Log. Univers. 14 (2020), 461–479

1661-8297/20/040461-19, *published online* September 21, 2020
https://doi.org/10.1007/s11787-020-00261-2

Logica Universalis

Bases for Structures and Theories II

Jeffrey Ketland

Abstract. In Part I of this paper (Ketland in Logica Universalis 14:357–381, 2020), I assumed we begin with a (relational) signature $P = \{P_i\}$ and the corresponding language L_P, and introduced the following notions: a *definition system* d_Φ for a set of new predicate symbols Q_i, given by a set $\Phi = \{\phi_i\}$ of defining L_P-formulas (these definitions have the form: $\forall \overline{x}(Q_i(x) \leftrightarrow \phi_i)$); a corresponding *translation function* $\tau_\Phi : L_Q \rightarrow L_P$; the corresponding *definitional image operator* D_Φ, applicable to L_P-structures and L_P-theories; and the notion of *definitional equivalence* itself: for structures $A + d_\Phi \equiv B + d_\Theta$; for theories, $T_1 + d_\Phi \equiv T_2 + d_\Theta$. Some results relating these notions were given, ending with two characterizations for definitional equivalence. In this second part, we explain the notion of a *representation basis*. Suppose a set $\Phi = \{\phi_i\}$ of L_P-formulas is given, and $\Theta = \{\theta_i\}$ is a set of L_Q-formulas. Then the original set Φ is called a *representation basis* for an L_P-structure A with inverse Θ iff an inverse explicit definition $\forall \overline{x}(P_i(\overline{x}) \leftrightarrow \theta_i)$ is true in $A + d_\Phi$, for each P_i. Similarly, the set Φ is called a *representation basis* for a L_P-theory T with inverse Θ iff each explicit definition $\forall \overline{x}(P_i(\overline{x}) \leftrightarrow \theta_i)$ is provable in $T + d_\Phi$. Some results about representation bases, the mappings they induce and their relationship with the notion of definitional equivalence are given. In particular, we show that T_1 (in L_P) is definitionally equivalent to T_2 (in L_Q), with respect to Φ and Θ, if and only if Φ is a *representation basis* for T_1 with inverse Θ and $T_2 \equiv D_\Phi T_1$.

Mathematics Subject Classification. Primary 03C07; Secondary 03C95.

Keywords. Definitional equivalence, Theories, Definability.

1. Introduction

Sometimes theories are formulated with different sets of non-logical primitives and yet are definitionally equivalent. There are many examples of theories—often involving formalized systems of arithmetic and set theory—formulated with rather different sets of primitives (aka signatures), which are nonetheless "equivalent".

2. Summary of Part I

In Part I ([4]), we considered a starting language L_P over a relational signature $P = \{P_i\}_{i \in I_P}$, and a set $\Phi = \{\phi_i\}_{i \in I}$ of L_P-formulas. Given Φ, introduce a disjoint set $Q = \{Q_i\}_{i \in I}$ of new relation symbols, with card Q = card Φ, and with the arity of Q_i matching the arity of ϕ_i. The extended language is denoted $L_{P,Q}$ and the language built from the new signature Q (with the implicitly induced arities) is denoted L_Q.[1]

Definition 1. Given $\Phi = \{\phi_i\}$, the *definition system* over Φ, which we write as,

$$d_\Phi$$

is the set of explicit definitions,

$$\forall x_1 \ldots x_{n_i}(Q_i(x_1, \ldots, x_{n_i}) \leftrightarrow \phi_i)$$

where $\{x_1, \ldots x_{n_i}\} = \mathrm{FV}(\phi_i)$, and $n_i = \mathrm{card}\, \mathrm{FV}(\phi_i)$ (the "arity" of ϕ_i). These define the new symbols Q_i in terms of the defining L_P-formulas ϕ_i. We shall sometimes write $\forall \overline{x}(Q_i(\overline{x}) \leftrightarrow \phi_i)$ instead of $\forall x_1 \ldots x_n(Q_i(x_1, \ldots, x_n) \leftrightarrow \phi_i)$.[2]

Definition 2. Let A be an L_P-structure. Then $A + d_\Phi$ is the unique definitional expansion $A^+ \models d_\Phi$ of A. ($A + d_\Phi$ is an $L_{P,Q}$-structure.)

Definition 3. Let T be an L_P-theory. The *the definitional extension* of T wrt Φ is $T + d_\Phi$. We say that T^+ in $L_{P,Q}$ is a *definitional extension* of T in L_P just if

$$T^+ \equiv T + d_\Phi,$$

for some definition system d_Φ, where Φ is some set of L_P-formulas.

Definition 4. Let a definition system d_Φ be given. Define the *translation*, induced by Φ

$$\tau^+_\Phi : L_{P,Q} \to L_P$$

as follows. For symbols P_i, Q_j, variables $x, y, \overline{x}$, and for $L_{P,Q}$-formulas $\alpha, \alpha_1, \alpha_2$:

$$(1) \quad \tau^+_\Phi(P_i(\overline{x})) := P_i(\overline{x})$$
$$(2) \quad \tau^+_\Phi(Q_j(\overline{x})) := (\phi_j)'$$
$$(3) \quad \tau^+_\Phi(x = y) := (x = y)$$
$$(4) \quad \tau^+_\Phi(\neg \alpha) := \neg\, \tau^+_\Phi(\alpha)$$
$$(5) \quad \tau^+_\Phi(\alpha_1 \# \alpha_2) := \tau^+_\Phi(\alpha_1) \;\#\; \tau^+_\Phi(\alpha_2)$$
$$(6) \quad \tau^+_\Phi(\mathbf{q}x\alpha) := \mathbf{q}x\; \tau^+_\Phi(\alpha).$$

---

[1] To simplify notation, I sometimes write "$\Phi = \{\phi_i\}$" to mean $\Phi = \{\phi_i\}_{i \in I}$, omitting the index set. In this case, the expression "$\{\phi_i\}$" does not denote the singleton set containing ϕ_i, but rather the indexed set $\{\phi_i \mid i \in I\}$ of such formulas.

[2] Strictly speaking, there is a distinct sequence $\overline{x}$ for each definition of a Q_i symbol; but it would merely make notation ugly to keep mentioning that. Likewise, d_Θ will be an abbreviation for the set of definitions of the form $\forall \overline{x}(P_i(\overline{x}) \leftrightarrow \theta_i)$, where the θ_i are L_Q-formulas.

where $\#$ is any binary connective, $\mathbf{q}$ is a quantifier and $(\phi_i)'$ is the result of ensuring that the free variables appearing ϕ_i are relabelled, to match those of $Q_i(\overline{x})$.[3] We call τ_Φ^+ the *translation* induced by Φ. It maps from the enriched language $L_{P,Q}$ back to the original language L_P. We let τ_Φ be the *restriction* of τ_Φ^+ to L_Q: thus, τ_Φ maps from the new language L_Q back to the original language L_P. (τ_Φ is also called the *translation* induced by Φ.)

Definition 5. Let $\tau_\Phi : L_Q \to L_P$ and $\tau_\Theta : L_P \to L_Q$ be translations induced by d_Φ and d_Θ. Let T_1 be an L_P theory. Let T_2 be an L_Q theory. Then τ_Θ is an *right inverse* of τ_Φ in T_1 iff, for any $\alpha \in L_P$,

$$T_1 \vdash \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha))$$

We write this more suggestively as:

$$(\tau_\Phi \tau_\Theta = 1)_{T_1}$$

And τ_Θ is an *left inverse* of τ_Φ in T_2 iff, for any $\beta \in L_Q$,

$$T_2 \vdash \beta \leftrightarrow \tau_\Theta(\tau_\Phi(\beta))$$

Likewise, we write this more suggestively as:

$$(\tau_\Theta \tau_\Phi = 1)_{T_2}$$

Definition 6. Let A be an L_P-structure. Then the L_Q-structure $D_\Phi A$ is defined by:

$$D_\Phi A := (A + d_\Phi)\restriction_{L_Q}$$

$D_\Phi A$ is called the *definitional image* of A with respect to Φ.

Definition 7. The *definitional image* of T, with respect to Φ, is the restriction of the deductive closure of $T + d_\Phi$ to the new language L_Q. The definitional image of T with respect to Φ is denoted $D_\Phi T$. That is,

$$D_\Phi T := \text{DedCl}(T + d_\Phi) \restriction_{L_Q} = \{\beta \in L_Q \mid T + d_\Phi \vdash \beta\}$$

Definition 8. Structures A and B are *definitionally equivalent* wrt d_Φ and d_Θ iff

$$A + d_\Phi \cong B + d_\Theta.$$

If this is so, we write:

$$A \overset{\Phi}{\underset{\Theta}{\longleftrightarrow}} B$$

Definition 9. Theories T_1 and T_2 are *definitionally equivalent* wrt d_Φ and d_Θ iff

$$T_1 + d_\Phi \equiv T_2 + d_\Theta.$$

To express this, we write:

$$T_1 \overset{\Phi}{\underset{\Theta}{\longleftrightarrow}} T_2$$

In Part I we established a fair number of "book-keeping lemmas". The three most important results can be summarized:

[3] The clauses (4)–(6) are usually read as saying "τ_Φ^+ commutes with the logical operators".

Lemma 1. *The following hold always:*[4]

$$(1)\ \text{If } A \models T \text{ then } D_\Phi A \models D_\Phi T$$
$$(2)\ \ D_\Phi[Mod(T)] \subseteq Mod(D_\Phi T)$$

Lemma 2. *The following are equivalent:*[5]

$$(1)\ T + d_\Phi \vdash d_\Theta$$
$$(2)\ (\tau_\Theta \tau_\Phi = 1)_T$$
$$(3)\ T \xleftrightarrow[\Theta]{\Phi} D_\Phi T$$

Lemma 3. *The following are equivalent:*[6]

$$(1)\ A + d_\Phi \models d_\Theta$$
$$(2)\ D_\Theta D_\Phi A = A$$
$$(3)\ \ A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$$

In addition to those "book-keeping lemmas", we established two conditions for definitional equivalence, one for structures and one for theories:

Theorem 1. *The following are equivalent:*

$$(1)\qquad A \xleftrightarrow[\Theta]{\Phi} B.$$
$$(2)\ B \cong D_\Phi A \text{ and } A \cong D_\Theta B.$$

Theorem 2. *The following are equivalent:*

$$(1)\qquad T_1 \xleftrightarrow[\Theta]{\Phi} T_2.$$
$$(2)\ (\tau_\Phi \tau_\Theta = 1)_{T_1} \text{ and } T_2 \equiv D_\Phi T_1.$$

3. Definitional Equivalence: Model-Theoretic Criteria

Theorem 2 above establishes a criterion for $T_1 \xleftrightarrow[\Theta]{\Phi} T_2$ in terms of *translation*: $(\tau_\Phi \tau_\Theta = 1)_{T_1}$ and $T_2 \equiv D_\Phi T_1$. Next, we establish *model-theoretic* criteria.

Definition 10. We write

$$(T_1, \Phi) \rightarrowtriangle (T_2, \Theta)$$

to mean:

$$\text{for any } A \models T_1, \text{ there is a } B \models T_2 \text{ st } A + d_\Phi \cong B + d_\Theta.$$

We write:

$$(T_1, \Phi) \leftrightarrowtriangle (T_2, \Theta)$$

to mean: $(T_1, \Phi) \rightarrowtriangle (T_2, \Theta)$ and $(T_2, \Theta) \rightarrowtriangle (T_1, \Phi)$.

[4]These correspond to Lemma 15(2) and 15(3), from Part I, [4].
[5]These correspond to Lemma 10(5), Lemma 11, Lemma 19 from Part I, [4].
[6]These are three conditions from Lemma 16 from Part I, [4].

Lemma 4. *If* $(T_1, \Phi) \multimap (T_2, \Theta)$, *then* $D_\Phi : Mod(T_1) \to Mod(T_2)$ *and, for any* $A \in Mod(T_1)$, *we have* $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$.

Proof. Let $(T_1, \Phi) \multimap (T_2, \Theta)$. So, for any $A \models T_1$, there is a $B \models T_2$ such that $A + d_\Phi \cong B + d_\Theta$.

Consider the operator D_Φ. Let $A \models T_1$. So, there is a $B \models T_2$ such that $A + d_\Phi \cong B + d_\Theta$. So, $B \cong D_\Phi A$. So, $D_\Phi A \models T_2$. So, $D_\Phi : \text{Mod}(T_1) \to \text{Mod}(T_2)$. And since $A + d_\Phi \cong B + d_\Theta$, we have $A + d_\Phi \cong D_\Phi A + d_\Theta$. So, $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$, as required. □

Lemma 5. *If* $D_\Phi : Mod(T_1) \to Mod(T_2)$ *and, for any* $A \in Mod(T_1)$, *we have* $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$, *then* $(T_1, \Phi) \multimap (T_2, \Theta)$.

Proof. Suppose that $D_\Phi : \text{Mod}(T_1) \to \text{Mod}(T_2)$ and, for any $A \in \text{Mod}(T_1)$, we have $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$.

Now suppose $A \models T_1$. We claim there is a $B \models T_2$ st $A + d_\Phi \cong B + d_\Theta$. Since $A \models T_1$, we have $D_\Phi A \models T_2$ and

$$A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$$

And so,

$$A + d_\Phi \cong D_\Phi A + d_\Theta$$

So there is a $B \models T_2$ such that $A + d_\Phi \cong B + d_\Theta$. So, $(T_1, \Phi) \multimap (T_2, \Theta)$, as claimed. □

Lemma 6. *The following are equivalent:*

(1) $$(T_1, \Phi) \multimap (T_2, \Theta).$$

(2) $D_\Phi : Mod(T_1) \to Mod(T_2)$ *and, for any* $A \in Mod(T_1)$, $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$.

Proof. Lemma 4 and Lemma 5. □

Lemma 7. *The following are equivalent:*

$$(1)\ T_1 + d_\Phi \vdash T_2 + d_\Theta.$$
$$(2)\ (T_1, \Phi) \multimap (T_2, \Theta).$$

Proof. (1) ⇒ (2). Assume $T_1 + d_\Phi \vdash T_2 + d_\Theta$. Recall Lemma 3, which implies:

$$(*) \qquad \text{If } A + d_\Phi \models d_\Theta, \text{ then } A + d_\Phi \cong D_\Phi A + d_\Theta.$$

We wish to show $(T_1, \Phi) \multimap (T_2, \Theta)$. Let $A \models T_1$. Thus, $A + d_\Phi \models T_1 + d_\Phi$. Thus, $A + d_\Phi \models T_2 + d_\Theta$. And so, $A + d_\Phi \models d_\Theta$. From $(*)$, it follows that $A + d_\Phi \cong D_\Phi A + d_\Theta$. Since, $A + d_\Phi \models T_2$, we have $D_\Phi A + d_\Theta \models T_2 + d_\Theta$. Thus, $D_\Phi A \models T_2$. And $A + d_\Phi \cong D_\Phi A + d_\Theta$. I.e., $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$. Since this holds in general, $(T_1, \Phi) \multimap (T_2, \Theta)$.

(2) ⇒ (1). We suppose, for any $A \models T_1$, there is some $B \models T_2$ such that $A + d_\Phi \cong B + d_\Theta$. For a contradiction, suppose $T_2 + d_\Theta \vdash \alpha$ and $T_1 + d_\Phi \nvdash \alpha$, for some formula α. This gives us $A + d_\Phi \models T_1 + d_\Phi$, and $A + d_\Phi \not\models \alpha$. Thus,

$A \models T_1$. Hence, there is a model $B \models T_2$ such that $A + d_\Phi \cong B + d_\Theta$. Thus, $B + d_\Theta \models T_2 + d_\Theta$. Since $T_2 + d_\Theta \vdash \alpha$, we have that $B + d_\Theta \models \alpha$. Since, $A + d_\Phi \cong B + d_\Theta$, we get $A + d_\Phi \models \alpha$. Contradiction. □

We then obtain a characterization theorem:

Theorem 3. *The following are equivalent*

$$(1) \qquad T_1 \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} T_2.$$

$$(2)\ (T_1, \Phi) \leftrightarrow (T_2, \Theta).$$

Proof. Reason as follows:

$$\begin{aligned} T_1 \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} T_2 &\Leftrightarrow T_1 + d_\Phi \equiv T_2 + d_\Theta \\ &\Leftrightarrow T_1 + d_\Phi \vdash T_2 + d_\Theta \text{ and } T_2 + d_\Theta \vdash T_1 + d_\Phi \\ &\Leftrightarrow (T_1, \Phi) \rightarrow (T_2, \Theta) \text{ and } (T_2, \Theta) \rightarrow (T_1, \Phi) \\ &\Leftrightarrow (T_1, \Phi) \leftrightarrow (T_2, \Theta). \end{aligned}$$

The relationship $(T_1, \Phi) \leftrightarrow (T_2, \Theta)$ asserts the existence of two functions $F : \mathrm{Mod}(T_1) \to \mathrm{Mod}(T_2)$ and $G : \mathrm{Mod}(T_2) \to \mathrm{Mod}(T_1)$ such that, for any $A \in \mathrm{Mod}(T_1), B \in \mathrm{Mod}(T_2)$, we have

$$F(A) \models T_2 \text{ and } A \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} F(A)$$

$$G(B) \models T_1 \text{ and } G(B) \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} B$$

Together, these imply the existence of a *bijection* (wrt $\cong$):[7]

$$H : \mathrm{Mod}(T_1) \leftrightarrow \mathrm{Mod}(T_2)$$

such that, for any $A \in \mathrm{Mod}(T_1)$, we have:

$$A \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} H(A)$$

To show this, first we prove a simple lemma about definitional equivalence of structures:

Lemma 8. *If* $A \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} B$ *and* $A' \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} B$, *then* $A \cong A'$.

Proof. Let $A + d_\Phi \cong B + d_\Theta$ and $A' + d_\Phi \cong B + d_\Theta$. Thus, $A + d_\Phi \cong A' + d_\Phi$. Thus, by right cancellation, $A \cong A'$. □

Lemma 9. *If* $(T_1, \Phi) \leftrightarrow (T_2, \Theta)$, *there exists a bijection* $H : Mod(T_1) \leftrightarrow Mod(T_2)$ *wrt* $\cong$ *such that, for any* $A \in Mod(T_1)$, $A \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} H(A)$.

[7] Whenever I use the notion of bijection relating classes of models, I mean "bijection up to isomorphism". We write $H : \mathrm{Mod}(T_1) \leftrightarrow \mathrm{Mod}(T_2)$ to mean the map $H : \mathrm{Mod}(T_1) \to \mathrm{Mod}(T_2)$ is a bijection.

Proof. The assumption $(T_1, \Phi) \leftrightarrow (T_2, \Theta)$ asserts the existence of two functions F, G which are linked by the parameters Φ, Θ. We show that G left-inverts F and F left-inverts G.

Let $A \in \text{Mod}(T_1)$. Then $F(A) \in \text{Mod}(T_2)$ and

$$A \xleftrightarrow[\Theta]{\Phi} F(A).$$

Now consider $G(F(A))$. We have $G(F(A)) \in \text{Mod}(T_1)$ and

$$G(F(A)) \xleftrightarrow[\Theta]{\Phi} F(A).$$

So, by the above cancellation lemma, $G(F(A)) \cong A$. Thus G is a left-inverse of F.

We may also show that F is a right inverse of G. For suppose $B \in \text{Mod}(T_2)$. Then $G(B) \in \text{Mod}(T_1)$ and

$$B \xleftrightarrow[\Phi]{\Theta} G(B).$$

Likewise, consider $F(G(B))$. We have $F(G(B)) \in \text{Mod}(T_2)$ and

$$F(G(B)) \xleftrightarrow[\Phi]{\Theta} G(B).$$

So, by the above cancellation lemma again, $F(G(B)) \cong B$. Thus F is a left-inverse of G.

Now if we have $F : X \to Y$ and $G : Y \to X$, and G is a left-inverse of F and F is a left-inverse of G, then $F : X \to Y$ is a bijection. Therefore, $F : \text{Mod}(T_1) \to \text{Mod}(T_2)$ satisfies the conditions stated. □

Lemma 10. *If $(T_1, \Phi) \leftrightarrow (T_2, \Theta)$, then $D_\Phi : Mod(T_1) \to Mod(T_2)$ is a bijection wrt $\cong$ such that, for any $A \in Mod(T_1)$, $A \xleftrightarrow[\Theta]{\Phi} D_\Phi(A)$.*

Proof. Suppose $(T_1, \Phi) \leftrightarrow (T_2, \Theta)$. So, by Lemma 9, there exists a bijection $H : \text{Mod}(T_1) \leftrightarrow \text{Mod}(T_2)$ wrt $\cong$ such that, for any $A \in \text{Mod}(T_1)$, $A \xleftrightarrow[\Theta]{\Phi} H(A)$. Thus, for any $A \in \text{Mod}(T_1)$, we have,

$$\text{(i)} \qquad A + d_\Phi \cong H(A) + d_\Theta$$

We will show that H and D_Φ are the same up to isomorphism: i.e.,

$$H(A) \cong D_\Phi A$$

Since $(T_1, \Phi) \rightarrow (T_2, \Theta)$, by Lemma 4, we know that $D_\Phi : \text{Mod}(T_1) \to \text{Mod}(T_2)$ and, for any $A \in \text{Mod}(T_1)$, we have $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$. So, for any $A \in \text{Mod}(T_1)$, we have

$$\text{(ii)} \qquad A + d_\Phi \cong D_\Phi A + d_\Theta$$

Combining (i) and (ii), we infer,

$$\text{(iii)} \qquad H(A) + d_\Theta \cong D_\Phi A + d_\Theta$$

And by cancellation, $H(A) \cong D_\Phi A$, as required. □

Furthermore, the converse is true.

Lemma 11. *If $D_\Phi : Mod(T_1) \leftrightarrow Mod(T_2)$ such that, for any $A \in Mod(T_1)$, $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$, then $(T_1, \Phi) \Leftrightarrow (T_2, \Theta)$.*

Proof. Let $D_\Phi : \mathrm{Mod}(T_1) \to \mathrm{Mod}(T_2)$ be a bijection such that, for any $A \in \mathrm{Mod}(T_1)$,

$$A + d_\Phi \cong D_\Phi A + d_\Theta.$$

We want to show, first, that for, any $A \models T_1$, there is a $B \models T_2$ st $A + d_\Phi \cong B + d_\Theta$. If $A \models T_1$, then $D_\Phi A$ is such a model. So, $(T_1, \Phi) \Rightarrow (T_2, \Theta)$. We want to show, second, that for, any $B \models T_2$, there is an $A \models T_1$ st $A + d_\Phi \cong B + d_\Theta$. If $B \models T_2$, then $D_\Phi^{-1} B$ is such a model. So, $(T_2, \Theta) \Rightarrow (T_1, \Phi)$. And therefore we have $(T_1, \Phi) \Leftrightarrow (T_2, \Theta)$. □

The two previous lemmas give a second characterization theorem:

Theorem 4. *The following are equivalent:*

(1) $(T_1, \Phi) \Leftrightarrow (T_2, \Theta)$.

(2) $D_\Phi : Mod(T_1) \leftrightarrow Mod(T_2)$ *st for any* $A \in Mod(T_1), A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$.

4. Mutual Definability Does Not Imply Definitional Equivalence

Andréka et al ([1]) show that mutual definability of a pair of theories in each other does not entail their definitional equivalence. First, their notion of model-theoretic definability is explained as follows:

> Let $Th1$ and $Th2$ be theories, maybe on different first-order languages. An explicit definition of $Th1$ over $Th2$ is a conjunction Δ of explicit definitions of the relation symbols of $Th1$ in terms of the language of $Th2$ such that the models of $Th1$ are exactly the reducts of the models of $Th2 \cup \Delta$ (to the language of $Th1$). Thus, we get the models of $Th1$ from those of $Th2$ by first defining the relations of $Th1$ via using Δ, and then forgetting the relations not present in the language of $Th2$. ([1]: 591)

If we switch their $Th1$ to T_2, $Th2$ to T_1, and Δ to d_Φ, this corresponds, in our terminology, to saying that Φ *model-theoretically defines* T_2 in T_1: i.e., $\mathrm{Mod}(T_2) = D_\Phi[\mathrm{Mod}(T_1)]$ (see [4], Definition 28, Part I).

It is relatively straightforward to prove:

Lemma 12. *Let $T_1 \xleftrightarrow[\Theta]{\Phi} T_2$. Then*

$$(1)\ Mod(T_2) = D_\Phi[Mod(T_1)]$$
$$(2)\ Mod(T_1) = D_\Theta[Mod(T_2)].$$

Proof. Let us suppose that $T_1 \xleftrightarrow[\Theta]{\Phi} T_2$. That is, $T_1 + d_\Phi \equiv T_2 + d_\Theta$.

We first wish to prove (1): $\mathrm{Mod}(T_2) = D_\Phi[\mathrm{Mod}(T_1)]$. That is, we wish to prove that, for any L_Q-structure B, we have:

$$B \models T_2 \text{ if and only if } B \cong D_\Phi A, \text{ for some } A \models T_1.$$

Suppose $B \models T_2$. Thus, $B + d_\Theta \models T_2 + d_\Theta$. So, $B + d_\Theta \models T_1 + d_\Phi$. Let $A = D_\Theta B$. Now $B + d_\Theta \models d_\Phi$. So, from Lemma 3 (but relabelling in terms of an L_Q-structure B instead of an L_P-structure A), we have: $B + d_\Theta \cong D_\Theta B + d_\Phi$. I.e., $B + d_\Theta \cong A + d_\Phi$. So, $A + d_\Phi \models T_1 + d_\Phi$. So, $A \models T_1$, as required.

Instead, suppose $B \cong D_\Phi A$, where $A \models T_1$. Thus, $A + d_\Phi \models T_1 + d_\Phi$. And thus, $A + d_\Phi \models T_2 + d_\Theta$. Since $A + d_\Phi \models d_\Theta$, we have, by Lemma 3 again, $A + d_\Phi \cong D_\Theta A + d_\Theta$. Thus, $A + d_\Phi \cong B + d_\Theta$. And so, $B + d_\Theta \models T_2 + d_\Theta$, which implies that $B \models T_2$, as required.

The proof of (2), $\text{Mod}(T_1) = D_\Theta[\text{Mod}(T_2)]$, is entirely analogous, just switching labels. □

On the other hand, the converse of Lemma 12 is not true. Andréka et al 2005 [1] provide a counter-example:

Theorem 5. *(Andréka et al 2005 [1]) There are theories T_1, T_2 and defining sets Φ, Θ such that the following all hold:*

(1) $\quad Mod(T_2) = D_\Phi[Mod(T_1)]$.

(2) $\quad Mod(T_1) = D_\Theta[Mod(T_2)]$.

(3) *not-*$(T_1 \xleftrightarrow[\Theta']{\Phi'} T_2,$ *for any* $\Phi', \Theta')$.

One way to understand the problem is that the translations associated with Φ and Θ are not *mutual inverses*. Below (Theorem 16) we show the "gap" connected to Lemma 12 and Theorem 5 closes, by requiring not only $\text{Mod}(T_2) = D_\Phi[\text{Mod}(T_1)]$ but also that Φ is a *representation basis* for T_1 with inverse Θ: if these conditions hold, then it follows that $T_1 \xleftrightarrow[\Theta]{\Phi} T_2$.[8]

5. Summary

We quickly summarize the results of the three previous sections.

Theorem 6. *The following three claims are equivalent:*

(1) $\quad T_1 + d_\Phi \vdash T_2 + d_\Theta$.

(2) $D_\Phi : Mod(T_1) \to Mod(T_2)$ *st, for any* $A \in Mod(T_1), A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$. (*Lemma* 6)

(3) $\quad (T_1, \Phi) \rightarrow (T_2, \Theta)$. (*Lemma* 7)

Theorem 7. *The following are equivalent:*

(1) $\quad A \xleftrightarrow[\Theta]{\Phi} B$

(2) $\quad A + d_\Phi \cong B + d_\Theta$. (*Definition* 8)

(3) $B \cong D_\Phi A$ *and* $A \cong D_\Theta B$. (*Theorem* 1)

[8]Moreover, the condition "Φ is a *representation basis* for T_1 with inverse Θ" is equivalent to the *mutual invertibility condition* of translations, that $(\tau_\Phi \tau_\Theta = 1)_{T_1}$. See Theorem 10 below (it follows immediately from Lemma 1).

Theorem 8. *The following are equivalent:*[9]

(1) $$T_1 \overset{\Phi}{\underset{\Theta}{\longleftrightarrow}} T_2.$$

(2) $$T_1 + d_\Phi \equiv T_2 + d_\Theta.\ (\textit{Definition } 9)$$

(3) $$(\tau_\Theta \tau_\Phi = 1)_{T_1} \textit{ and } T_2 \equiv D_\Phi T_1.\ (\textit{Theorem } 2)$$

(4) $$(T_1, \Phi) \leftrightarrow\!\!\!\!\!\!\!\leftrightarrow (T_2, \Theta).\ (\textit{Theorem } 3)$$

(5) $D_\Phi : Mod(T_1) \leftrightarrow Mod(T_2)$ *st, for any* $A \in Mod(T_1), A \overset{\Phi}{\underset{\Theta}{\longleftrightarrow}} D_\Phi A$. (*Theorem* 4)

6. Representation Basis

We next move on to defining the notion of "*representation basis*" for a structure and for a theory.

The underlying intuitive concept is fairly simple. Given a structure A in a signature P, we may wish to consider a special set Φ of L_P-formulas, and then examine the "internal structure" defined by them in A: this is what we have called the "definitional image", $D_\Phi A$. What condition should we impose if we wish to reconstruct A from its image $D_\Phi A$?

Clearly, the condition is that the definitional expansion $A + d_\Phi$—which we used to define $D_\Phi A$ prior to forgetting the P-relations—should satisfy an *invertibility condition*, namely that each original P_i be *explicitly definable* from a formula, say θ_i, in the new Q-language. This then means that the original set Φ of L_P-formulas, in some sense, does not "omit" any structural content built into A itself. The formulas ϕ_i merely "encode" that content differently.

That condition is then, simply, that there is a set $\Theta = \{\theta_i\}_{i \in I_P}$ of Q-formulas such that

$$\boxed{A + d_\Phi \models d_\Theta}$$

holds. Or, more explicitly, for each atomic formula $P_i(\overline{x})$ of L_P, we have an explicit inverse definition,

$$A + d_\Phi \models \forall \overline{x}(P_i(\overline{x}) \leftrightarrow \theta_i)$$

for some $\theta_i \in L_Q$.

The same intuition motivates an analogous account for theories. The definitional image $D_\Phi T$ is a smaller theory which "lives inside" the original T (though it is not a sub-theory). It is kind of filtering or projection. But what condition would permit reconstruction of the original?

Again, it is the condition that there is a set $\Theta = \{\theta_i\}_{i \in I_P}$ of Q-formulas such that

$$\boxed{T + d_\Phi \vdash d_\Theta}$$

[9] Aside from the analysis of definitional equivalence in terms of "concept algebras", these are similar to the four characterizations recently given in the lecture notes Andréka & Németi 2014 ([2], p. 40), except that my second condition (Theorem 2) seems to be much stronger than theirs.

holds. Similarly, we can more explicitly express this by saying that, for each atomic formula $P_i(\overline{x})$ of L_P, we have an explicit inverse definition,

$$T + d_\Phi \vdash \forall \overline{x}(P_i(\overline{x}) \leftrightarrow \theta_i)$$

for some $\theta_i \in L_Q$.

As the reader may have noticed above, throughout earlier sections, we have examined what kinds of consequences follow from precisely this invertibility assumption. Based on these informal explanations, we then given the two main definitions:

Definition 11. Φ is a *representation basis* for A with inverse Θ iff $A + d_\Phi \models d_\Theta$.

Definition 12. Φ is a *representation basis* for T with inverse Θ iff $T + d_\Phi \vdash d_\Theta$.

In each case, the formulas θ_i in the set Θ are called *inversion formulas.*

The following two lemmas are entirely straightforward.

Lemma 13. *If T is inconsistent, any Φ is a representation basis for T.*

Lemma 14. *Suppose $T + d_\Phi \vdash d_\Theta$ and $T + d_\Phi \vdash d_\Psi$, for inverses $\Theta = \{\theta_i\}$ and $\Psi = \{\psi_i\}$. Then, for all i, we have: $T + d_\Phi \vdash \theta_i \leftrightarrow \psi_i$.*

For the case where T is the empty theory in L_P (i.e., pure logic), we define the notion of a *logical* representation basis (for L_P):

Definition 13. Φ is a *logical representation basis* for L_P with inverse Θ iff $d_\Phi \vdash d_\Theta$.

It is easy to see that being a representation basis is preserved under theory extension (or structure expansion):

Lemma 15. *If Φ is a representation basis for T (or structure A), then Φ is also a representation basis for every extension of T (resp. every expansion of A). In particular, if Φ is a* logical *representation basis (for L_P), then it is a representation basis for* every *theory T (in L_P).*

The converse of this lemma is not true. Φ might be a representation basis for T, but not a representation basis for a sub-theory or a weaker theory. Similarly, Φ might be a representation basis for A, but not a representation basis for a reduct of A.

We note in passing that while we have defined representation basis *syntactically* in terms of explicit definability of the P_i, a mathematically equivalent *model-theoretic* definition can be given, in terms of implicit definability:

Definition 14. A set Φ of L_P-formulas is a *semantic* representation basis for T just in case, each primitive L_P-symbol P_i is implicitly definable in $T + d_\Phi$.

It then follows, using Beth's definability theorem, that the syntactic definition of representation basis (Definition 12) and the model-theoretic one (Definition 14) are equivalent:

Theorem 9. *Φ is a representation basis for T iff Φ is a semantic representation basis for T.*

7. Examples

We give three examples of this notion in operation and an interesting example involving definitional image. The first three are fairly simple. The fourth provides an example of a logical representation basis Φ for (a propositional language) L_P but where the definitional image of *logic* in L_P under D_Φ is *not logically true*.

Example 1. Given a signature $P = \{P_i\}_{i \in I_P}$, let $\Phi = \{P_i(x_1, \ldots, x_{a(P_i)})\}_{i \in I_P}$, consisting of the atomic formulas of the language L_P. Then Φ is a logical representation basis for L_P, and indeed a representation basis for any theory T in L_P. In this case, each new Q_i is simply defined as P_i. This means they are equivalent in the extension $T + d_\Phi$. And then, trivially, the inversion conditions hold.

The next two examples are slightly modified from examples given by David Miller in his series of papers explaining the language-dependence problem for explications of the concept of truthlikeness (Miller 1974 [5], 1975 [6], 1978 [7]).

Example 2. Let $P = \{p_1, p_2\}$ be a propositional signature and consider the L_P-formulas

$$\phi_1 := p_1$$
$$\phi_2 := p_1 \leftrightarrow p_2$$

Then $\{\phi_1, \phi_2\}$ is a logical representation basis for L_P. For consider the definitions of the new symbols, Q_1 and Q_2:

$$Q_1 \leftrightarrow p_1$$
$$Q_2 \leftrightarrow (p_1 \leftrightarrow p_2)$$

Let d_Φ be $\{Q_1 \leftrightarrow p_1, Q_2 \leftrightarrow (p_1 \leftrightarrow p_2)\}$. Then d_Φ implies:

$$p_1 \leftrightarrow Q_1$$
$$p_2 \leftrightarrow (Q_1 \leftrightarrow Q_2)$$

So, given d_Φ, p_1 and p_2 can be explicitly defined in terms of Q_1 and Q_2 (this happens essentially because $p_1 \leftrightarrow (p_1 \leftrightarrow p_2)$ is logically equivalent to p_2).

Example 3. Let $P = \{P_1, P_2\}$, where P_1, P_2 are unary predicates. Consider the L_P-formulas ϕ_1, ϕ_2:

$$\phi_1 := P_1(x)$$
$$\phi_2 := P_1(x) \leftrightarrow P_2(x)$$

Then the pair $\{\phi_1, \phi_2\}$ is a logical representation basis for L_P. For the explicit definitions,

$$Q_1(x) \leftrightarrow P_1(x)$$
$$Q_2(x) \leftrightarrow (P_1(x) \leftrightarrow P_2(x))$$

imply (in logic alone) the inversions:

$$P_1(x) \leftrightarrow Q_1(x)$$
$$P_2(x) \leftrightarrow (Q_1(x) \leftrightarrow Q_2(x))$$

for the same reason as the previous example.

Example 4. Consider, given Φ, the definitional image of T when T is the *empty theory*. Let Log_P be the set of L_P-sentences which are theorems of logic in L_P. The definitional image Log_P under Φ is given by:

$$D_\Phi \mathsf{Log}_P := \{\beta \in \mathrm{Sent}(L_Q) \mid d_\Phi \vdash \beta\}$$

Then:

Observation. There is a signature P and (logical) representation basis Φ for logic in L_P such that the L_Q theory $D_\Phi \mathsf{Log}_P$ isn't logically true.

For example, let $P = \{p_1, p_2\}$ be a propositional signature and consider the three L_P-formulas:

$$\phi_1 := p_1 \wedge p_2$$
$$\phi_2 := p_1 \wedge \neg p_2$$
$$\phi_3 := \neg p_1 \wedge p_2$$

Then one can show that $\{\phi_1, \phi_2, \phi_3\}$ is a logical representation basis for L_P. Introduce the definition system d_Φ for the new symbols, Q_1, Q_2 and Q_3:

$$Q_1 \leftrightarrow p_1 \wedge p_2$$
$$Q_2 \leftrightarrow p_1 \wedge \neg p_2$$
$$Q_3 \leftrightarrow \neg p_1 \wedge p_2$$

Then inversions—i.e., explicit definitions of p_1, p_2 in terms of the Q_i—can be obtained as follows:

$$d_\Phi \vdash p_1 \leftrightarrow (Q_1 \vee Q_2)$$
$$d_\Phi \vdash p_2 \leftrightarrow (Q_1 \vee Q_3)$$

However, not every L_Q-theorem of d_Φ is logically true. For example,

$$d_\Phi \vdash Q_1 \rightarrow \neg Q_2$$
$$d_\Phi \vdash Q_1 \rightarrow \neg Q_3$$
$$d_\Phi \vdash Q_2 \rightarrow \neg Q_3$$

In each case, we have $\beta \in L_Q$, with $d_\Phi \vdash \beta$. But, for each β, we have $\nvdash \beta$.[10] And thus the theory $D_\Phi \mathsf{Log}_P$ is not logically true. △

[10]Note that $\vdash \tau_\Phi(\beta)$.

8. Basis, Translation and Equivalence

We can now begin assemble the various pieces of this rather complicated jigsaw. In Subsection 8.1, we shall establish equivalent ways of expressing "Φ is a representation basis for T with inverse Θ". Then, in Subsection 8.2, we shall see how imposing this condition leads to strengthened properties of the D_Φ operator. Finally, in Subsection 8.3, we see how to include being a representation basis as a further criterion for expressing "T_1 is definitionally equivalent to T_2, wrt Φ and Θ".

8.1. Criteria for Being a Representation Basis

First, we establish equivalents for Φ being a representation basis for T with inverse Θ.

Theorem 10. *The following conditions are equivalent:*

(1) Φ *is a representation basis for* T *with inverse*Θ.

(2) $T + d_\Phi \vdash d_\Theta$.

(3) $(\tau_\Phi \tau_\Theta = 1)_T$.

(4) $T \xleftrightarrow[\Theta]{\Phi} D_\Phi T$.

(5) *For any* $A \models T, A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$.

(6) *For any* $A \models T, D_\Theta D_\Phi A = A$.

Proof. (1) $\Leftrightarrow$ (2) is simply Definition 12. (2) $\Leftrightarrow$ (3) follows immediately from Lemma 2. Similarly, (2) $\Leftrightarrow$ (4) follows immediately from Lemma 2.

For (2) $\Leftrightarrow$ (5): first, suppose $T + d_\Phi \vdash d_\Theta$. So, by Lemma 2: $T + d_\Phi \equiv D_\Phi T + d_\Theta$. Next, suppose $A \models T$. So, $A + d_\Phi \models d_\Theta$. So, by Lemma 3: $A + d_\Phi \cong D_\Phi A + d_\Theta$. So, $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$, as claimed.

For the converse, suppose $A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$, for any $A \models T$. By Lemma 3, we infer that $A + d_\Phi \models d_\Theta$, for any $A \models T$. Now suppose $B \models T + d_\Phi$. Let $B = A' + d_\Phi$. So, $A' \models T$. And therefore, $A' + d_\Phi \models d_\Theta$. So, $B \models d_\Theta$. Therefore, $T + d_\Phi \vdash d_\Theta$, as claimed.

For (5) $\Leftrightarrow$ (6): notice that, by Lemma 3,

$$D_\Theta D_\Phi A = A \text{ iff } A \xleftrightarrow[\Theta]{\Phi} D_\Phi A$$

□

Next, specializing to the case where T is pure logic, we obtain:

Theorem 11. *The following are equivalent:*

(1) Φ *is a logical representation basis for* L_P *with inverse* Θ.

(2) $$d_\Phi \vdash d_\Theta.$$

(3) $$\vdash \alpha \leftrightarrow \tau_\Phi(\tau_\Theta(\alpha)), \quad \textit{for any } \alpha \in L_P.$$

(4) $$A + d_\Phi \models d_\Theta, \quad \textit{for any } A.$$

(5) $$A \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} D_\Phi A, \quad \textit{for any } A.$$

(6) $$D_\Theta D_\Phi A \cong A, \quad \textit{for any } A.$$

Proof. The equivalences of (1), (2) and (3) follow immediately from Theorem 10 by setting T as the empty theory. The equivalences of (4), (5) and (6) follow immediately from Lemma 3. And the equivalence of (2) and (5) follows from the equivalence of conditions (2) and (5) in Theorem 10, by setting T as the empty theory. □

8.2. Consequences of Φ Being a Representation Basis for T

Theorem 12. *Let* Φ *be a representation basis for* T *with inverse* Θ. *Then:*

(1) $$D_\Phi : Mod(T) \leftrightarrow Mod(D_\Phi T).$$

(2) *If* $A, B \in Mod(T)$ *and* $D_\Phi A \cong D_\Phi B$, *then* $A \cong B$.

(3) *For any* $B \models D_\Phi T$, *there is* $A \models T$ *st* $B \cong D_\Phi A$.

(4) $$A \models T \textit{ iff } D_\Phi A \models D_\Phi T.$$

(5) $$D_\Phi[Mod(T)] = Mod(D_\Phi T).$$

Proof. Let us suppose Φ is a representation basis for T with inverse Θ. In particular, by Theorem 10(4), we have:

$$(*) \quad T \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} D_\Phi T$$

Second, recall Theorem 8. Conditions (1) and (5) tell us that $T_1 \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} T_2$ iff $D_\Phi : \text{Mod}(T_1) \leftrightarrow \text{Mod}(T_2)$ st, for any $A \in \text{Mod}(T_1)$, $A \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} D_\Phi A$. So, we have, using $(*)$:

$$(**) \quad D_\Phi : \text{Mod}(T) \leftrightarrow \text{Mod}(D_\Phi T) \text{ st, for any } A \in \text{Mod}(T), A \underset{\Theta}{\overset{\Phi}{\longleftrightarrow}} D_\Phi A.$$

We can then prove the results quickly.

(1) is a consequence of $(**)$. And (2) and (3) are obvious consequences of (1).

For (4). We already know that if $A \models T$, then $D_\Phi A \models D_\Phi T$ (Lemma 1(1)). So, instead, suppose that $D_\Phi A \models D_\Phi T$. Let $B = D_\Phi A$. It follows from (**) that there is a model $A' \models T$ such that $B \cong D_\Phi A'$. So $B \cong D_\Phi A'$ and $B \cong D_\Phi A$, which implies: $D_\Phi A' \cong D_\Phi A$. And therefore, $A \cong A'$, since D_Φ is injective. So, $A \models T$.

For (5). We already know that $D_\Phi[\text{Mod}(T)] \subseteq \text{Mod}(D_\Phi T)$ (Lemma 1(2)). We wish to prove the converse: $\text{Mod}(D_\Phi T) \subseteq D_\Phi[\text{Mod}(T)]$. So let $B \in \text{Mod}(D_\Phi T)$.

We claim $B \in D_\Phi[\mathrm{Mod}(T)]$. That is, we claim there is some $A \models T$ such that $B \cong D_\Phi A$. But this follows immediately from (3). □

The next two results reveal the sense in which moving between different bases, say $\Phi = \{\phi_i\}$ and $\Phi^* = \{\phi_i^*\}$, is somewhat analogous to moving between bases for a vector space or between co-ordinate system on a manifold. That is, so long as both are representation bases, then they are interdefinable:

Theorem 13. *Let an L_P-structure A be given. Let $\Phi = \{\phi_i\}$ be a representation basis for A defining the Q_i as ϕ_i, and with inverse Θ. Let $\Phi^* = \{\phi_i^*\}$ be a representation basis for A, defining the Q_j^* as ϕ_j^*, and with inverse Θ^*. Then:*

(1) *For each Q_j^*, there is $\alpha \in L_Q$ st $A + d_\Phi + d_{\Phi^*} \models \forall \overline{x}\ (Q_j^*(\overline{x}) \leftrightarrow \alpha)$.*

(2) *For each Q_i, there is $\beta \in L_{Q^*}$ st $A + d_\Phi + d_{\Phi^*} \models \forall \overline{x}\ (Q_i(\overline{x}) \leftrightarrow \beta)$.*

Proof. One has definitions, in the definitional expansion $A + d_\Phi + d_{\Phi^*}$, of the symbols Q_i and the Q_i^* in terms of the P_i symbols, as well as "inverse definitions" of the P_i in terms of the Q_i and the Q_i^*. One can then verify that there are definitions of the Q_i in terms of the Q_i^* and vice versa. □

Theorem 14. *Let an L_P-theory T be given. Let $\Phi = \{\phi_i\}$ be a representation basis for T defining the Q_i as ϕ_i, and with inverse Θ. Let $\Phi^* = \{\phi_i^*\}$ be a representation basis for T, defining the Q_j^* as ϕ_j^*, and with inverse Θ^*. Then:*

(1) *For each Q_j^*, there is $\alpha \in L_Q$ st $T + d_\Phi + d_{\Phi^*} \vdash \forall \overline{x}\ (Q_j^*(\overline{x}) \leftrightarrow \alpha)$.*

(2) *For each Q_i, there is $\beta \in L_{Q^*}$ st $T + d_\Phi + d_{\Phi^*} \vdash \forall \overline{x}\ (Q_i(\overline{x}) \leftrightarrow \beta)$.*

Proof. Analogously to the previous result, one has definitions, in the definitional extension $T + d_\Phi + d_{\Phi^*}$, of the symbols Q_i and the Q_i^* in terms of the P_i symbols, as well as "inverse definitions" of the P_i in terms of the Q_i and the Q_i^*. One can then verify that there are definitions of the Q_i in terms of the Q_i^* and vice versa. □

8.3. Criteria for Definitional Equivalence

We extend the criteria for definitional equivalence (Theorem 8) with a sixth condition, now formulated in terms of "representation basis":

Theorem 15. *The following are equivalent:*

(1) $$T_1 \xleftrightarrow[\Theta]{\Phi} T_2.$$

(2) $$T_1 + d_\Phi \equiv T_2 + d_\Theta.$$

(3) $$(\tau_\Theta \tau_\Phi = 1)_{T_1} \text{ and } T_2 \equiv D_\Phi T_1.$$

(4) $$(T_1, \Phi) \Leftrightarrow (T_2, \Theta).$$

(5) $$D_\Phi : Mod(T_1) \leftrightarrow Mod(T_2) \text{ st, for any } A \models T_1, A \xleftrightarrow[\Theta]{\Phi} D_\Phi A.$$

(6) *Φ is a representation basis for T_1 with inverse Θ and $T_2 \equiv D_\Phi T_1$.*

Proof. The equivalence of the first five criteria is stated in Theorem 8. To establish (3) ⇔ (6), note that, using Theorem 10(1, 3), Φ is a representation basis for T_1 with inverse Θ if and only if $(\tau_\Theta \tau_\Phi = 1)_{T_1}$. □

A (final) corollary is:

Theorem 16. *Suppose* Φ *is a representation basis for* T_1 *with inverse* Θ. *Then*

$$(1) \qquad T_1 \xleftrightarrow[\Theta]{\Phi} T_2.$$
$$(2)\ Mod(T_2) = D_\Phi[Mod(T_1)].$$

are equivalent.

Proof. For (1) $\Rightarrow$ (2), let us suppose first that $T_1 \xleftrightarrow[\Theta]{\Phi} T_2$. Then this already implies that Φ is a representation basis for T_1 with inverse Θ and, furthermore, $T_2 \equiv D_\Phi T_1$. So, by Theorem 12(5), $D_\Phi[\text{Mod}(T_1)] = \text{Mod}(D_\Phi T_1)$. And since $T_2 \equiv D_\Phi T_1$, we conclude that $\text{Mod}(T_2) = D_\Phi[\text{Mod}(T_1)]$.

For (2) $\Rightarrow$ (1), let us suppose instead that Φ is a representation basis for T_1 with inverse Θ and $\text{Mod}(T_2) = D_\Phi[\text{Mod}(T_1)]$. By Theorem 12(5), $D_\Phi[\text{Mod}(T_1)] = \text{Mod}(D_\Phi T_1)$. Hence, $\text{Mod}(T_2) = \text{Mod}(D_\Phi T_1)$. And so, $T_2 \equiv D_\Phi T_1$. And by condition (6) of Theorem 15, we conclude that $T_1 \xleftrightarrow[\Theta]{\Phi} T_2$. □

Thus, if Φ is a representation basis for T_1 with inverse Θ *and* $\text{Mod}(T_2) = D_\Phi[\text{Mod}(T_1)]$, it follows that $T_1 \xleftrightarrow[\Theta]{\Phi} T_2$. However, the counterexample from [1] shows that $\text{Mod}(T_2) = D_\Phi[\text{Mod}(T_1)]$ is too weak for this conclusion: the missing ingredient is that the defining set Φ be a *representation basis* for T_1 (with inverse Θ).

9. Theories and Basis Dependence

There are several theories which have been carefully studied in mathematical logic—usually involving arithmetic and set theory—known to be definitionally equivalent.

But this is not a minor topic of narrow interest only to mathematical logicians. First, philosophers of science have long been interested in what constitutes either the empirical equivalence, or the full equivalence, of scientific theories. In the mid 70s, David Miller introduced the "language dependence" problem for theories of truthlikeness: the core of his argument being that changing which predicates are taken as *primitive* can affect comparisons of truthlikeness for false theories ([5–7]). In metaphysics and epistemology, there is the famous example of "grue" and "bleen" predicates introduced by Nelson Goodman ([3]).

Second, there is an intuitive idea within mathematics and in physics of trying to eliminate dependence on "arbitrary choices", and that one prefers "basis independent" descriptions of mathematical objects. In a recent paper, Albert Visser notes:

> The study of interpretability is, in part, about the escape from the tyranny of signature. Specific choices for the language are implementation artifacts introduced because, after all is said an done, we

> have to do things one way or another. Good mathematical properties of theories should be independent of these arbitrary choices. (Visser 2015 [8], p. 2 of preprint)

An example is the basis of a vector space V. Every vector space has a basis $\{\mathbf{e}_a\}$ and the space itself is reconstructed as the linear span of the basis. But any such basis itself is an "arbitrary parametrization" or "implementation" of V and vectors in V don't "care" what basis they are expanded relative to: given two bases $\{\mathbf{e}_a\}$ and $\{\mathbf{e}'_a\}$, and some $\mathbf{v} \in V$, there are expansions:

$$X^a \mathbf{e}_a = \mathbf{v} = Y^a \mathbf{e}'_a$$

Sometimes one basis is much more convenient to work with than others (e.g., certain matrices get diagonalized; certain operators take simpler forms). But these efficiency features have nothing to do with vector space itself.

Similarly, any particular chart (U, φ) in a topological manifold is, in some sense, arbitrary. For the points in M shouldn't really "care" what co-ordinates they are given in overlapping charts; say, charts (U, φ) and (V, ψ) where $U \cap V \neq \varnothing$. And similarly, in the overlap region, there is an invertible mapping (a local homeomorphism) between the charts. If M is a 3-manifold and I am told that $\varphi(x) = (0, 0, 0)$, then I have been told nothing intrinsic to M itself; since any point $x \in M$ can be given those co-ordinates by some chart φ.

One wonders if structures and theories could be treated similarly. Here we have explained (Sect. 6) what it is for a set $\Phi = \{\phi_i\}_{i \in I}$ of formulas of a language L to form a "representation basis" for an L-structure A or for an L-theory T. The atomic formulas automatically do. But complex logical compounds may also form a representation basis for A or T, so long as the corresponding system of definitions is "invertible". If the set $\Phi = \{\phi_i\}_{i \in I}$ is a *representation basis*, each primitive P_i of the language L can be given an explicit definition of the form $\forall \overline{x}(P_i(\overline{x}) \leftrightarrow \theta_i)$, where θ_i contains only new predicates Q_i introduced via explicit definition to represent the $\phi_i \in \Phi$. These intertranslatable representation bases then provide "definitionally equivalent reparametrizations" of a structure A or a theory T (relative to L).[11]

[11] I would like to acknowledge support from a grant from the National Science Centre in Kraków (NCN) (grant number 2018/29/B/HS1/01832).

References

[1] Andréka, H., Madarász, J.X., Németi, I.: Mutual definability does not imply definitional equivalence, a simple example. Math. Log. Quart. **51**(6), 591–597 (2005)

[2] Andréka, H., Németi, I.: Definability Theory Course Notes. https://old.renyi.hu/pub/algebraic-logic/DefThNotes0828.pdf. (2014)

[3] Goodman, N.: Fact, Fiction, and Forecast. Athlone Press, London (1954)

[4] Ketland, J.: Bases for structures and theories I. Logica Universalis **14**, 357–381 (2020)

[5] Miller, D.W.: On popper's qualitative theory of verisimilitude. Br. J. Philos. Sci. **25**, 166–177 (1974)

[6] Miller, D.W.: The accuracy of predictions. Synthese **30**, 159–191 (1975)

[7] Miller, D.W.: The distance between constituents. Synthese **38**, 197–212 (1978)

[8] Visser, A.: Extension and Interpretability. Preprint, (2015)

Jeffrey Ketland
Institute of Philosophy
University of Warsaw
Krakowskie Przedmieście 3
00-927 Warszawa
Poland
e-mail: jeffreyketland@gmail.com

Received: May 26, 2020.
Accepted: September 13, 2020.