash \mathsf{Id}_1(x(\mathsf{Id}_2y)){:}\kappa_j^{(l)}$. Lemma 4.11(4) implies $\vdash \mathsf{Id}_1{:}\chi_{i_j}^{(l)} \to \kappa_j^{(l)}$ and

$\vdash \mathsf{Id}_2 : \nu_j^{(l)} \to \mu_{i_j}^{(l)}$ for all $l \in L$ and $j \in J_l$. By induction for all $l \in L$ and $j \in J_l$:

$$\chi_{i_j}^{(l)} \leq^\vee \kappa_j^{(l)} \text{ and } \#(\mathsf{Id}_1) \supseteq \bigcup_{l \in L} \bigcup_{j \in J_l} e(\chi_{i_j}^{(l)} \leq^\vee \kappa_j^{(l)})$$
$$\nu_j^{(l)} \leq^\wedge \mu_{i_j}^{(l)} \text{ and } \#(\mathsf{Id}_2) \supseteq \bigcup_{l \in L} \bigcup_{j \in J_l} e(\nu_j^{(l)} \leq^\wedge \mu_{i_j}^{(l)})$$

By definition of e (Figure 4) $\bigcup_{l \in L} e(\sigma_l \leq^\wedge \tau_l) = \bigcup_{l \in L} \bigcup_{j \in J_l} (\swarrow \cdot e(\nu_j^{(l)} \leq^\wedge \mu_{i_j}^{(l)}) \cup \searrow \cdot e(\chi_{i_j}^{(l)} \leq^\vee \kappa_j^{(l)})) \subseteq \#(\mathsf{Id})$.

(2). By induction on Id. If $\mathsf{Id} = \lambda x.x$, then $\#(\mathsf{Id}) = \{\epsilon\}$ and the proof is immediate.

If $\mathsf{Id}\ _\beta\!\longleftarrow \lambda xy.\mathsf{Id}_1(x(\mathsf{Id}_2 y))$ let $\sigma = \bigwedge_{i \in I}(\mu_i \to \chi_i)$ (the proof for the case of the union is similar). Let $I = J \cup H$, with $J \cap H = \emptyset$, and J be the maximum subset of I such that $\vdash \mathsf{Id}:(\mu_j \to \chi_j) \to \mu_j \to \chi_j$ for all $j \in J$.

The proof starts by showing that J can not be empty. Theorem 3.4 assures that for all $i \in I$ there is $j_i \in I$ such that $\vdash \mathsf{Id}:(\mu_{j_i} \to \chi_{j_i}) \to \mu_i \to \chi_i$. If there are $i_1,\ldots,i_n$ such that $\vdash \mathsf{Id}:(\mu_{i_l} \to \chi_{i_l}) \to \mu_{i_{l+1}} \to \chi_{i_{l+1}}$ for $1 \leq l \leq n-1$ and $\vdash \mathsf{Id}:(\mu_{i_n} \to \chi_{i_n}) \to \mu_{i_1} \to \chi_{i_1}$, then also $\vdash \mathsf{Id}:(\mu_{i_l} \to \chi_{i_l}) \to \mu_{i_l} \to \chi_{i_l}$ for $1 \leq l \leq n$, since $\mathsf{Id}\ _\beta\!\longleftarrow \mathsf{Id} \circ \mathsf{Id}$.

From $\vdash \mathsf{Id}:(\mu_j \to \chi_j) \to \mu_j \to \chi_j$ Corollary 2.2(1) and Lemma 4.11(3) give $\mathsf{Id}_2 : \mu_j \to \mu_j$ and $\mathsf{Id}_1 : \chi_j \to \chi_j$. By induction $\mu_j \propto \#(\mathsf{Id}_2)$ and $\chi_j \propto \#(\mathsf{Id}_1)$ for all $j \in J$, which imply $\bigwedge_{j \in J}(\mu_j \to \chi_j) \propto \#(\mathsf{Id})$. Moreover by assumption for all $h \in H$ there is $j_h \in J$ such that $\mathsf{Id}:(\mu_{j_h} \to \chi_{j_h}) \to \mu_h \to \chi_h$. Point (1) implies $\mu_{j_h} \to \chi_{j_h} \leq^\diamond \mu_h \to \chi_h$ and $\#(\mathsf{Id}) \supseteq e(\mu_{j_h} \to \chi_{j_h} \leq^\diamond \mu_h \to \chi_h)$. The second erasure rule gives $\sigma \Longrightarrow \bigwedge_{j \in J}(\mu_j \to \chi_j)$. Therefore σ would not be a normal type. So H must be empty. $\qquad\square$

Theorem 4.14.   *1. If $\vdash \mathsf{Id}: \bigwedge_{j \in J}\chi_j \to \bigwedge_{i \in I}\chi_i$ and $J \subset I$, then $\bigwedge_{i \in I}\chi_i$ is not a normal type.*

    2. If $\vdash \mathsf{Id}: \bigvee_{i \in I}\alpha_i \to \bigvee_{j \in J}\alpha_j$ and $J \subset I$, then $\bigvee_{i \in I}\alpha_i$ is not a normal type.

Proof. (1). Assume ad absurdum that $\bigwedge_{i \in I}\chi_i$ is a normal type. Let $I = K \cup H$, with $K \cap H = \emptyset$, and H be the maximum subset of I such that for all $h \in H$ there is $k_h \in K$ such that $\mathsf{Id}:(\mu_{k_h} \to \chi_{k_h}) \to \mu_h \to \chi_h$. Notice that by construction $H \supseteq I - J$, therefore H can not be empty. By Lemma 4.13(1) $\mu_{k_h} \to \chi_{k_h} \leq^\diamond \mu_h \to \chi_h$ and $\#(\mathsf{Id}) \supseteq e(\mu_{k_h} \to \chi_{k_h} \leq^\diamond \mu_h \to \chi_h)$. Moreover by assumption $\vdash \mathsf{Id}:\chi_k \to \chi_k$ for all $k \in K$. Lemma 4.13(2) implies $\bigwedge_{k \in K}\chi_k \propto \#(\mathsf{Id})$. The first erasure rule gives $\bigwedge_{i \in I}\chi_i \Longrightarrow \bigwedge_{k \in K}\chi_k$, proving that $\bigwedge_{i \in I}\chi_i$ is not a normal type.

(2). Similar to the proof of (1), using the last erasure rule. $\qquad\square$

Theorem 4.15. *Let $\bigwedge_{i \in I}(\bigvee_{h \in H_i}\alpha_h^{(i)}) \approx \bigwedge_{j \in J}(\bigvee_{k \in K_j}\beta_k^{(j)})$ and both types be normal. Then $I = J$, $H_i = K_i$ and $\alpha_h^{(i)} \approx \beta_h^{(i)}$ for all $h \in H_i$ and $i \in I$.*

Proof. Let $< P, P^{-1} >$ prove the isomorphism and let $\chi_i = \bigvee_{h \in H_i}\alpha_h^{(i)}$ and $\kappa_j = \bigvee_{k \in K_j}\beta_k^{(j)}$.

Assume ad absurdum that $I \subset J$. By Theorem 3.4 for all $j \in J$ there is $i_j \in I$ such that $\vdash P:\chi_{i_j} \to \kappa_j$ and for $i_j \in I$ there is $j_{i_j} \in J$ such that $\vdash P^{-1}:\kappa_{j_{i_j}} \to \chi_{i_j}$. This implies $\vdash P \circ P^{-1}:\kappa_{j_{i_j}} \to \kappa_j$ and, for cardinality reasons, there are i, j such that $j_{i_j} \neq j$. This, together with $\vdash P \circ P^{-1}: \bigwedge_{j \in J}\kappa_j \to \bigwedge_{j \in J}\kappa_j$, gives $\vdash P \circ P^{-1}: \bigwedge_{j \in J'}\kappa_j \to \bigwedge_{j \in J}\kappa_j$ for some $J' \subseteq I$. By Theorem 4.14(1) $\bigwedge_{j \in J}\kappa_j$ is not a normal type. Then $I = J$ and $j_{i_j} = j$. Therefore the indexes can be chosen to get $\chi_i \approx \kappa_i$ for all $i \in I$.

Assume ad absurdum that $K_i \subset H_i$. By Theorem 3.2 for all $h \in H_i$ there is $k_h \in K_i$ such that $\vdash P:\alpha_h^{(i)} \to \beta_{k_h}^{(i)}$ and for all $k_h \in K_i$ there is $h_{k_h} \in H_i$ such that $\vdash P^{-1}:\beta_{k_h}^{(i)} \to \alpha_{h_{k_h}}^{(i)}$. This implies $\vdash P^{-1} \circ P:\alpha_h^{(i)} \to \alpha_{h_{k_h}}^{(i)}$ and there are h, k such that $h_{k_h} \neq h$. This fact, together with $\vdash P^{-1} \circ P: \bigvee_{h \in H_i}\alpha_h^{(i)} \to \bigvee_{h \in H_i}\alpha_h^{(i)}$, gives $\vdash P^{-1} \circ P: \bigvee_{h \in H_i}\alpha_h^{(i)} \to \bigvee_{h \in H'}\alpha_h^{(i)}$ for some $H' \subseteq K_i$. By Theorem 4.14(2) $\bigvee_{h \in H_i}\alpha_h^{(i)}$ is not a normal type. Then $H_i = K_i$ and $h_{k_h} = h$. Therefore the indexes can be chosen to get $\alpha_h^{(i)} \approx \beta_h^{(i)}$ for all $h \in H_i$ and $i \in I$. $\quad\square$

5 Conclusion

This paper introduces a system with intersection and union types for linear λ-terms. The system enjoys subject conversion owing to the linearity restriction. The types that can be derived for the λ-terms proving type isomorphism are studied. A main achievement of this paper is the definition of rules to reduce types to normal form, while preserving isomorphism. These rules are the building blocks for characterising type isomorphism by means of a syntactic equivalence relation between types. This characterisation is the content of [4], where all proofs given in the present paper are omitted. The present paper and [4] can be considered as the first and the second part of a unique work.

References

[1] Franco Barbanera, Mariangiola Dezani-Ciancaglini & Ugo de'Liguoro (1995): *Intersection and Union Types: Syntax and Semantics*. Information and Computation 119(2), pp. 202–230, doi:10.1006/inco.1995.1086.

[2] Kim Bruce, Roberto Di Cosmo & Giuseppe Longo (1992): *Provable isomorphisms of types*. Mathematical Structures in Computer Science 2(2), pp. 231–247, doi:10.1017/S0960129500001444.

[3] Kim Bruce & Giuseppe Longo (1985): *Provable Isomorphisms and Domain Equations in Models of Typed Languages*. In Robert Sedgewick, editor: STOC'85, ACM Press, pp. 263 – 272, doi:10.1145/22145.22175.

[4] Mario Coppo, Mariangiola Dezani-Ciancaglini, Ines Margaria & Maddalena Zacchi (2013): *Isomorphism of Intersection and Union Types*. Mathematical Structures in Computer Science To appear.

[5] Nachum Dershowitz (1982): *Orderings for term-rewriting systems*. Theoretical Computer Science 17(3), pp. 279 – 301, doi:10.1016/0304-3975(82)90026-3.

[6] Mariangiola Dezani-Ciancaglini (1976): *Characterization of Normal Forms Possessing an Inverse in the $\lambda\beta\eta$-calculus*. Theoretical Computer Science 2(3), pp. 323–337, doi:10.1016/0304-3975(76)90085-2.

[7] Mariangiola Dezani-Ciancaglini, Roberto Di Cosmo, Elio Giovannetti & Makoto Tatsuta (2010): *On Isomorphisms of Intersection Types*. ACM Transactions on Computational Logic 11(4), pp. 1–22, doi:10.1145/1805950.1805955.

[8] Roberto Di Cosmo (1995): *Second order Isomorphic Types. A proof theoretic study on second order λ-calculus with surjective pairing and terminal object*. Information and Computation 119(2), pp. 176–201, doi:10.1006/inco.1995.1085.

[9] Roberto Di Cosmo (2005): *A Short Survey of Isomorphisms of Types*. Mathematical Structures in Computer Science 15, pp. 825–838, doi:10.1017/S0960129505004871.

[10] Donald Knuth & Peter Bendix (1970): *Simple word problems in universal algebras*. In J. Leech, editor: Computational Problems in Abstract Algebra, Pergamon Press, pp. 263–297.

[11] David MacQueen, Gordon Plotkin & Ravi Sethi (1986): *An Ideal Model for Recursive Polymorphic Types*. Information and Control 71(1-2), pp. 95–130, doi:10.1016/S0019-9958(86)80019-5.

[12] Dag Prawitz (1965): *Natural Deduction*. Almqvist & Wiksell.

[13] Richard Routley & Robert K. Meyer (1972): *The Semantics of Entailment III*. Journal of Philosophical Logic 1(2), pp. 192–208, doi:10.1007/BF00650498.

[14] Sergei Soloviev (1993): *A Complete Axiom System for Isomorphism of Types in Closed Categories*. In A. Voronkov, editor: LPAR'93, Lecture Notes in Computer Science 698, Springer-Verlag, pp. 360–371, doi:10.1007/3-540-56944-8_71.

[15] Sergei V. Soloviev (1983): *The category of finite sets and Cartesian Closed Categories*. Journal of Soviet Mathematics 22(3), pp. 1387–1400, doi:10.1007/BF01084396. English translation of the original paper in Russian published in Zapiski Nauchnykh Seminarov LOMI, v.105, 1981.