§5. Equivalence

Let $\boldsymbol{L}$ be a fixed DSV, $\boldsymbol{K}$ the full subcategory of its kinds.

We have defined what an $\boldsymbol{L}$-structure is; even, what a $\boldsymbol{C}$-valued $\boldsymbol{L}$-structure is, for any $\boldsymbol{C}$ with finite limits. In what follows, we will make the minimal assumption that $\boldsymbol{C}$ is a regular category (which is equivalent to saying that $\mathcal{P}(\boldsymbol{C})$, with "total" $\mathcal{Q}$, is a $\wedge\exists$-fibration: just ignore $\mathbf{f}$ and $\vee$ in the definition of $\wedge\vee\exists$-fibration).

The category of $\boldsymbol{C}$-valued $\boldsymbol{L}$-structures, $\mathrm{Str}_{\boldsymbol{C}}(\boldsymbol{L})$, has objects the $\boldsymbol{C}$-valued $\boldsymbol{L}$-structures, and morphisms natural transformations; $\mathrm{Str}_{\boldsymbol{C}}(\boldsymbol{L})$ is a full subcategory of $\boldsymbol{C}^{\boldsymbol{L}}$ (with $\boldsymbol{L}$ in its last occurrence understood as a mere category). We write $\mathrm{Str}(\boldsymbol{L})$ for $\mathrm{Str}_{\mathrm{Set}}(\boldsymbol{L})$.

Given $M\in\mathrm{Str}_{\boldsymbol{C}}(\boldsymbol{L})$, we have $M\restriction\boldsymbol{K}:\boldsymbol{K}\longrightarrow\boldsymbol{C}$, its $\boldsymbol{K}$-reduct, the structure of kinds associated to M. For any $R\in\mathrm{Rel}(\boldsymbol{L})$, we have the canonical monomorphism $\mathrm{m}_R:M(R)\rightarrowtail M[R] = (M\restriction\boldsymbol{K})[R]$ (see §1). For a natural transformation $(f:U\longrightarrow V)\in\boldsymbol{C}^{\boldsymbol{K}}$, we have the canonical arrow $f_{[R]}:U[R]\longrightarrow V[R]$ for which

$$\begin{array}{ccc} U[R] & \xrightarrow{\quad f_{[R]}\quad} & V[R] \\ {\scriptstyle \pi^U_p}\downarrow & \circ & \downarrow{\scriptstyle \pi^V_p} \\ U(\mathrm{K}_p) & \xrightarrow[\quad h_{\mathrm{K}_p}\quad]{} & V(\mathrm{K}_p) \end{array}$$

for all $p\in R|\boldsymbol{L}$. If $(h:M\longrightarrow N)\in\mathrm{Str}(\boldsymbol{L})$, then

$$\begin{array}{ccc} M(R) & \xrightarrow{\quad h_R\quad} & N(R) \\ {\scriptstyle \mathrm{m}^M_R}\downarrow & \circ & \downarrow{\scriptstyle \mathrm{m}^N_R} \\ (M\restriction\boldsymbol{K})[R] & \xrightarrow[\quad h_{[R]}\quad]{} & (N\restriction\boldsymbol{K})[R] \end{array}$$

which shows that $h\restriction\boldsymbol{K}:M\restriction\boldsymbol{K}\longrightarrow N\restriction\boldsymbol{K}$ determines h (if any).

We have the forgetful functor $\mathcal{E}_{\boldsymbol{C},\boldsymbol{L}}=\mathcal{E}:\mathtt{Str}_{\boldsymbol{C}}(\boldsymbol{L})\longrightarrow\boldsymbol{C}^{\boldsymbol{K}}$; $\mathcal{E}$ is faithful, by the last remark. $\mathcal{E}$ is a fibration. Indeed, given $f:U\to V$ in $\boldsymbol{C}^{\boldsymbol{K}}$, and N over V (that is, $N\restriction\boldsymbol{K}=V$), then the Cartesian arrow $h:M\to N$ over f is obtained by defining M and h such that $M\restriction\boldsymbol{K}=U$, $h\restriction\boldsymbol{K}=f$ and, for all $R\in\mathtt{Rel}(\boldsymbol{L})$,

$$\begin{array}{ccc} M(R) & \xrightarrow{\;h_R\;} & N(R) \\ {\scriptstyle m^M_R}\big\downarrow & \square & \big\downarrow{\scriptstyle m^N_R} \\ U[R] & \xrightarrow[\;f_{[R]}\;]{} & V[R] \end{array}$$

is a pullback (it is immediate to see that h so defined is Cartesian). As usual with fibrations, let us denote M so defined by $f^*(N)$, and the Cartesian arrow h by $\theta_f:f^*(N)\to N$.

$\mathcal{E}$ is a fibration with fibers that are preorders.

When in particular $\boldsymbol{C}=\mathtt{Set}$ (which is the most important case), a functor $U:\boldsymbol{K}\to\mathtt{Set}$ is called *separated* if $U(K)\cap U(K')=\varnothing$ whenever K, K' are distinct objects of $\boldsymbol{K}$. For a separated U, we define $|U|=\bigcup_{K\in\boldsymbol{K}}U(K)$; for a general U, we would put $|U|=\bigsqcup_{K\in\boldsymbol{K}}U(K)=\{(K,a):k\in\boldsymbol{K},\ a\in U(K)\}$. Of course, every functor is isomorphic to a separated one. When $f:U\to V$, and U is separated, for $a\in|U|$ we may write $h(a)$ without ambiguity for $h_K(a)$ for which $a\in U(K)$. For notational simplicity, we will restrict attention to separated functors $\boldsymbol{K}\to\mathtt{Set}$.

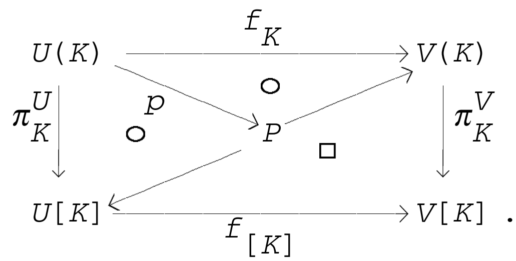
I will now isolate a property of a natural transformation $f:U\to V$ in $\boldsymbol{C}^{\boldsymbol{K}}$. Let first $\boldsymbol{C}=\mathtt{Set}$. We say that f is *very surjective* if whenever $K\in\boldsymbol{K}$, $\langle a_p\rangle_{p\in K|\boldsymbol{K}}\in U[K]$, the mapping

$$f_{\langle a_p\rangle_{p\in K|\boldsymbol{K}}}:UK(\langle a_p\rangle_{p\in K|\boldsymbol{K}})\longrightarrow VK(\langle fa_p\rangle_{p\in K|\boldsymbol{K}})\ :\ a\longmapsto f(a)$$

(see (3) in §1) is surjective.

For a general $\boldsymbol{C}$ (assumed to be regular), $f:U\to V$ in $\boldsymbol{C}^{\boldsymbol{K}}$ is *very surjective* if for every

$K \in \boldsymbol{K}$, the canonical map $p: U(K) \to P = U[K] \times_{V[K]} V(K)$ from the diagram below is surjective (a regular epimorphism):



It is clear that if f is an isomorphism (in $\boldsymbol{C}^{\boldsymbol{K}}$), then it is very surjective. It is easy to see (by induction on the level of $K \in \boldsymbol{K}$) that very surjective implies surjective (being a regular epimorphism in $\boldsymbol{C}^{\boldsymbol{K}}$), but not necessarily conversely.

In this section, we consider logic with dependent sorts only without equality; all $\boldsymbol{L}$-formulas are without equality.

(1) Let $f: U \to V$ in $\boldsymbol{C}^{\boldsymbol{K}}$ be very surjective, and any $N \in \mathrm{Str}_{\boldsymbol{C}}(\boldsymbol{L})$ over V. Let $h = \theta_f : M = f^*(N) \to N$.

(a) Let first $\boldsymbol{C} = \mathrm{Set}$. h is *elementary with respect to logic without equality* in the sense that for any context $\mathcal{X}$ and $\boldsymbol{L}$-formula φ (in logic with dependent sorts and without equality) with $\mathrm{Var}(\varphi) \subset \mathcal{X}$, and any $\langle a_x \rangle_{x \in \mathcal{X}} \in M[\mathcal{X}]$,

$$M \models \varphi[\langle a_x \rangle_{x \in \mathcal{X}}] \iff N \models \varphi[\langle ha_x \rangle_{x \in \mathcal{X}}] \quad .$$

(b) For a general $\boldsymbol{C}$ which is a Heyting category (to interpret all $\boldsymbol{L}$-formulas), for any φ and $\mathcal{X}$ as above, there is a pullback

$$\begin{array}{ccc} M[\mathcal{X}:\varphi] & \longrightarrow & N[\mathcal{X}:\varphi] \\ \downarrow & \square & \downarrow \\ U[\mathcal{X}] & \xrightarrow[f_{\mathcal{X}}]{} & V[\mathcal{X}] \end{array} \tag{1b}$$

(the vertical monomorphisms are representatives for the subobjects $M[\mathcal{X}:\varphi]\in S(U[\mathcal{X}])$, $N[\mathcal{X}:\varphi]\in S(V[\mathcal{X}])$; in other words, (1b) says $M[\mathcal{X}:\varphi] = (f_{\mathcal{X}})^{*}N[\mathcal{X}:\varphi]$).

here, $f_{\mathcal{X}}$ is the canonical map determined through by the definition of $U[\mathcal{X}]$, $V[\mathcal{X}]$ as limits in $\boldsymbol{C}$.

Obviously, (b) generalizes (a).

The proof for (a) can be given as a straightforward induction on the complexity of φ. The clause for atomic formulas is essentially the definition of M. For the propositional connectives, the induction step is automatic. By the duality in Set between $\exists$ and $\forall$, it is enough to handle the inductive step involving $\exists$, which is done using the "very surjective" assumption. In Appendix B, I will take a "fibrational" view of the notion of equivalence, and give a detailed proof of the more general form (b) .

Let M, N be $\boldsymbol{C}$-valued $\boldsymbol{L}$-structures. We say that they are $\boldsymbol{L}$-*equivalent*, and we write $M\sim_{\boldsymbol{L}}N$, if there is a diagram

$$M \xleftarrow{\;\bar{m}\;} P \xrightarrow{\;\bar{n}\;} N$$

in $\mathrm{Str}_{\boldsymbol{C}}(\boldsymbol{L})$ such that $\bar{m}\restriction\boldsymbol{K}$, $\bar{n}\restriction\boldsymbol{K}$ are very surjective, and $\bar{m}$ and $\bar{n}$ are Cartesian arrows in the fibration $\mathcal{E}_{\boldsymbol{C},\boldsymbol{L}}$. Paraphrased, this means that there exists a functor $W\in\boldsymbol{C}^{\boldsymbol{K}}$ and very surjective maps $m:W\to M\restriction\boldsymbol{K}$, $n:W\to N\restriction\boldsymbol{K}$ such that $m^{*}(M) = n^{*}(N)$, that is, for all $R\in\mathrm{Rel}(\boldsymbol{L})$,

$$\begin{array}{ccccccc}
M(R) & \longleftarrow & M(R)\times_{M[R]}W[R] & = & N(R)\times_{N[R]}W[R] & \longrightarrow & N(R) \\
\downarrow & & \square & \downarrow & \square & & \downarrow \\
M[R] & \xleftarrow{\quad m_{[R]}\quad} & & W[R] & & \xrightarrow{\quad n_{[R]}\quad} & N[R]
\end{array} \qquad (1')$$

(where the equality means equality of subobjects of $W[R]$). In case $\boldsymbol{C}=\mathrm{Set}$, (1') means that if $R\in\mathrm{Rel}(\boldsymbol{L})$, $\langle c_p\rangle_{p\in R|\boldsymbol{K}}\in W[R]$, then

$$\langle mc_p \rangle_{p\in R|\mathbf{K}} \in M(R) \iff \langle nc_p \rangle_{p\in R|\mathbf{K}} \in N(R) \quad . \tag{1"}$$

The data (W, m, n) are said to form an $\mathbf{L}$-*equivalence* of M and N; in notation, $(W, m, n) : M \underset{\mathbf{L}}{\longleftrightarrow} N$.

It is easy to see that the relation $\sim_{\mathbf{L}}$ is an equivalence relation (for a proof, see Appendix B). It is also clear that isomorphism of $\mathbf{L}$-structures implies $\mathbf{L}$-equivalence.

Let us write $M \equiv_{\mathbf{L}} N$ for: $M \models \sigma \iff N \models \sigma$ for all $\mathbf{L}$-sentences in logic with dependent sorts and without equality. We have

(2)(a) $M \sim_{\mathbf{L}} N \Longrightarrow M \equiv_{\mathbf{L}} N$.

This immediately follows from (1).

The word "equivalence" is used in " $\mathbf{L}$-equivalence" because of the relationship to the various notions of "equivalence" used in category theory; see later.

At this point, the reader may want to look at Appendix C, which may help understand the concept of $\mathbf{L}$-equivalence.

We now will exploit the fact that we have specified variables "with arbitrary parameters". In what follows, a *context* is a, not necessarily finite, set $\mathcal{Y}$ of variables such that $y \in \mathcal{Y}$, $x \in \mathrm{Dep}(y)$ imply that $x \in \mathcal{Y}$. When we want to refer to the previous sense of "context", we will say "finite context". A *specialization* is a map of contexts whose restriction to all finite subcontexts of the domain is a specialization in the original sense. Just as in case of finite contexts, there is a correspondence between contexts and functors $F : \mathbf{K} \to \mathrm{Set}$ which is an equivalence of the categories $\mathrm{Set}^{\mathbf{K}}$ and $\mathrm{Con}_{\infty}[\mathbf{K}]$, the category of all (small) contexts and specializations.

Given a context $\mathcal{Y}$ and an $\mathbf{K}$-structure M, the set $M[\mathcal{Y}]$ is defined by the formula (1), §1 (which was the definition of $M[\mathcal{Y}]$ for finite $\mathcal{Y}$). Given a formula φ with $\mathrm{Var}(\varphi)\subset\mathcal{Y}$, $M[\mathcal{Y}:\varphi]$ is the subset of $M[\mathcal{Y}]$ for which, for any $\langle a_y\rangle_{y\in\mathcal{Y}}\in M[\mathcal{Y}]$,

$$\langle a_y\rangle_{y\in\mathcal{Y}}\in M[\mathcal{Y}:\varphi] \iff \langle a_y\rangle_{y\in\mathcal{Y}'}\in M[\mathcal{Y}':\varphi]$$

for any (equivalently, some) finite context $\mathcal{Y}'$ with $\mathrm{Var}(\varphi)\subset\mathcal{Y}'\subset\mathcal{Y}$. As before, we write also $M\models\varphi[\langle a_y\rangle_{y\in\mathcal{Y}}]$ for $\langle a_y\rangle_{y\in\mathcal{Y}}\in M[\mathcal{Y}:\varphi]$.

Suppose $\mathcal{X}$ is a context, M, N $\boldsymbol{L}$-structures, $\vec{a}=\langle a_x\rangle_{x\in\mathcal{X}}\in M[\mathcal{X}]$, $\vec{b}=\langle b_x\rangle_{x\in\mathcal{X}}\in N[\mathcal{X}]$. We write

$$(W,m,n):(M,\vec{a})\xleftrightarrow[\boldsymbol{L}]{}(N,\vec{b}) \qquad (3)$$

if $(W,m,n):M\xleftrightarrow[\boldsymbol{L}]{}N$ and there is $\langle s_x\rangle_{x\in\mathcal{X}}\in W[\mathcal{X}]$ such that $ms_x=a_x$ and $ns_x=b_x$ for all $x\in\mathcal{X}$. We write $(M,\vec{a})\sim_{\boldsymbol{L}}(N,\vec{b})$ if there is (W,m,n) such that (3) holds.

With $M,N,\mathcal{X},\vec{a},\vec{b}$ as above, we write $(M,\vec{a})\equiv_{\boldsymbol{L}}(N,\vec{b})$ for: for all $\boldsymbol{L}$-formulas φ with $\mathrm{Var}(\varphi)\subset\mathcal{X}$, we have $M\models\varphi[\langle a_x\rangle_{x\in\mathcal{X}}] \iff N\models\varphi[\langle b_x\rangle_{x\in\mathcal{X}}]$.

We have the following generalization of (2)(a) :

(2)(b) $(M,\vec{a})\sim_{\boldsymbol{L}}(N,\vec{b}) \implies (M,\vec{a})\equiv_{\boldsymbol{L}}(N,\vec{b})$;

this also follows immediately from (1) .

Let $\mathcal{Y}$ be a context, x a variable such that $x\notin\mathcal{Y}$ but $\mathcal{Y}\dot{\cup}\{x\}$ is a context (thus, $x_{x,p}\in\mathcal{Y}$ for all $p\in K_x|\boldsymbol{K}$), and let Φ be a set of formulas in logic with dependent sorts over $\boldsymbol{L}$ such that $\mathrm{Var}(\Phi)=\bigcup_{\varphi\in\Phi}\mathrm{Var}(\varphi)\subset\mathcal{Y}\dot{\cup}\{x\}$; such Φ is called a *$\mathcal{Y}$-set* (of formulas; with x any variable as described with respect to $\mathcal{Y}$). Let M be an $\boldsymbol{L}$-structure, and $\vec{a}=\langle a_y\rangle_{y\in\mathcal{Y}}\in M[\mathcal{Y}]$. We say that Φ is *satisfiable in* $(M,\vec{a})$ if there is $a\in|M|$ (more precisely, $a\in M K_x[\langle a_{x_{x,p}}\rangle_{p\in K_x|\boldsymbol{K}}]$) such that $M\models\varphi[\vec{a}, a/x]$ (of course, $\vec{a}, a/x$ stands for $\langle a'_y\rangle_{y\in\mathcal{Y}\dot{\cup}\{x\}}$ for which $a'_y=a_y$ for $y\in\mathcal{Y}$, and $a'_x=a$). Φ is *finitely satisfiable in* $(M, \vec{a})$ if every finite subset of Φ is satisfiable in $(M, \vec{a})$. M is said to be $\mathcal{Y}$-$\boldsymbol{L}$-*saturated* if for every $\vec{a}\in M[\mathcal{Y}]$ and every $\mathcal{Y}$-set Φ, if Φ is finitely satisfiable in $(M, \vec{a})$, then Φ is satisfiable in $(M, \vec{a})$.

Let κ be an infinite cardinal. We say that M is κ, $\boldsymbol{L}$-*saturated* if it is $\mathcal{Y}$-$\boldsymbol{L}$-saturated for every context $\mathcal{Y}$ with cardinality smaller than κ.

For saturated models for ordinary first order logic, see [CK]. In [MR2], one can find a detailed introduction to saturated and special models for multisorted logic; the basic facts and their proofs in the multisorted context do not essentially differ from the original one-sorted versions.

κ, $\boldsymbol{L}$-saturation is κ-saturation with respect to $\boldsymbol{L}$-formulas. Since $\boldsymbol{L}$-formulas form a part of the multisorted formulas over $|\boldsymbol{L}|$, it is clear that if M, an $\boldsymbol{L}$-structure, is κ-saturated as a structure for the similarity type $|\boldsymbol{L}|$, then M is κ, $\boldsymbol{L}$-saturated. More generally, suppose that we have "interpreted" $\boldsymbol{L}$ in a theory S in ordinary multisorted first-order logic; that is, we have a $\boldsymbol{C}$-valued $\boldsymbol{L}$-structure $I:\boldsymbol{L}\longrightarrow\boldsymbol{C}$, for $\boldsymbol{C}$ the Lindenbaum-Tarski category $[S]$ of S (see [MR]; $[S]$ is a Boolean category). Then if M is a model of S, or equivalently, $M:\boldsymbol{C}\to\mathrm{Set}$ is a coherent functor, and M is κ-saturated in the ordinary sense, then the $\boldsymbol{L}$-structure $M\upharpoonright\boldsymbol{L}=MI:\boldsymbol{L}\to\mathrm{Set}$ is κ, $\boldsymbol{L}$-saturated.

By the *cardinality of* the structure M, $\#M$, we mean the cardinality of its underlying set $|M|$.

(4) Suppose the $\boldsymbol{L}$-structures M, N are κ, $\boldsymbol{L}$-saturated, and both are of cardinality $\leq\kappa$. Then the converses of (2)(a) and (2)(b) hold:

$$M\equiv_{\boldsymbol{L}} N \Longrightarrow M\sim_{\boldsymbol{L}} N ;$$

and more generally, if $\mathcal{X}$ is a context of size $<\kappa$, $\vec{a}\in M[\mathcal{X}]$, $\vec{b}\in N[\mathcal{X}]$, then

$$(M, \vec{a}) \equiv_{\boldsymbol{L}} (N, \vec{b}) \implies (M, \vec{a}) \sim_{\boldsymbol{L}} (N, \vec{b}) \ .$$

Proof.

For a given infinite cardinal κ, and a given context $\mathcal{X}$ of cardinality less than κ, let $\mathcal{U}=\mathcal{U}[\kappa, \mathcal{X}]$ be a context such that $\#\mathcal{U} = \kappa$, $\mathcal{X}\subset\mathcal{U}$, and for every sort X with $\mathtt{Var}(X)\subset\mathcal{U}$, the cardinality of the set of variables $x\in\mathcal{U}$ with $x{:}X$ is equal to κ. It is easy to see that such an $\mathcal{U}$ exists; we define contexts $\mathcal{U}_i$ by recursion on $i\le k$ for k the height of $\boldsymbol{K}$; let $\mathcal{U}_0=\emptyset$; if $\mathcal{U}_i$ has been defined, pick, for every sort X whose kind is of level i and for which $\mathtt{Var}(X) \subset \mathcal{U}_i$, a set V_X of variables $v{:}X$ such that $\#V_X=\kappa$, and let $\mathcal{U}_{i+1}$ be the union of $\mathcal{U}_i$ and all the V_X for all such X; if $k=\omega$, let $\mathcal{U}_\omega=\bigcup_{i<\omega}\mathcal{U}_i$; let $\mathcal{U}=\mathcal{U}_k$.

Next, enumerate $\mathcal{U}$ as a sequence $\langle u_\alpha\rangle_{\alpha<\kappa}$ in such a way that for each $\beta<\kappa$, $\langle u_\alpha\rangle_{\alpha<\beta}$ is a context; equivalently, such that for each $\beta<\kappa$, $\mathtt{Dep}(u_\beta)\subset\{u_\alpha : \alpha<\beta\}$. Note first of all that for any finite context $\mathcal{Y}$, there is an enumeration $\mathcal{Y}=\{y_i : i<n\}$ such that $\langle y_i\rangle_{i<j}$ is a context for all $j<n$; enumerate first the level-0 variables, next the level-1 ones, etc. Call such an enumeration of $\mathcal{Y}$ "good". Now, take first an arbitrary enumeration $\langle v_\alpha\rangle_{\alpha<\kappa}$ of $\mathcal{U}$; define the increasing sequence $\langle \beta_\alpha\rangle_{\alpha<\kappa}$ of ordinals and the partial enumeration $\langle u_\gamma\rangle_{\gamma<\beta_\alpha}$ by induction on α as follows. For a limit ordinal α, $\beta_\alpha=\lim_{\delta<\alpha}\beta_\delta$. For $\alpha=\delta+1$, let $\langle u_{\beta_\delta+i}\rangle_{i<n}$ be a good enumeration of $\mathtt{Dep}(v_\delta)\cup\{v_\delta\}$, and let $\beta_\alpha=\beta_\delta+n$.

For every sort X such that $\mathtt{Var}(X)\subset\mathcal{U}$, let $\langle u_{\alpha_{X,\nu}}\rangle_{\nu<\kappa}$ be an enumeration in increasing order of all u_α of sort X for which $u_\alpha\notin\mathcal{X}$. Finally, for any $\alpha<\kappa$, let $\nu[\alpha]$ be the ordinal ν for which $\alpha_{X,\nu} = \alpha$ where X is the sort of u_α.

Assume $\mathcal{X}$ is a context of size $<\kappa$, $\#M,\#N\le\kappa$, $\vec{a}=\langle a_x\rangle_{x\in\mathcal{X}}\in M[\mathcal{X}]$, $\vec{b}=\langle b_x\rangle_{x\in\mathcal{X}}\in N[\mathcal{X}]$, and $(M, \vec{a}) \equiv_{\boldsymbol{L}} (N, \vec{b})$. For any M-sort $MK(\langle c_p\rangle_{p\in K|\boldsymbol{K}})=MK(\vec{c})$, let us fix an enumeration $\langle e_\xi\rangle_{\xi<\lambda} = \langle e_{K,\vec{c},\xi}\rangle_{\xi<\lambda_{K,\vec{c}}}$ of the set $MK(\vec{c})$; here, $\lambda_{K,\vec{c}} \le \kappa$.

Consider $\mathcal{U}=\mathcal{U}[\kappa, \mathcal{X}]$ constructed above.

We define a context $\mathcal{Z}$, a subset of $\mathcal{U}$, by deciding, recursively on $\alpha<\kappa$, whether u_α belongs to $\mathcal{Z}$ or not; furthermore, we also define, for each $u_\alpha\in\mathcal{Z}$, elements $c_\alpha\in|M|$ and $d_\alpha\in|N|$. Let $\mathcal{Z}_\alpha$ denote the set of all u_β with $\beta<\alpha$ for which $u_\beta\in\mathcal{Z}$, and $\vec{c}[\alpha]$ be the sequence $\langle c_z\rangle_{z\in\mathcal{X}\cup\mathcal{Z}_\alpha}\in M[\mathcal{X}\cup\mathcal{Z}_\alpha]$ for which $c_x=a_x$ ($x\in\mathcal{X}$) and $c_{u_\beta}=c_\beta$ ($u_\beta\in\mathcal{Z}_\alpha$). Similarly, we have $\vec{d}[\alpha]\in N[\mathcal{X}\cup\mathcal{Z}_\alpha]$. The induction hypothesis of the construction is that

$$(M,\vec{c}[\alpha+1])\equiv_{\boldsymbol{L}}(M,\vec{d}[\alpha+1])\ . \tag{5}$$

Suppose $\alpha<\kappa$, and $\mathcal{Z}_\alpha$, $\vec{c}[\alpha]$, $\vec{d}[\alpha]$ have been defined so that, for all $\beta<\alpha$, $(M,\vec{c}[\beta+1])\equiv_{\boldsymbol{L}}(M,\vec{d}[\beta+1])$. Since in the definition of " $\equiv_{\boldsymbol{L}}$ ", formulas with finitely many free variables are involved, we can conclude that

$$(M,\vec{c}[\alpha])\equiv_{\boldsymbol{L}}(M,\vec{d}[\alpha])\ . \tag{6}$$

Look at the variable u_α and its sort X. If $u_\alpha\in\mathcal{X}$, we let $u_\alpha\in\mathcal{Z}$, $c_\alpha=a_{u_\alpha}$, $d_\alpha=b_{u_\alpha}$. (5) is now an automatic consequence of (6).

If not all the variables in X (which are u_β's for $\beta<\alpha$) are in $\mathcal{Z}$, then $u_\alpha\notin\mathcal{Z}$, and we are finished with the stage α.

Assume that $u_\alpha\notin\mathcal{X}$ and all the variables in X are in $\mathcal{Z}$. Look at the ordinal $\nu=\nu[\alpha]$; write ν in the form $\nu=2\cdot\mu$ or $\nu=2\cdot\mu+1$ as the case may be. Let first $\nu=2\cdot\mu$. With $X=K(\langle u_{\beta_p}\rangle_{p\in K|\boldsymbol{K}})$, consider the M-sort $MK(\langle c_{\beta_p}\rangle_{p\in K|\boldsymbol{K}})=MK(\vec{c})$ and its previously fixed enumeration $\langle e_\xi\rangle_{\xi<\lambda}$ ($=\langle e_{K,\vec{c},\xi}\rangle_{\xi<\lambda_{K,\vec{c}}}$). If $\mu\geq\lambda$, then again $u_\alpha\notin\mathcal{Z}$. If, however, $\mu<\lambda$, then $u_\alpha\in\mathcal{Z}$. Moreover, $c_\alpha \overset{\text{def}}{=} e_\mu$.

Let Φ be the $\mathcal{X}\cup\mathcal{Z}_\alpha$-set of all formulas φ with $\mathrm{Var}(\varphi)\subset\mathcal{X}\cup\mathcal{Z}_\alpha\dot{\cup}\{u_\alpha\}$ for which $M\models\varphi[\vec{c}[\alpha],e_\mu/u_\alpha]$. I claim that Φ is finitely satisfiable in $(N,\vec{d}[\alpha])$. Let Ψ be a finite subset of Φ. For $\varphi=\bigwedge\Psi$, we have $M\models\varphi[\vec{c}[\alpha],e_\mu/u]$, hence, $M\models(\exists u_\alpha\varphi)[\vec{c}[\alpha]]$ (note that $\exists u_\alpha\varphi$ is well-formed, since for every $z\in\mathrm{Var}(\varphi)$, $z\neq u_\alpha$, we have $z\in\mathcal{X}\cup\mathcal{Z}_\alpha$, hence $\mathrm{Dep}(z)\subset\mathcal{X}\cup\mathcal{Z}_\alpha$, and $u_\alpha\notin\mathrm{Dep}(z)$). As a consequence, by (6),

$N \models (\exists u\varphi)[\vec{d}[\alpha]]$. This means that Ψ is satisfiable in $(N, \vec{d}[\alpha])$ as desired.

Since $\#(\mathcal{X} \cup \mathcal{Z}_\alpha) < \kappa$, and N is κ, $\boldsymbol{L}$-saturated, Φ is satisfiable in $(N, \vec{d}[\alpha])$, by $d_\alpha \in NK(\langle d_{\beta_p} \rangle_{p \in K | \boldsymbol{K}})$, say. The choice of Φ ensures that (5) holds.

In case $\nu = 2 \cdot \mu + 1$, we proceed similarly, with the roles of M and N interchanged.

With the construction completed, we put $\mathcal{Z} = \bigcup_{\alpha < \kappa} \mathcal{Z}_\alpha$. We let W be the functor $F_{\mathcal{Z}} : \boldsymbol{K} \to \mathrm{Set}$ associated with the context $\mathcal{Z}$ (see §4). $m : W \to M \upharpoonright \boldsymbol{K}$, $n : W \to N \upharpoonright \boldsymbol{K}$ are defined by $m(u_\alpha) = c_\alpha$, $n(u_\alpha) = d_\alpha$ ($u_\alpha \in \mathcal{Z}$). The definition ensures that $\mathcal{X} \subset \mathcal{Z}$ and $m(x) = a_x$, $n(x) = b_x$ ($x \in \mathcal{X}$).

Let us see that m is very surjective. Let $K \in \boldsymbol{K}$. $W[K]$ is the set of all tuples $\langle z_p \rangle_{p \in K | \boldsymbol{K}}$ for which each $z_p \in \mathcal{Z}$, and $X = K(\langle z_p \rangle_{p \in K | \boldsymbol{K}})$ is a (well-formed) sort; $WK(\langle z_p \rangle_{p \in K | \boldsymbol{K}})$ is the set of all $z \in \mathcal{Z}$ such that $z : X$. So, assume that $X = K(\langle z_p \rangle_{p \in K | \boldsymbol{K}}) = K(\langle u_{\beta_p} \rangle_{p \in K | \boldsymbol{K}})$ is a sort, and

$$a \in MK(\langle mz_p \rangle_{p \in K | \boldsymbol{K}}) = MK(\langle c_{\beta_p} \rangle_{p \in K | \boldsymbol{K}}) = MK(\vec{c}) \quad .$$

Then $a = e_{K, \vec{c}, \mu}$ for some $\mu < \lambda_{K, \vec{c}}$, and for $\alpha = \alpha_{X, 2 \cdot \mu}$, the construction at stage α puts $u_\alpha : X$ into $\mathcal{Z}$; that is, $u_\alpha \in WK(\langle z_p \rangle_{p \in K | \boldsymbol{K}})$, with $a = c_\alpha = mu_\alpha$ as desired.

The fact that n is very surjective is seen analogously.

We have that $(W, m, n) : M \xleftrightarrow[\boldsymbol{L}]{} N$, since (1") is a consequence of (5) being true for all $\alpha < \kappa$; one has to apply (5) to atomic formulas.

This completes the proof of (4).

Let $\boldsymbol{C}$ be a small Boolean category. By a *model of* $\boldsymbol{C}$ we mean a functor $M : \boldsymbol{C} \to \mathrm{Set}$ preserving the Boolean structure (that is, M is a coherent functor). We write $M \models \boldsymbol{C}$ to say that M is a model of $\boldsymbol{C}$.

There is a theory $\mathrm{T}_{\boldsymbol{C}}=(\mathrm{L}_{\boldsymbol{C}},\Sigma_{\boldsymbol{C}})$ in multisorted first-order logic, with $\mathrm{L}_{\boldsymbol{C}}$ the underlying graph of $\boldsymbol{C}$, such that the models of $\boldsymbol{C}$ are *the same* as the models of $\mathrm{T}_{\boldsymbol{C}}$ (note that both the models of $\boldsymbol{C}$ and the models of $\mathrm{T}_{\boldsymbol{C}}$ are particular diagrams $\mathrm{L}_{\boldsymbol{C}}\to\mathrm{Set}$). Moreover, for any subobject $\varphi\in\mathrm{S}_{\boldsymbol{C}}(A)$, $A\in\boldsymbol{C}$, there is a (simply defined) $\mathrm{L}_{\boldsymbol{C}}$-formula $\underline{\varphi}(x)$ with a single free variable $x:A$ such that for every $M\models\boldsymbol{C}$ and $a\in M(A)$, $M\models\underline{\varphi}[a]$ ($\Longleftrightarrow M\models\underline{\varphi}[a/x]$) iff $a\in M(\varphi)$ ($\subset M(A)$). See [MR].

For $\sigma\in\mathrm{S}(1_{\boldsymbol{C}})$, a subobject of the terminal object in $\boldsymbol{C}$, we write $M\models\sigma$ for $M(\sigma)=1$ in Set. We will call a subobject of $1_{\boldsymbol{C}}$ a *sentence* in $\boldsymbol{C}$.

Let $I:\boldsymbol{L}\to\boldsymbol{C}$ a $\boldsymbol{C}$-valued $\boldsymbol{L}$-structure (in particular, $I:\boldsymbol{L}\to\boldsymbol{C}$ is a functor from $\boldsymbol{L}$ as a category). When $\boldsymbol{C}$ is the Lindenbaum-Tarski category $[S]$ of a theory $S=(\mathrm{L}_S,\Sigma_S)$ in ordinary multisorted logic (see [MR] or [M?]), then such an I is what we should consider an *interpretation* of the DS vocabulary $\boldsymbol{L}$ in the theory S. An example is obtained by taking $S=(|\boldsymbol{L}|,\Sigma[\boldsymbol{L}])$ (for $\Sigma[\boldsymbol{L}]$, see §1), and for $I:\boldsymbol{L}\to[S]$ the $[S]$-structure defined by $I(A)=[a:\mathbf{t}]$ for $A\in\boldsymbol{L}$ where $a:A$, and for $f:A\to B$, $I(f)=\langle a\mapsto b:fa=b\rangle:[a:\mathbf{t}]\to[b:\mathbf{t}]$. $I:\boldsymbol{L}\to[S]$ is the *canonical* interpretation of logic with dependent types in multisorted logic. In this case, for any formula φ of FOLDS over $\boldsymbol{L}$, with $\mathrm{Var}(\varphi)\subset\mathcal{X}$, we have $I[\mathcal{X}:\varphi]=m^*[\mathcal{X}:\varphi^*]$; here, $m:I[\mathcal{X}:\varphi]\rightarrowtail\{\mathcal{X}\}\overset{=}{\mathrm{def}}\prod_{x\in\mathcal{X}}K_x$ is the canonical monomorphism, m^* denotes pulling back along m; φ^* was defined in §1.

For a general $I:\boldsymbol{L}\to\boldsymbol{C}$, and for an $\boldsymbol{L}$-sentence θ, let us write $I(\theta)$ for the sentence $I[\emptyset:\theta]$ of $\boldsymbol{C}$. In case $\boldsymbol{C}=[S]$, $I(\theta)$ also stands for any one of the S-equivalent L_S-sentences which are the representatives of the $\boldsymbol{C}$-subobject $I(\theta)$.

When $M\models\boldsymbol{C}$, the composite $MI:\boldsymbol{L}\to\mathrm{Set}$ is an $\boldsymbol{L}$-structure. We also write $M\restriction\boldsymbol{L}$ for MI; $M\restriction\boldsymbol{L}$ is the *$\boldsymbol{L}$-reduct of* M (via I).

Let $\boldsymbol{C}$ and $\boldsymbol{D}$ be small Boolean categories, $I:\boldsymbol{L}\to\boldsymbol{C}$ and $J:\boldsymbol{L}\to\boldsymbol{D}$. Notational conventions introduced above for $I:\boldsymbol{L}\to\boldsymbol{C}$ are valid for $J:\boldsymbol{L}\to\boldsymbol{D}$, *mutatis mutandis.*

(7)(a) Assume that σ is a sentence of $\boldsymbol{C}$, τ a sentence of $\boldsymbol{D}$, and for all $M\models\boldsymbol{C}$, $N\models\boldsymbol{D}$,

$$M \models \sigma \;\; \& \;\; M{\upharpoonright}\mathbf{L} \sim_{\mathbf{L}} N{\upharpoonright}\mathbf{L} \;\; \Longrightarrow \;\; N \models \tau \; .$$

Then there is an $\mathbf{L}$-sentence θ in logic with dependent sorts without equality such that for all $M \models \mathbf{C}$, $N \models \mathbf{D}$, we have

$$M \models \sigma \Longrightarrow M{\upharpoonright}\mathbf{L} \models \theta \qquad \text{and} \qquad N{\upharpoonright}\mathbf{L} \models \theta \Longrightarrow N \models \tau \; .$$

For a more general formulation, consider a finite $\mathbf{L}$-context $\mathcal{X}$, and the object $I[\mathcal{X}] \in \mathbf{C}$. $I[\mathcal{X}]$ is defined as a finite limit in $\mathbf{C}$; see the end of §1; let $\pi_{[x]} : I[\mathcal{X}] \to I(\mathrm{K}_x)$ be the limit projections ($x \in \mathcal{X}$). Given any $M \models \mathbf{C}$, we have similar projections $\rho_{[x]} : (M{\upharpoonright}\mathbf{L})[\mathcal{X}] \to MI(\mathrm{K}_x)$ in $\mathtt{Set}$, and a canonical isomorphism $\mu : (M{\upharpoonright}\mathbf{L})[\mathcal{X}] \xrightarrow{\cong} M(I[\mathcal{X}])$ making each diagram

$$\begin{array}{ccc} (M{\upharpoonright}\mathbf{L})[\mathcal{X}] & \xrightarrow[\cong]{\mu} & M(I[\mathcal{X}]) \\ & {\scriptstyle \rho_{[x]}} \searrow \quad \circ \quad \swarrow {\scriptstyle M(\pi_{[x]})} & \\ & MI(\mathrm{K}_x) & \end{array} \tag{7'}$$

commute. If $\vec{a} = \langle a_x \rangle_{x \in \mathcal{X}} \in (M{\upharpoonright}\mathbf{L})[\mathcal{X}]$, we write $\langle \vec{a} \rangle$ for $\mu(\vec{a}) \in M(I[\mathcal{X}])$. Once again, similar conventions apply in the context of $J : \mathbf{L} \to \mathbf{D}$.

(7)(b) Assume that $\mathcal{X}$ is a finite $\mathbf{L}$-context, $\sigma \in \mathrm{S}_{\mathbf{C}}(I[\mathcal{X}])$, $\tau \in \mathrm{S}_{\mathbf{D}}(J[\mathcal{X}])$, and for all $M \models \mathbf{C}$, $N \models \mathbf{D}$, $\vec{a} \in (M{\upharpoonright}\mathbf{L})[\mathcal{X}]$, $\vec{b} \in (N{\upharpoonright}\mathbf{L})[\mathcal{X}]$,

$$\langle \vec{a} \rangle \in M(\sigma) \;\; \& \;\; (M{\upharpoonright}\mathbf{L}, \vec{a}) \sim_{\mathbf{L}} (N{\upharpoonright}\mathbf{L}, \vec{b}) \;\; \Longrightarrow \;\; \langle \vec{b} \rangle \in N(\tau) \; . \tag{8}$$

Then there is an $\mathbf{L}$-formula θ in logic with dependent sorts without equality with $\mathrm{Var}(\varphi) \subset \mathcal{X}$ such that

$$\sigma \leq_{I[\mathcal{X}]} I[\mathcal{X} : \theta] \; , \qquad J[\mathcal{X} : \theta] \leq_{J[\mathcal{X}]} \tau \; . \tag{8'}$$

Note that (8') may be written equivalently as for all $M\models\boldsymbol{C}$, $N\models\boldsymbol{D}$, $\vec{a}\in(M\upharpoonright\boldsymbol{L})[\mathcal{X}]$ and $\vec{b}\in(N\upharpoonright\boldsymbol{L})[\mathcal{X}]$,

$$\langle\vec{a}\rangle\in M(\sigma) \Longrightarrow M\upharpoonright I\models\theta[\vec{a}] \text{ and } N\upharpoonright J\models\theta[\vec{b}] \Longrightarrow \langle\vec{b}\rangle\in N(\tau) .$$

Proof. Let us extend the vocabulary $\mathrm{L}_{\boldsymbol{C}}$ to $\mathrm{L}_{\boldsymbol{C}}(c)$ by adding a single new individual constant c of sort $A \underset{\mathrm{def}}{=} I[\mathcal{X}]$. For any $\varphi\in\mathrm{S}_{\boldsymbol{C}}(A)$, let $\varphi(c)$ denote $\underline{\varphi}(c/x)$, the result of substituting c for x in $\underline{\varphi}(x)$. For an $\boldsymbol{L}$-formula θ with $\mathrm{Var}(\theta)\subset\mathcal{X}$, let $\theta(c)$ stand for $(I[\mathcal{X}:\theta])(c)$. Similarly, we introduce $d:B\underset{\mathrm{def}}{=}J[\mathcal{X}]$; for $\psi\in\mathrm{S}_{\boldsymbol{D}}(B)$, $\psi(d)$ and for θ as before, $\theta(d)$.

Let Θ be the set of all $\boldsymbol{L}$-formulas θ with $\mathrm{Var}(\theta)\subset\mathcal{X}$ such that $\sigma\leq_A I[\mathcal{X}:\theta]$. Consider the set $\Sigma \underset{\mathrm{def}}{=} \Sigma_{\boldsymbol{D}}\cup\{\theta(d):\theta\in\Theta\}$ of $\mathrm{L}_{\boldsymbol{D}}(d)$-sentences. I claim that

$$(\mathrm{L}_{\boldsymbol{D}}(d),\Sigma)\models\tau(d) . \tag{9}$$

Once the claim is proved, by compactness there are finitely many $\theta_i\in\Theta$ ($i<n$) such that $(\mathrm{L}_{\boldsymbol{D}}(d),\ \Sigma_{\boldsymbol{D}}\cup\{\theta_i(d):i<n\})\models\tau(d)$, which means, for $\theta=\bigwedge_{i<n}\theta_i\in\Theta$ that $(\mathrm{L}_{\boldsymbol{D}}(d),\Sigma_{\boldsymbol{D}})\models\theta(d)\to\tau(d)$, that is, $(\mathrm{L}_{\boldsymbol{D}}(d),\Sigma_{\boldsymbol{D}})\models\forall x{:}B.(\underline{\theta}(x)\to\underline{\tau}(x))$, which means $J[\mathcal{X}:\theta]\leq_B\tau$; thus, it is enough to see the claim.

Assume that there is an infinite cardinal $\lambda\geq\#\mathrm{L}_{\boldsymbol{C}}$ such that $\lambda^{+}=2^{\lambda}$ (see below for the legitimacy of this assumption). Let $\kappa=\lambda^{+}$. According to the existence theorem for saturated models (see [CK], [MR2]), any $\mathrm{L}_{\boldsymbol{D}}(d)$-structure is elementarily equivalent to a κ-saturated structure of cardinality $\leq\kappa$. Therefore, to show (9), take $(N,b/d)$, a κ-saturated model of cardinality $\leq\kappa$ of $(\mathrm{L}_{\boldsymbol{D}}(d),\Sigma)$, to show $(N,b/d)\models\tau(d)$.

Let Φ be the set of $\boldsymbol{L}$-formulas φ with $\mathrm{Var}(\varphi)\subset\mathcal{X}$ such that $b\in N(I[\mathcal{X}:\varphi])\subset NB$; for every $\boldsymbol{L}$-formula φ with $\mathrm{Var}(\varphi)\subset\mathcal{X}$, exactly one of φ, $\neg\varphi$ belongs to Φ. Since $(N,b/d)$ is a model of $(\mathrm{L}_{\boldsymbol{D}}(d),\Sigma)$, with Σ defined as it is, we have $\Theta\subset\Phi$. I make the subclaim that the theory

$$(\mathrm{L}_{\boldsymbol{C}}(c),\ \Sigma_{\boldsymbol{C}}\cup\{\sigma(c)\}\cup\{\varphi(c):\varphi\in\Phi\}) \tag{10}$$

is consistent. Consider a finite subset $\{\varphi_i:i<n\}$ of Φ. If

$(\mathrm{L}_{\boldsymbol{C}}(c),\ \Sigma_{\boldsymbol{C}}\cup\{\sigma(c)\}\cup\{\varphi_i(c):i<n\})$ were not consistent, then we would have, for $\varphi=\bigwedge_{i<n}\varphi_i\in\Phi$, that $\sigma\leq_A I[\mathcal{X}:\neg\varphi]$, which would mean that $\neg\varphi\in\Theta\subset\Phi$, contradicting $\varphi\in\Phi$. This shows the subclaim.

Now, let $(M,\ a/c)$ be a κ-saturated model of (10) of cardinality $\leq\kappa$. Let $\vec{a}\in(M\restriction\boldsymbol{L})[\mathcal{X}]$ such that $a=\langle\vec{a}\rangle$ (see (7')) and $\vec{b}\in(N\restriction\boldsymbol{L})[\mathcal{X}]$ such that $b=\langle\vec{b}\rangle$. Then, for any $\boldsymbol{L}$-formula θ with $\mathrm{Var}(\theta)\subset\mathcal{X}$ such that $M\restriction\boldsymbol{L}\models\theta[\vec{a}]$, we have $\neg\theta\notin\Phi$, hence $\theta\in\Phi$, hence $N\restriction\boldsymbol{L}\models\theta[\vec{b}]$. This says that $(M\restriction\boldsymbol{L},\ \vec{a})\equiv_{\boldsymbol{L}}(N\restriction\boldsymbol{L},\ \vec{b})$. By (4), $(M\restriction\boldsymbol{L},\ \vec{a})\sim_{\boldsymbol{L}}(N\restriction\boldsymbol{L},\ \vec{b})$, and by the (8), the assumption of the proposition, $\langle\vec{b}\rangle\in N(\tau)$, that is, $N\models\underline{\tau}[\langle\vec{b}\rangle/x]$, that is, $(N,\ b/d)\models\tau(d)$ as promised.

The set-theoretic assumption used in the proof is redundant, by a general absoluteness theorem (arithmetic statements are absolute with respect to the constructible universe, in which the Generalized Continuum Hypothesis (GCH) holds; see [J]). On the other hand, one may use "special" models in place of saturated ones, and avoid the use of GCH; see [CK], [MR2].

(11)(a) Assume that S is a theory in multisorted logic, and $I:\boldsymbol{L}\rightarrow[S]$ is an interpretation of the DSV $\boldsymbol{L}$ in S. Suppose that the class $\mathrm{Mod}(S)$ of models of S is invariant under $\boldsymbol{L}$-equivalence in the sense that for any L_S-structures M and N, $M\in\mathrm{Mod}(S)$ and $M\restriction\boldsymbol{L}\sim_{\boldsymbol{L}}N\restriction\boldsymbol{L}$ imply that $N\in\mathrm{Mod}(S)$. Then S is $\boldsymbol{L}$-axiomatizable; that is, for a set Θ of $\boldsymbol{L}$-sentences, $\mathrm{Con}_{\mathrm{L}_S}(\{I(\theta):\theta\in\Theta\})=\mathrm{Con}_{\mathrm{L}_S}(\Sigma_S)$; here, $\mathrm{Con}_L(\Phi)$ is the set of L-sentences that are consequences of the theory $(L,\ \Phi)$.

Note that the conclusion can also be expressed by saying that for any L_S-structure M, $M\models\Sigma_S$ iff $M\restriction\boldsymbol{L}\models\Theta$.

(11)(b) More generally, assume, in addition to S and $I:\boldsymbol{L}\rightarrow[S]$, a theory T in a language extending that of S ($\mathrm{L}_S\subset\mathrm{L}_T$) such that

for any $M,\ N\in\mathrm{Mod}(T)$, $M\restriction\mathrm{L}_S\in\mathrm{Mod}(S)$ and $M\restriction\boldsymbol{L}\sim_{\boldsymbol{L}}N\restriction\boldsymbol{L}$ imply that

$N{\restriction}L_S \in \mathrm{Mod}(S)$.

Then, there is a set Θ of $\boldsymbol{L}$-sentences such that, for any $M \models T$, $M \models \Sigma_S$ iff $M{\restriction}\boldsymbol{L} \models \Theta$.

(11)(a) is the special case when $T=(L_S, \emptyset)$.

Proof of (11)(b). For any $\tau \in \Sigma_S$, $M \models T$ and $N \models T$, we have

$$M \models \Sigma_S \ \&\ M{\restriction}\boldsymbol{L} \sim_{\boldsymbol{L}} N{\restriction}\boldsymbol{L} \implies N \models \tau .$$

By appropriately coding the condition $M{\restriction}\boldsymbol{L} \sim_{\boldsymbol{L}} N{\restriction}\boldsymbol{L}$ in first order logic with suitable additional primitives, and by applying compactness, we can find $\sigma[\tau]$, a finite conjunction of elements of Σ_S, such that for any $M \models T$ and $N \models T$,

$$M \models \sigma[\tau] \ \&\ M{\restriction}\boldsymbol{L} \sim_{\boldsymbol{L}} N{\restriction}\boldsymbol{L} \implies N \models \tau .$$

Then by (7)(a), applied to $\boldsymbol{C}=\boldsymbol{D}=[T]$, and $I=J:\boldsymbol{L} \xrightarrow{I} [S] \xrightarrow{\mathrm{incl}} [T]$, we can find $\theta[\tau]$, an $\boldsymbol{L}$-sentence, such that $T \models \sigma[\tau] \longrightarrow I(\theta[\tau])$, $T \models I(\theta[\tau]) \longrightarrow \tau$. Clearly, $\Theta = \{\theta[\tau] : \tau \in \Sigma_S\}$ is then appropriate for the assertion.

We leave it to the reader to formulate a version of (11) with formulas in a given context $\mathcal{X}$ instead of sentences.

The following, which is a special case of (7)(b), says that a first-order property invariant under $\boldsymbol{L}$-equivalence is expressible in logic with dependent types over $\boldsymbol{L}$.

(12) Let $I:\boldsymbol{L} \to \boldsymbol{C}$ be as before. Assume that $\mathcal{X}$ is a finite $\boldsymbol{L}$-context, $\sigma \in \mathrm{S}(I[\mathcal{X}])$, and for all $M, N \models \boldsymbol{C}$ and $\vec{a} \in (M{\restriction}\boldsymbol{L})[\mathcal{X}]$, $\vec{b} \in (N{\restriction}\boldsymbol{L})[\mathcal{X}]$,

$$\langle \vec{a} \rangle \in M(\sigma) \ \ \&\ \ (M{\restriction}\boldsymbol{L}, \vec{a}) \sim_{\boldsymbol{L}} (N{\restriction}\boldsymbol{L}, \vec{b}) \implies \langle \vec{b} \rangle \in N(\sigma) .$$

Then there is an $\boldsymbol{L}$-formula θ in logic with dependent sorts without equality with $\mathrm{Var}(\theta) \subset \mathcal{X}$ such that $\sigma =_{I[\mathcal{X}]} I[\mathcal{X}:\theta]$.

The notion of $\boldsymbol{L}$-equivalence as defined is relevant to FOLDS without equality. However, frequently we deal with FOLDS with restricted equality. As explained in §1, when M is an $\boldsymbol{L}$-structure, it can be considered as an $\boldsymbol{L}^{\mathrm{eq}}$-structure, with the additional relations E_K interpreted as true equality; let us write M for the resulting "standard" $\boldsymbol{L}^{\mathrm{eq}}$-structure as well. What does it mean to have an equivalence $(W, m, n): M \xleftrightarrow[\boldsymbol{L}^{\mathrm{eq}}]{} N$ for $\boldsymbol{L}$-structures M, N?

Clearly, this is to say that $(W, m, n): M \xleftrightarrow[\boldsymbol{L}]{} N$ and, for any maximal kind K, and $\vec{c} \in W[K]$, $c_1, c_2 \in WK(m\vec{c})$, we have that $mc_1 = mc_2$ iff $nc_1 = nc_2$. Let us write $(W, m, n): M \xleftrightarrow[\boldsymbol{L}]{\approx} N$ for $(W, m, n): M \xleftrightarrow[\boldsymbol{L}^{\mathrm{eq}}]{} N$, and let us call such (W, m, n) an $\boldsymbol{L}$, $\approx$-equivalence; also, write $M \approx_{\boldsymbol{L}} N$ for $M \sim_{\boldsymbol{L}^{\mathrm{eq}}} N$; note that throughout, M and N are $\boldsymbol{L}$-structures.

Let us define $M \equiv_{\boldsymbol{L}^{=}} N$ as we did $M \equiv_{\boldsymbol{L}} N$ above, except that we refer to logic with equality. Then, using the translation $\varphi \mapsto \hat{\varphi}$ mentioned in §1, we obviously have $M \equiv_{\boldsymbol{L}^{=}} N \iff M \equiv_{\boldsymbol{L}^{\mathrm{eq}}} N$. Thus, by (2)(a) we have

(13) For $\boldsymbol{L}$-structures M and N, $M \approx_{\boldsymbol{L}} N \implies M \equiv_{\boldsymbol{L}^{=}} N$.

$\boldsymbol{L}$,$\approx$-equivalences can be "normalized" in a certain way, which will be useful for us later.

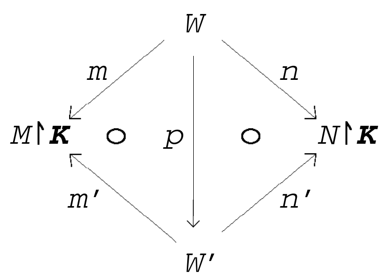
Let $U, V \in \mathrm{Set}^{\boldsymbol{K}}$. A very surjective morphism $f: U \to V$ is *normal* if for any maximal kind K, and any $\vec{a} \in U[K]$, " f is 1-1 in the fiber over $\vec{a}$ ", that is, if $b, c \in UK(\vec{a})$, then $f(b) = f(c)$ implies $b = c$. Together with the very surjective condition, this says that f induces a bijection $UK(\vec{a}) \xrightarrow{\cong} VK(f\vec{a})$.

Let M, N be $\boldsymbol{L}$-structures. A *normal* $\boldsymbol{L}$,$\approx$-equivalence $(W, m, n): M \xleftrightarrow[\boldsymbol{L}]{\approx} N$ is an $\boldsymbol{L}$,$\approx$-equivalence in which both m and n are normal. We have the fact

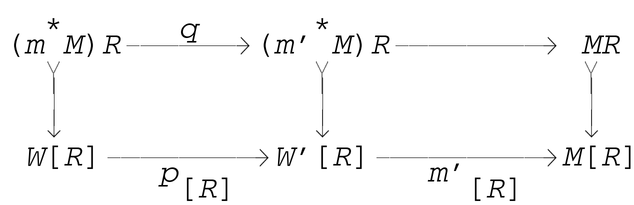
(14) For any $\boldsymbol{L}$-structures M, N, if $M \approx_{\boldsymbol{L}} N$, then there is a normal $\boldsymbol{L}$,$\approx$-equivalence

$(W, m, n) : M \overset{\approx}{\underset{\boldsymbol{L}}{\longleftrightarrow}} N$.

The argument is as follows. Start with any $\boldsymbol{L}$,≈-equivalence $(W, m, n) : M \overset{\approx}{\underset{\boldsymbol{L}}{\longleftrightarrow}} N$. Define $W' \in \mathrm{Set}^{\boldsymbol{K}}$ by setting $W'K = WK$ for all $K \in \boldsymbol{K}$ except the maximal ones; for a maximal K, $W'K \underset{\mathrm{def}}{=} WK/\sim$, where $\sim$ is the equivalence relation on WK for which $b \sim c$ iff b and c are over the same $\vec{a} \in W[K]$, and $m(b) = m(c)$. When in this definition, we replace m by n, the result is the same; this is because (W, m, n) being an $\boldsymbol{L}$,≈-equivalence, $m(b) = m(c)$ iff $n(b) = n(c)$ for b, c over the same element in $W[K]$. For an arrow $p : K \to K_p$, $W'(p) = W(p)$ when K is not maximal (in which case K_p is not maximal either); and for K maximal, $(W'p)(b/\sim) = (Wp)(b)$; the latter is well-defined, since by the definition of $\sim$, if $b \sim c$, then $(Wp)(b) = (Wp)(c)$. Clearly, $W' : \boldsymbol{K} \to \mathrm{Set}$ is well-defined, and we have obvious maps $p : W \to W'$, $m' : W' \to M \restriction \boldsymbol{K}$, $n' : W' \to N \restriction \boldsymbol{K}$ such that



I claim that $(W', m', n') : M \overset{\approx}{\underset{\boldsymbol{L}}{\longleftrightarrow}} N$; the normality condition is clearly satisfied. Consider a relation R in $\boldsymbol{L}$. In the commutative diagram



the outside rectangle and the right-hand square are pullbacks. It follows that the left-hand square is a pullback too. Obviously, $p_{[R]}$ is surjective. It follows that q is surjective. This determines the subobject $(m'^{*}M)R \rightarrowtail W'[R]$ as the image of $(m^{*}M)R \rightarrowtail W[R]$ under $p_{[R]}$. Switching to N from M, $(n'^{*}N)R \rightarrowtail W'[R]$ is the image of $(n^{*}N)R \rightarrowtail W[R]$

under $p_{[R]}$. Since $(m^{*}M)R =_{W[R]} (n^{*}N)R$, it follows that $(m'^{*}M)R =_{W'[R]} (n'^{*}N)R$ as desired. The additional condition concerning equality is clearly satisfied.

Notice that the above proof works for an essentially arbitrary $\boldsymbol{C}$ in place of $\mathtt{Set}$.

Note that if $m:W\to M\restriction\boldsymbol{K}$ is normal, then $m^{*}M$ formed from M as a standard $\boldsymbol{L}^{\mathrm{eq}}$-structure is a standard $\boldsymbol{L}^{\mathrm{eq}}$-structure too. Put in another way, the standard fiberwise equality relations on the maximal kinds in $m^{*}M$ are formed by the same pullback operation from the corresponding relation on M as any primitive $\boldsymbol{L}$-relation.

We have the following variant of (12).

(15) Let $\boldsymbol{C}$ be a small Boolean category, $I:\boldsymbol{L}\to\boldsymbol{C}$. Assume that $\mathcal{X}$ is a finite $\boldsymbol{L}$-context, $\sigma\in\mathrm{S}(I[\mathcal{X}])$, and for all $M, N\models\boldsymbol{C}$ and $\vec{a}\in(M\restriction\boldsymbol{L})[\mathcal{X}]$, $\vec{b}\in(N\restriction\boldsymbol{L})[\mathcal{X}]$,

$$\langle\vec{a}\rangle\in M(\sigma) \ \ \& \ \ (M\restriction\boldsymbol{L}, \vec{a}) \approx_{\boldsymbol{L}} (N\restriction\boldsymbol{L}, \vec{b}) \quad\Longrightarrow\quad \langle\vec{b}\rangle\in N(\sigma) \quad .$$

Then there is an $\boldsymbol{L}$-formula θ in logic with dependent sorts *with* equality with $\mathrm{Var}(\theta)\subset\mathcal{X}$ such that $\sigma =_{I[\mathcal{X}]} I[\mathcal{X}:\theta]$.

Proof. By definition, for each maximal K , $I[\mathrm{E}_K]=I(K)\times_{I[K]}I(K)$. Let us form $I^{\mathrm{eq}}:\boldsymbol{L}^{\mathrm{eq}}\longrightarrow\boldsymbol{C}$ extending $I:\boldsymbol{L}\to\boldsymbol{C}$ by specifying that, $I^{\mathrm{eq}}(\mathrm{E}_K) = I[E_K]$, with $I^{\mathrm{eq}}(\mathrm{e}_{K0})=I^{\mathrm{eq}}(\mathrm{e}_{K1})=1_{I[\mathrm{E}_K]}$. We apply (12) to $I^{\mathrm{eq}}:\boldsymbol{L}^{\mathrm{eq}}\to\boldsymbol{C}$. For $M\models\boldsymbol{C}$, $M\restriction\boldsymbol{L}^{\mathrm{eq}}=M\circ I^{\mathrm{eq}}$ is, clearly, the same as $M\restriction\boldsymbol{L}$ as a standard $\boldsymbol{L}^{\mathrm{eq}}$-structure. Thus,

$$(M\restriction\boldsymbol{L}^{\mathrm{eq}}, \vec{a}) \sim_{\boldsymbol{L}^{\mathrm{eq}}} (N\restriction\boldsymbol{L}^{\mathrm{eq}}, \vec{b}) \quad\Longleftrightarrow\quad (M\restriction\boldsymbol{L}, \vec{a}) \approx_{\boldsymbol{L}} (N\restriction\boldsymbol{L}, \vec{b}) \quad .$$

Thus, from the hypothesis of (15), that of (12) follows. By (12), we have some θ in FOLDS without equality over $\boldsymbol{L}^{\mathrm{eq}}$ such that $\sigma =_{I[\mathcal{X}]} I^{\mathrm{eq}}[\mathcal{X}:\theta]$; but clearly, for θ' in FOLDS with equality over $\boldsymbol{L}$ such that $\hat{\theta}' = \theta$, we have $I[\mathcal{X}:\theta'] = I^{\mathrm{eq}}[\mathcal{X}:\theta]$; thus $\sigma =_{I[\mathcal{X}]} I[\mathcal{X}:\theta']$ as required.