

## ABSTRACT BETH DEFINABILITY IN INSTITUTIONS

MARIUS PETRIA\* AND RĂZVAN DIACONESCU

**Abstract.** This paper studies definability within the theory of institutions, a version of abstract model theory that emerged in computing science studies of software specification and semantics. We generalise the concept of definability to arbitrary logics, formalised as institutions, and we develop three general definability results. One generalises the classical Beth theorem by relying on the interpolation properties of the institution. Another relies on a meta Birkhoff axiomatizability property of the institution and constitutes a source for many new actual definability results, including definability in (fragments of) classical model theory. The third one gives a set of sufficient conditions for ‘borrowing’ definability properties from another institution via an ‘adequate’ encoding between institutions.

The power of our general definability results is illustrated with several applications to (many-sorted) classical model theory and partial algebra, leading for example to definability results for (quasi-)varieties of models or partial algebras. Many other applications are expected for the multitude of logical systems formalised as institutions from computing science and logic.

### §1. Introduction.

**1.1. Institution-independent model theory.** The theory of “institutions” [26] is a categorical abstract model theory which formalises the intuitive notion of logical system, including syntax, semantics, and the satisfaction between them. It provides the most complete form of abstract model theory, the only one including signature morphisms, model reducts, and even mappings (morphisms) between logics as primary concepts. Institution have been recently also extended towards proof theory [40, 19] in the spirit of categorical logic [31].

The concept of institution arose within computing science (algebraic specification) in response to the population explosion among logics in use there, with the ambition of doing as much as possible at a level of abstraction independent of commitment to any particular logic [26, 44, 21]. Besides its extensive use in specification theory (it has become the most fundamental mathematical structure in algebraic specification theory), there have been several substantial developments towards an “institution-independent” (abstract) model theory [48, 49, 14, 16, 15, 25, 24]. A textbook dedicated to this topic is under preparation [20] and [18] is a recent survey.

The significance of institution-independent model theory is manifold. First, it provides model theoretic results and analysis for various logics in a generic way. Apart of reformulation of standard concepts and results in a very general setting, thus applicable to many logical systems, institution-independent model theory has already produced a serie of new significant results in classical model theory [16, 25].

---

Received November 6, 2005.

\*On leave from the Institute of Mathematics of the Romanian Academy.

Then, institution-independent model theory provides a new *top-down* way of doing model theory, making explicit the generality and power of concepts by placing them at the right level of abstraction and thus extracting the essence of the results independently of the largely irrelevant details of the particular logic in use. This leads to a deeper conceptual understanding guided by a structurally clean causality. Concepts come naturally as presumptive features that “a logic” might exhibit or not, hypotheses are kept as general as possible and introduced on a by-need basis, results and proofs are modular and easy to track down despite their sometimes very deep content.

**1.2. Summary and contributions of this work.** In this paper we study the (Beth) definability problem within an abstract institutional framework, and by applying our general results to actual institutions we obtain a series of concrete results (some known, others new) in classical model theory and in partial algebra.

The basis of this approach is given by our novel institution-independent concept of definability for (arbitrary) signature morphisms, which is not only a natural abstraction of the situation when one considers (the definability of) a new symbol, but also generalises the classical concept of definability from inclusive signature morphisms to *any* signature morphism. More explicitly, the classical definability problem of a new (relational) symbol  $\pi$  with respect to a given signature  $\Sigma$ , which determines a signature inclusion  $\Sigma \hookrightarrow \Sigma \cup \{\pi\}$  is generalised and abstracted to any signature morphism  $\varphi: \Sigma \rightarrow \Sigma'$  in any institution. We argue that this is the right concept of definability.

At such level of generality, even the inclusion of explicit definability into the implicit definability is not a trivial problem anymore. We show that in order for this to hold, it is sufficient to impose only a very mild restriction on the signature morphisms, which in the actual (many sorted) situations requires only surjectivity of the sorts mapping.

The core of our paper consists of the study of the other inclusion, of the implicit definability into the explicit definability. In one section we develop a generic Beth theorem generalising the classical one to an institution-independent setting assuming Craig-Robinson interpolation [46, 52, 22], which although in general is stronger than the usual Craig interpolation, is in fact equivalent to the latter when the actual institution has implications and is compact [22].

In another section we develop another definability result which has a complementary range of applications with respect to the definability result via interpolation. This is based on assuming a meta Birkhoff axiomatizability property for the institution rather than Craig-Robinson interpolation, which is formalised by the “Birkhoff institutions” of [16]. It is interesting to notice that our definability result via meta Birkhoff axiomatizability requires rather different conditions than the interpolation result of [16]. This can be seen as a further indication that interpolation cannot be used for this class of definability results and demounts the common view of the causality relation between interpolation and definability. We illustrate the power of our general definability via axiomatizability theorem by developing several applications in (fragments of) classical model theory and partial algebra, most of them new up to our knowledge. These include definability results for various (quasi-)varieties of first order models and partial algebras. Other similar concrete results can be derived for a multitude of other logical systems just by following the same steps as for the above mentioned institutions.

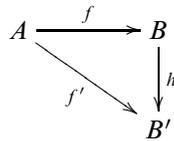
The next section studies a completely different kind of technique, very much in the spirit of institution theory, for establishing definability results. Instead of developing directly a definability result within an actual institution, one may ‘borrow’ it from a simpler, or better understood, institution via an adequate encoding, expressed as institution ‘comorphisms’ [28], of the former into the latter. Here we develop a general ‘borrowing’ definability theorem and illustrate its applicability power with several examples. For example, we can export smoothly the definability property of (full) first order logic to (full) first order partial algebra, and we can also obtain again the definability results for quasi-varieties of partial algebras in an alternative way without having to rely upon a Quasi-Variety Theorem for partial algebras.

Although our paper focuses on definability, it also needs to review a series of institution-independent model theoretic concepts, most of them developed quite recently, such as elementary diagrams [15], internal logic [47, 14], filtered products [14], interpolation [47, 16], Birkhoff institutions [16].

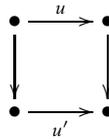
**§2. Institutions.**

**Categories.** We assume the reader is familiar with basic notions and standard notations from category theory; e.g., see [33] for an introduction to this subject. Here we recall very briefly some of them. By way of notation,  $|\mathbb{C}|$  denotes the class of objects of a category  $\mathbb{C}$ ,  $\mathbb{C}(A, B)$  the set of arrows with domain  $A$  and codomain  $B$ , and composition is denoted by “;” and in diagrammatic order. The category of sets (as objects) and functions (as arrows) is denoted by  $Set$ , and  $CAT$  is the category of all categories.<sup>1</sup> The opposite of a category  $\mathbb{C}$  (obtained by reversing the arrows of  $\mathbb{C}$ ) is denoted  $\mathbb{C}^{op}$ .

For any object  $A \in |\mathbb{C}|$ , the *comma category*  $A/\mathbb{C}$  has pairs  $(B, f : A \rightarrow B)$  as objects and  $h \in \mathbb{C}(B, B')$  with  $f;h = f'$  as arrows  $(B, f) \rightarrow (B', f')$ .



A class of arrows  $\mathcal{S} \subseteq \mathbb{C}$  in a category  $\mathbb{C}$  is *stable under pushouts* if for any pushout square in  $\mathbb{C}$



$u' \in \mathcal{S}$  whenever  $u \in \mathcal{S}$ .

**Institutions.** An *institution*  $\mathcal{I} = (\text{Sig}^{\mathcal{I}}, \text{Sen}^{\mathcal{I}}, \text{Mod}^{\mathcal{I}}, \models^{\mathcal{I}})$  consists of

1. a category  $\text{Sig}^{\mathcal{I}}$ , whose objects are called *signatures*,
2. a functor  $\text{Sen}^{\mathcal{I}}: \text{Sig}^{\mathcal{I}} \rightarrow Set$ , giving for each signature a set whose elements are called *sentences* over that signature,

<sup>1</sup>Strictly speaking, this is only a quasi-category living in a higher set-theoretic universe.

3. a functor  $\text{Mod}^{\mathcal{I}}: (\text{Sig}^{\mathcal{I}})^{\text{op}} \rightarrow \text{CAT}$  giving for each signature  $\Sigma$  a category whose objects are called  $\Sigma$ -models, and whose arrows are called  $\Sigma$ -(model) morphisms, and
4. a relation  $\models_{\Sigma}^{\mathcal{I}} \subseteq |\text{Mod}^{\mathcal{I}}(\Sigma)| \times \text{Sen}^{\mathcal{I}}(\Sigma)$  for each  $\Sigma \in |\text{Sig}^{\mathcal{I}}|$ , called  $\Sigma$ -satisfaction,

such that for each morphism  $\varphi: \Sigma \rightarrow \Sigma'$  in  $\text{Sig}^{\mathcal{I}}$ , the *satisfaction condition*

$$M' \models_{\Sigma'}^{\mathcal{I}} \text{Sen}^{\mathcal{I}}(\varphi)(\rho) \text{ iff } \text{Mod}^{\mathcal{I}}(\varphi)(M') \models_{\Sigma}^{\mathcal{I}} \rho$$

holds for each  $M' \in |\text{Mod}^{\mathcal{I}}(\Sigma')|$  and  $\rho \in \text{Sen}^{\mathcal{I}}(\Sigma)$ . We denote the *reduct* functor  $\text{Mod}^{\mathcal{I}}(\sigma)$  by  $\_ \upharpoonright_{\varphi}$  and the sentence translation  $\text{Sen}^{\mathcal{I}}(\varphi)$  by  $\varphi(\_)$ . When  $M = M' \upharpoonright_{\varphi}$  we say that  $M$  is a  $\varphi$ -reduct of  $M'$ , and that  $M'$  is a  $\varphi$ -expansion of  $M$ . When there is no danger of ambiguity, we may skip the superscripts from the notations of the entities of the institution; for example  $\text{Sig}^{\mathcal{I}}$  may be simply denoted  $\text{Sig}$ .

**General assumption:** We assume that all our institutions are such that satisfaction is invariant under model isomorphism, i.e., if  $\Sigma$ -models  $M, M'$  are isomorphic, denoted  $M \cong_{\Sigma} M'$ , then  $M \models_{\Sigma} \rho$  iff  $M' \models_{\Sigma} \rho$  for all  $\Sigma$ -sentences  $\rho$ .

In any institution, a signature morphism  $\varphi: \Sigma \rightarrow \Sigma'$  is *conservative* when each  $\Sigma$ -model has at least one  $\varphi$ -expansion.

An institution is *compact* if for each set of sentences  $E$  and each sentence  $e$ , if  $E \models e$  then there exists a finite subset  $E' \subseteq E$  such that  $E' \models e$ .

**EXAMPLE 2.1.** Let **FOL** be the institution of *many sorted first order logic with equality*. Its signatures  $(S, F, P)$  consist of a set of sort symbols  $S$ , a set  $F$  of function symbols, and a set  $P$  of relation symbols. Each function or relation symbol comes with a string of argument sorts, called *arity*, and for functions symbols, a result sort.  $F_{w \rightarrow s}$  denotes the set of function symbols with arity  $w$  and sort  $s$ , and  $P_w$  the set of relation symbols with arity  $w$ . We assume that each sort has at least one constant (null arity function symbol). Signature morphisms map the three components in a compatible way.

Models  $M$  are first order structures interpreting each sort symbol  $s$  as a set  $M_s$ , each function symbol  $\sigma$  as a function  $M_{\sigma}$  from the product of the interpretations of the argument sorts to the interpretation of the result sort, and each relation symbol  $\pi$  as a subset  $M_{\pi}$  of the product of the interpretations of the argument sorts. Note that each sort interpretation  $M_s$  is non-empty since it contains the interpretation of at least one constant.

Sentences are the usual first order sentences built from equational and relational atoms by iterative application of logical connectives and quantifiers. Sentence translations rename the sorts, function, and relation symbols. For each signature morphism  $\varphi$ , the reduct  $M' \upharpoonright_{\varphi}$  of a model  $M'$  is defined by  $(M' \upharpoonright_{\varphi})_x = M'_{\varphi(x)}$  for each  $x$  sort, function, or relation symbol from the domain signature of  $\varphi$ . The satisfaction of sentences by models is the usual Tarskian satisfaction defined inductively on the structure of the sentences.

The institution **PL** of *propositional logic* can be obtained as the sub-institution of **FOL** by considering only the signatures for which the set of sorts is empty.

A *universal Horn sentence* in **FOL** for a signature  $(S, F, P)$  is a sentence of the form  $(\forall X)H \Rightarrow C$ , where  $H$  is a finite conjunction of (relational or equational) atoms and  $C$  is a (relational or equational) atom, and  $H \Rightarrow C$  is the implication

of  $C$  by  $H$ . The sub-institution **HCL**, *Horn clause logic*, of **FOL** has the same signatures and models as **FOL** but only universal Horn sentences as sentences.

An *algebraic signature*  $(S, F)$  is just a **FOL** signature without relation symbols. The sub-institution of **HCL** which restricts the signatures only to the algebraic ones and the sentences to universally quantified equations is called *equational logic* and is denoted by **EQL**.

The extension of **FOL** allowing conjunctions of sets of sentences is denoted  $\mathbf{FOL}_{\infty, \omega}$ , the extension of **HCL** allowing infinitary conjunctions in the premises  $H$  of the Horn sentences  $(\forall X)H \Rightarrow C$  is denoted  $\mathbf{HCL}_{\infty}$ , the sub-institution of **FOL** with universal disjunctions of atoms as sentences by  $\forall\forall$ , its infinitary extension by  $\forall\forall_{\infty}$ .

**EXAMPLE 2.2.** The institution **PA** of partial algebra [9] is defined as follows.

A *partial algebraic signature* is a tuple  $(S, TF, PF)$ , where  $TF$  is the set of *total* operations and  $PF$  is the set of *partial* operations.

A partial algebra is just like an ordinary algebra but interpreting the operations of  $PF$  as partial rather than total functions. A *partial algebra homomorphism*  $h: A \rightarrow B$  is a family of (total) functions  $\{h_s: A_s \rightarrow B_s\}_{s \in S}$  indexed by the set of sorts  $S$  of the signature such that  $h_w(A_{\sigma}(a)) = B_{\sigma}(h_s(a))$  for each operation  $\sigma: w \rightarrow s$  and each string of arguments  $a \in A_w$  for which  $A_{\sigma}(a)$  is defined.

The sentences have three kinds of atoms: *definedness*  $def(t)$ , *strong equality*  $t \stackrel{s}{=} t'$ , and *existence equality*  $t \stackrel{e}{=} t'$ . The definedness  $def(t)$  of a term  $t$  holds in a partial algebra  $A$  when the interpretation  $A_t$  of  $t$  is defined. The strong equality  $t \stackrel{s}{=} t'$  holds when both terms are undefined or both of them are defined and are equal. The existence equality  $t \stackrel{e}{=} t'$  holds when both terms are defined and are equal.<sup>2</sup> The sentences are formed from these atoms by logical connectives and quantification over total variables.

A (*universal*) *quasi-existence equation* [9] is an infinitary Horn sentence in the infinitary extension  $\mathbf{PA}_{\infty, \omega}$  of **PA** of the form

$$(\forall X) \bigwedge_{i \in I} (t_i \stackrel{e}{=} t'_i) \Rightarrow (t \stackrel{e}{=} t').$$

Let  $QE(\mathbf{PA})$  be the sub-institution of  $\mathbf{PA}_{\infty, \omega}$  which restricts the sentences only to quasi-existence equations,  $QE_1(\mathbf{PA})$  the institution of the quasi-existence equations that have *either*  $t$  or  $t'$  ‘already defined’,<sup>3</sup> and  $QE_2(\mathbf{PA})$  institution of the quasi-existence equations that have *both*  $t$  and  $t'$  ‘already defined’, and let  $QE_k^{\omega}(\mathbf{PA}) = \mathbf{PA} \cap QE_k(\mathbf{PA})$  be their finitary versions.

**NOTATION 2.3** (Classes of signature morphisms). A **FOL** (or **PA**) signature morphism is an  $(xyz)$ -*morphism*, with  $x, y, z \in \{i, s, b, *\}$  (where  $i$  stands for ‘injective’,  $s$  for ‘surjective’,  $b$  for ‘bijective’, and  $*$  for ‘all’) when the sort component has the property  $x$ , the operation (total operation) component has the property  $y$ , and the relation (partial operation) component has the property  $z$ .

<sup>2</sup>Notice that  $def(t)$  is equivalent to  $t \stackrel{e}{=} t$  and that  $t \stackrel{s}{=} t'$  is equivalent to  $(t \stackrel{e}{=} t') \vee (\neg def(t) \wedge \neg def(t'))$ .

<sup>3</sup>They occur as subterms of the terms of the equations in the premise or are formed only from total operation symbols.

For example, a  $(ss^*)$ -morphism of signatures in **FOL** is surjective on the sorts and on the operations, while a  $(bis)$ -morphism of signatures in **PA** is bijective on the sorts, is injective on the total operations, and is surjective on the partial operations.

A brief random list of examples of institutions in use in computing science may also include rewriting [36], higher-order [7], polymorphic [45], temporal [23], process [23], behavioural [5], coalgebraic [12], object-oriented [27], and multi-algebraic (for non-determinism) [32] logics.

**Theories.** For any signature  $\Sigma$  in an institution  $\mathcal{I}$ , a  $\Sigma$ -theory is any set of  $\Sigma$ -sentences.

- For each  $\Sigma$ -theory  $E$ , let  $E^* = \{M \in \text{Mod}(\Sigma) \mid M \models_{\Sigma} e \text{ for each } e \in E\}$ , and
- For each class  $\mathbb{M}$  of  $\Sigma$ -models, let  $\mathbb{M}^* = \{e \in \text{Sen}(\Sigma) \mid M \models_{\Sigma} e \text{ for each } M \in \mathbb{M}\}$ .

If  $E$  and  $E'$  are theories of the same signature, then  $E' \subseteq E^{**}$  is denoted by  $E \models E'$ .

Two sentences,  $\rho_1$  and  $\rho_2$  of the same signature are *semantically equivalent*, denoted  $\models$  when  $\rho_1 \models \rho_2$  and  $\rho_2 \models \rho_1$ . Two models,  $M_1$  and  $M_2$  of the same signature are *elementarily equivalent*, denoted  $M_1 \equiv M_2$ , when they satisfy the same sentences, i.e.,  $\{M_1\}^* = \{M_2\}^*$ . A class  $\mathbb{M}$  of models (of the same signature) is *elementary* when  $\mathbb{M} = \mathbb{M}^{**}$ .

A *theory morphism*  $\varphi: (\Sigma, E) \rightarrow (\Sigma', E')$  is just a signature morphism  $\varphi: \Sigma \rightarrow \Sigma'$  such that  $E' \models \varphi(E)$ . The *institution*  $\mathcal{I}^T$  of  $\mathcal{I}$ -theories has the category of theories  $\text{Th}^{\mathcal{I}}$  of  $\mathcal{I}$  as its category of signatures,  $\text{Sen}^{\mathcal{I}^T}(\Sigma, E) = \text{Sen}^{\mathcal{I}}(\Sigma)$ , and  $\text{Mod}^{\mathcal{I}^T}(\Sigma, E)$  is the full subcategory of  $\text{Mod}^{\mathcal{I}}$  consisting of the  $\Sigma$ -models satisfying  $E$ .

The rest of this section is devoted to a brief presentation of two of the most used properties in institution-independent model theory, namely model amalgamation and elementary diagrams.

**Model amalgamation.** Exactness properties for institutions formalise the possibility of amalgamating models of different signatures when they are consistent on some kind of ‘intersection’ of the signatures (formalised as a pushout square). An institution  $\mathcal{I}$  is *exact* if and only if the model functor  $\text{Mod}^{\mathcal{I}}: (\text{Sig}^{\mathcal{I}})^{\text{op}} \rightarrow \text{CAT}$  preserves finite limits. The institution is *semi-exact* if and only if  $\text{Mod}^{\mathcal{I}}$  preserves pullbacks.

Semi-exactness is everywhere. Virtually all institutions formalising conventional or non-conventional logics are at least semi-exact. In general the institutions of many-sorted logics are exact, while those of unsorted (or one-sorted) logics are only semi-exact [21]. However, in applications the important amalgamation property is the semi-exactness rather than the full exactness. Moreover, in practice often the weak<sup>4</sup> version of exactness suffices [13, 51, 39].

The following amalgamation property is a direct consequence of semi-exactness. The commuting square of signature morphisms

$$\begin{array}{ccc}
 \Sigma & \xrightarrow{\varphi_1} & \Sigma_1 \\
 \varphi_2 \downarrow & & \downarrow \theta_1 \\
 \Sigma_2 & \xrightarrow{\theta_2} & \Sigma'
 \end{array}$$

<sup>4</sup>In the sense of ‘weak’ universal properties [33] not requiring uniqueness.

is an *amalgamation square* if and only if for each  $\Sigma_1$ -model  $M_1$  and a  $\Sigma_2$ -model  $M_2$  such that  $M_1 \upharpoonright_{\varphi_1} = M_2 \upharpoonright_{\varphi_2}$ , there exists a unique  $\Sigma'$ -model  $M'$ , denoted  $M_1 \otimes M_2$ , such that  $M' \upharpoonright_{\theta_1} = M_1$  and  $M' \upharpoonright_{\theta_2} = M_2$ . We can notice easily that in a semi-exact institution each pushout square of signature morphisms is an amalgamation square.

**The method of diagrams.** The method of diagrams is one of the most important conventional model theoretic methods. At the level of institution-independent model theory, cf. [15] this is reflected as a categorical property which formalises the idea that the class of model morphisms from a model  $M$  can be represented (by a natural isomorphism) as a class of models of a theory in a signature extending the original signature with syntactic entities determined by  $M$ . Elementary diagrams can be seen as a coherence property between the semantic structure and the syntactic structure of an institution. By following the basic principle that a structure is defined by its homomorphisms, the semantical structure of an institution is given by its model morphisms. On the other hand the syntactical structure of an institution is essentially determined by its atomic sentences.

An institution  $\mathcal{I}$  has *elementary diagrams* [15] iff for each signature  $\Sigma$  and each  $\Sigma$ -model  $M$ , there exists a signature morphism  $\iota_\Sigma(M): \Sigma \rightarrow \Sigma_M$ , “functorial” in  $\Sigma$  and  $M$ , and a set  $E_M$  of  $\Sigma_M$ -sentences such that  $\text{Mod}(\Sigma_M, E_M)$  and the comma category  $M/\text{Mod}(\Sigma)$  are naturally isomorphic, i.e., the following diagram commutes by the isomorphism  $i_{\Sigma, M}$  “natural” in  $\Sigma$  and  $M$

$$\begin{array}{ccc} \text{Mod}(\Sigma_M, E_M) & \xrightarrow{i_{\Sigma, M}} & (M/\text{Mod}(\Sigma)) \\ & \searrow \text{Mod}(\iota_\Sigma(M)) & \downarrow \text{forgetful} \\ & & \text{Mod}(\Sigma) \end{array}$$

The signature morphism  $\iota_\Sigma(M): \Sigma \rightarrow \Sigma_M$  is called the *elementary extension of  $\Sigma$  via  $M$*  and the set  $E_M$  of  $\Sigma_M$ -sentences is called the *elementary diagram* of the model  $M$ . Note that  $i_{\Sigma, M}^{-1}(1_M)$  is the initial model of  $(\Sigma_M, E_M)$ , which we denote as  $M_M$ .

It is also easy to notice that for a given system of elementary extensions, the canonical isomorphisms  $i_{\Sigma, M}$  imply that the deductive closure  $E_M^{**}$  of the elementary diagrams  $E_M$  are unique.

**EXAMPLE 2.4.** The standard system of diagrams for **FOL** is defined as follows. For any  $(S, F, P)$ -model  $M$ , let  $(F_M)_{\rightarrow s} = F_{\rightarrow s} \cup M_s$ , otherwise let  $(F_M)_{w \rightarrow s} = F_{w \rightarrow s}$ , and let  $M_M$  be the  $(S, F_M, P)$ -expansion of  $M$  such that  $M_m = m$  for each  $m \in M$ . Then  $E_M$  is the set of all (relational or equational) atoms satisfied by  $M_M$ .

However, by varying the concept of model homomorphism one may also get other elementary diagrams for the corresponding sub-institutions of **FOL**. For example, when one restricts model homomorphisms to injective ones,  $E_M$  consists of all atoms and negations of atomic equations satisfied by  $M_M$ , when one restricts them to the closed ones (a  $(S, F, P)$ -model homomorphism  $h: M \rightarrow N$  is *closed* if  $M_\pi = h^{-1}(N_\pi)$  for each  $\pi \in P$ ),  $E_M$  consists of all atoms and negations of atomic relations satisfied by  $M_M$ , and when one restricts them to closed injective model homomorphisms,  $E_M$  consists of all atoms and all negations of atoms satisfied by  $M_M$ .

In similar ways, many institutions either from conventional logic or from computing science, have elementary diagrams [15, 20].

EXAMPLE 2.5. The standard elementary diagrams of the institution  $\mathbf{PA}$  of partial algebras is defined such that given a partial algebra  $A$ , the elementary extension  $\iota(A)$  of its signature via  $A$  adds its elements as *total* constants and the elementary diagram  $E_A$  of  $A$  consists of all existence equations satisfied by  $A_A$ , where  $A_A$  is the  $\iota(A)$ -expansion of  $A$  interpreting each of its elements by itself. Notice that  $\mathbf{PA}$ ,  $QE(\mathbf{PA})$  and  $QE_1(\mathbf{PA})$  admit the same elementary diagrams, but these elementary diagrams are not  $QE_2(\mathbf{PA})$ -sentences.

The institution-independent concept of elementary diagrams presented above has been successfully used in a rather crucial way for developing several results in institution-independent model theory, including (quasi-)variety theorems and existence of free models for theories [15, 20], Robinson consistency and Craig interpolation [25], Tarski elementary chain theorem [24], existence of (co)limits of theory models [15], etc., while a quite different institution-independent version of the method of diagrams has been used for developing quasi-variety theorems and existence of free models within the framework of the so-called ‘abstract algebraic institutions’ [48, 49].

**§3. Abstract Beth definability.** The classical definability problem in model theory can be formulated as follows (see [11, 30]): for any  $\mathbf{FOL}$ -signature  $(S, F, P)$ , a new relation symbol  $\pi$  is ‘implicitly’ defined by a theory  $E$  if and only if it is ‘explicitly’ defined by the same theory.  $\pi$  is implicitly defined when the forgetful reduct  $\text{Mod}^{\mathbf{FOL}}((S, F, P \uplus \{\pi\}), E) \rightarrow \text{Mod}^{\mathbf{FOL}}(S, F, P)$  is injective, which in this case can be formulated in a more syntactic but equivalent way as

$$E \cup E[\pi/\pi'] \models_{(S, F, P \uplus \{\pi, \pi'\})} (\forall X)(\pi(X) \Leftrightarrow \pi'(X))$$

for any other new relation symbol  $\pi'$  of the same arity and where  $E[\pi/\pi']$  is the copy of  $E$  in which  $\pi$  is replaced by  $\pi'$ , while  $\pi$  is explicitly defined if  $\pi$  can be ‘defined’ by an  $(S, F \uplus X, P)$ -sentence  $E_\pi$ , i.e.,

$$E \models_{(S, F, P \uplus \{\pi\})} (\forall X)(\pi(X) \Leftrightarrow E_\pi)$$

where  $X$  a string of variables matching the arity of  $\pi$ .

Definability problem can be naturally formulated at the level of abstraction of arbitrary institutions by abstracting signature inclusions  $(S, F, P) \hookrightarrow (S, F, P \uplus \{\pi\})$  to arbitrary signature morphisms. However the formulation of explicit definability needs a little bit of preparation concerning the ‘internal logic’ of an institution [14, 47].

For any signature  $\Sigma$  in an arbitrary institution, for any  $\Sigma$ -sentences  $\rho_1$  and  $\rho_2$ , a  $\Sigma$ -model  $M$  satisfies  $\rho_1 \Leftrightarrow \rho_2$ , denoted  $M \models \rho_1 \Leftrightarrow \rho_2$ , when  $M \models \rho_1$  if and only if  $M \models \rho_2$ . Similarly, one may easily define other ‘internal logical connectives’ such as conjunction, disjunction, negation, implication, falsum, etc.

For any signature morphism  $\chi: \Sigma \rightarrow \Sigma'$  in an arbitrary institution, for any  $\Sigma'$ -sentence  $\rho$  and any  $\Sigma$ -model  $M$ , we say that  $M$  satisfies  $(\forall \chi)\rho$ , denoted by  $M \models (\forall \chi)\rho$ , if and only if each  $\chi$ -expansion of  $M$  satisfies  $\rho$  in the institution. The institution has *universal  $\mathcal{D}$ -quantification* for a class  $\mathcal{D}$  of signature morphisms,

when for each  $(\chi: \Sigma \rightarrow \Sigma') \in \mathcal{D}$  and each  $\rho \in \text{Sen}(\Sigma')$  there exists a  $\Sigma$ -sentence semantically equivalent to  $(\forall \chi)\rho$ .<sup>5</sup> Notice that the concept of ‘internal quantification’ captures ordinary quantification of the actual institutions, for example **FOL** has  $\mathcal{D}$ -quantification for  $\mathcal{D}$  the class of signature extensions with a finite number of constants, while in the case of second order logic  $\mathcal{D}$  is the class of signature (finite) extensions with any relation and any operation symbols.

It is important to notice that one may use such ‘internal sentences’ in a pure model-theoretic meaning even if they do not correspond to actual sentences of the institution.

**DEFINITION 3.1.** Let  $\varphi: \Sigma \rightarrow \Sigma'$  be a signature morphism and  $E'$  be a  $\Sigma'$ -theory. Then  $\varphi$

- is *defined implicitly* by  $E'$  if the reduct functor  $\text{Mod}(\Sigma', E') \rightarrow \text{Mod}(\Sigma)$  is injective, and
- is *defined (finitely) explicitly* by  $E'$  if for each signature morphism  $\theta: \Sigma \rightarrow \Sigma_1$ , and each sentence  $\rho \in \text{Sen}(\Sigma')$ , there exists a (finite) set of sentences  $E_\rho \subseteq \text{Sen}(\Sigma_1)$  such that

$$E' \models_{\Sigma'} (\forall \theta')(\rho \Leftrightarrow \varphi_1(E_\rho))$$

where

$$\begin{array}{ccc} \Sigma & \xrightarrow{\varphi} & \Sigma' \\ \theta \downarrow & & \downarrow \theta' \\ \Sigma_1 & \xrightarrow{\varphi_1} & \Sigma'_1 \end{array}$$

is any pushout square of the span  $\Sigma_1 \xleftarrow{\theta} \Sigma \xrightarrow{\varphi} \Sigma'$  of signature morphisms.

**REMARK 3.2.** Note that  $E_\rho$  is a (finite) *set* of sentences rather than a single sentence as in the classical formulations of definability. Although the ‘set of sentences’ and ‘the single sentence’ formulations coincide when the institution has conjunctions, only the former gets the right concept of definability for institutions without conjunctions, such as **EQL**, **HCL**, etc. This situation is very similar to that of interpolation, where the concept of interpolant which is meaningful for institutions not necessarily having conjunctions is given by a set of sentences rather than by a single sentence [43, 21, 16]; see also the definition of institution-independent interpolation presented below and the discussion after.

One may define the concept of explicit definability such that the quantification involved is admitted by the institution by requiring  $\theta$  to belong to a class  $\mathcal{D}$  of signature morphisms stable under pushouts such that the institution has universal  $\mathcal{D}$ -quantification. Because such condition would not affect the results of our paper, for the simplicity of presentation we prefer the unrestricted version of the explicit definability with  $\theta$  *any* signature morphism.

**REMARK 3.3.** In actual institutions, it is common to have atomic sentences corresponding to (some) symbols in signatures. For example, in **FOL** for each relation

<sup>5</sup>Existential quantification is defined similarly.

symbol  $\pi$  we have the atom  $\pi(X)$ . Similarly, in **PA** for each partial operation symbol  $\sigma$ , we have the atom  $def(\sigma(X))$ . This means that explicit definability ensures a *uniform* elimination of the symbol  $\pi$  from the sentences. Although this uniformity cannot be expected at the level of Definition 3.1, it can be established easily in the concrete applications on the basis of such correspondences between symbols of signatures and atomic sentences.

One of the most important aspects of definability theory is to establish the relationship between the implicit and the explicit definability. Although in classical model theory and in most of the actual institutions, explicit definability implies very easily the implicit definability, the abstract model theoretic framework shows this is in fact a conditioned property holding for the signature morphisms satisfying a certain condition which can be formulated by relying upon model amalgamation and elementary diagrams.

**DEFINITION 3.4.** In any semi-exact institution with elementary diagrams  $\iota$ , a signature morphism  $\varphi: \Sigma \rightarrow \Sigma'$  is  $\iota$ -tight when for all  $\Sigma'$ -models  $M'$  and  $N'$  with a common  $\varphi$ -reduct,  $M' \otimes M_M \equiv N' \otimes N_N$  implies  $M' = N'$  (where  $M = M' \upharpoonright_{\varphi} = N' \upharpoonright_{\varphi} = N$ ).

$$\begin{array}{ccc}
 \Sigma & \xrightarrow{\varphi} & \Sigma' \\
 \downarrow \iota_{\Sigma}(M) & & \downarrow \theta' \\
 \Sigma_M & \xrightarrow{\varphi_1} & \Sigma'_1
 \end{array}$$

**EXAMPLE 3.5.** Consider the classical situation when  $\varphi$  is a signature morphism in **FOOL** adding one relation symbol  $\pi$ . Then the only possible difference between  $M'$  and  $N'$  could only be found in the difference between  $M'_\pi$  and  $N'_\pi$ . But  $M'_\pi = \{X \mid M' \otimes M_M \models \pi(X)\} = \{X \mid N' \otimes N_N \models \pi(X)\} = N'_\pi$ .

**REMARK 3.6.** The situation of the above example is quite symptomatic for most of the actual institutions.  $M' \otimes M_M$  is just the expansion of  $M'$  interpreting the elements of  $M$  by themselves. Therefore  $M' \otimes M_M \equiv N' \otimes N_N$  implies that each atom in the extended signature is satisfied either by none or by both models, which means that each symbol newly added by  $\varphi$  gets the same interpretation in  $M'$  and  $N'$ . This argument holds in all actual institutions in which models interpret the symbols of the signatures as sets and functions, such institutions can be formalised by the so-called *concrete institutions* of [6, 38].

The following helps to characterise the tight signature morphisms in the actual institutions.

**FACT 3.7.** Let  $\varphi: \Sigma \rightarrow \Sigma'$  be a  $\iota$ -tight signature morphism in a semi-exact institution with elementary diagrams  $\iota$ . Then any two  $\Sigma'$ -models which are isomorphic by a  $\varphi$ -expansion of an identity, are equal.

**PROOF.** Let  $h: M' \rightarrow N'$  be a  $\Sigma'$ -isomorphism such that  $h \upharpoonright_{\varphi}$  is identity. Let  $M = M' \upharpoonright_{\varphi}$  and  $N = N' \upharpoonright_{\varphi}$ . For the diagram of Definition 3.4 consider the amalgamation  $h \otimes 1_{M_M}$ ; this is also an isomorphism. Therefore  $M' \otimes M_M$  and  $N' \otimes N_N$  are isomorphic, hence they are elementarily equivalent. By the definition of  $\varphi$  being tight, we get that  $M' = N'$ .  $\dashv$

**COROLLARY 3.8.** In **FOL** and **PA** (considered with the standard systems of elementary diagrams  $\iota$ ), a signature morphism is  $\iota$ -tight if and only if it is an  $(s^{**})$ -morphism.

**PROOF.** The surjectivity on the sorts is necessary because otherwise, given a  $\Sigma'$ -model  $M'$  we may consider another  $\Sigma'$ -model  $N'$  which is like  $M'$  but interprets the sorts outside image of the tight signature morphism  $\varphi: \Sigma \rightarrow \Sigma'$  differently but isomorphically to  $M'$ . This gives a  $\Sigma'$ -isomorphism expanding a  $\Sigma$ -identity between different  $\Sigma'$ -models, thus contradicting Fact 3.7.

The surjectivity on the sorts is also sufficient. We treat here only the case of **FOL**, since **PA** may get a similar treatment. Consider the diagram of Definition 3.4. If  $M' \otimes M_M \equiv N' \otimes N_N$  for  $\Sigma'$ -models  $M', N'$  with  $(M =)M' \upharpoonright_\varphi = N' \upharpoonright_\varphi (= N)$ , then for all operation symbols  $(\sigma: w \rightarrow s) \in \Sigma'$  and all  $a \in M'_w = N'_w$  and  $b \in M'_s = N'_s$ , we have that  $M' \otimes M_M \models \sigma(a) = b$  iff  $N' \otimes N_N \models \sigma(a) = b$ . This means that  $M'_\sigma = N'_\sigma$ . This argument can be extended to relation symbols too.  $\dashv$

**PROPOSITION 3.9.** In any semi-exact institution with elementary diagrams  $\iota$ , each  $\iota$ -tight signature morphism is defined implicitly whenever it is defined explicitly.

**PROOF.** Let  $\varphi: \Sigma \rightarrow \Sigma'$  be a tight signature morphism which is explicitly defined by  $E' \subseteq \text{Sen}(\Sigma')$ . We show that  $\varphi$  is defined implicitly by  $E'$ . Let  $M', N' \in |\text{Mod}(\Sigma', E')|$  with  $M' \upharpoonright_\varphi = N' \upharpoonright_\varphi$ .

It suffices to show that  $M' \otimes M_M$  is elementarily equivalent to  $N' \otimes N_N$ , where  $M = M' \upharpoonright_\varphi = N' \upharpoonright_\varphi = N$ .

Let  $M' \otimes M_M \models \rho$ . Because  $\varphi$  is explicitly defined by  $E'$ , there exists  $E_\rho \subseteq \text{Sen}(\Sigma_M)$  such that  $E' \models (\forall \theta')(\varphi_1(E_\rho) \Leftrightarrow \rho)$ . Therefore  $M' \models E'$  implies  $M' \models (\forall \theta')(\varphi_1(E_\rho) \Leftrightarrow \rho)$ . Because  $M' \otimes M_M$  is a  $\theta'$ -expansion of  $M'$ , we get that  $M' \otimes M_M \models \varphi_1(E_\rho) \Leftrightarrow \rho$ , which means that  $M' \otimes M_M \models \varphi_1(E_\rho)$ . By the Satisfaction Condition applied successively in both directions we get that  $N_N = M_M \models E_\rho$  and that  $N' \otimes N_N \models \varphi_1(E_\rho)$ . But  $N' \models E'$  implies  $N' \models (\forall \theta')(\varphi_1(E_\rho) \Leftrightarrow \rho)$ , which further implies that  $N' \otimes N_N \models \varphi_1(E_\rho) \Leftrightarrow \rho$ . Since we have already shown that  $N' \otimes N_N \models \varphi_1(E_\rho)$ , we deduce that  $N' \otimes N_N \models \rho$ .

Because in this case the choice between  $M'$  and  $N'$  is immaterial, we have that  $M' \otimes M_M \equiv N' \otimes N_N$ .  $\dashv$

**REMARK 3.10.** Notice that our usage of elementary diagrams here does involve only the elementary extensions  $I_\Sigma(M): \Sigma \rightarrow \Sigma_M$  and the existence of  $M_M$  as a ‘canonical’  $I_\Sigma(M)$ -expansion of  $M$ . This is weaker than the full requirement of existence of elementary diagrams and can be fulfilled by institutions with a rather poor sentence functor, such as  $QE_2(\mathbf{PA})$  for example. However, the sentence functor should be rich enough in order to allow the existence of tight signature morphisms. For example, in an institution with an empty sentence functor, any signature morphism is explicitly defined (by the empty set of sentences) but not necessarily implicitly defined.

Therefore by means of the above Proposition 3.9 one can easily establish in the actual institutions that the implicit definability contains the explicit definability. The real definability problem is thus given by the reverse implication, which constitutes the topic of the rest of our paper.

DEFINITION 3.11. A signature morphism  $\varphi$  has the (*finite*) *definability property* iff a theory defines  $\varphi$  (finitely) explicitly whenever it defines  $\varphi$  implicitly.

Before focusing on various methods for obtaining the definability property, let us give without proof<sup>6</sup> some structural properties of definability:

- PROPOSITION 3.12. 1. In any institution the classes of signature morphisms which are defined implicitly/explicitly form a category.  
 2. Moreover, if the institution is semi-exact, these classes of signature morphisms are also stable under pushouts.  
 3. In any semi-exact institution with universal  $\mathcal{D}$ -quantification for a class  $\mathcal{D}$  of signature morphisms which is stable under pushouts, for any pushout square of signature morphisms

$$\begin{array}{ccc} \Sigma & \xrightarrow{\varphi} & \Sigma' \\ \theta \downarrow & & \downarrow \theta' \\ \Sigma_1 & \xrightarrow{\varphi_1} & \Sigma'_1 \end{array}$$

such that  $\theta \in \mathcal{D}$  and is conservative,  $\varphi$  has the definability property with respect to  $E'$  whenever  $\varphi_1$  has the definability property with respect to  $\theta'(E')$ .

**§4. Definability via interpolation.** In classical model theory, Beth definability theorem is often presented as one of the applications of Craig interpolation [11, 30]. In this section we develop an institution-independent proof of Beth theorem based on interpolation properties. Let us first recall how interpolation is conceptualised at the level of arbitrary institutions.

For any classes  $\mathcal{L}$  and  $\mathcal{R}$  of signature morphisms in an institution  $\mathcal{I}$ , the institution has the *Craig-Robinson  $\langle \mathcal{L}, \mathcal{R} \rangle$ -interpolation property*, if for any pushout in  $\text{Sig}$  such that  $\varphi_1 \in \mathcal{L}$  and  $\varphi_2 \in \mathcal{R}$ , any set of  $\Sigma_1$ -sentences  $E_1$  and any sets of  $\Sigma_2$ -sentences  $E_2$  and  $\Gamma_2$  with  $\theta_1(E_1) \cup \theta_2(\Gamma_2) \models \theta_2(E_2)$  there exists a set of  $\Sigma$ -sentences  $E$  (called the *interpolant*) such that  $E_1 \models \varphi_1(E)$  and  $\varphi_2(E) \cup \Gamma_2 \models E_2$ .

$$\begin{array}{ccc} \Sigma & \xrightarrow{\varphi_1} & \Sigma_1 \\ \varphi_2 \downarrow & & \downarrow \theta_1 \\ \Sigma_2 & \xrightarrow{\theta_2} & \Sigma' \end{array}$$

The restriction given by  $\Gamma_2$  being empty is called *Craig  $\langle \mathcal{L}, \mathcal{R} \rangle$ -interpolation*.

This generalises the conventional formulations of interpolation in several ways:

- From intersection-union squares of signatures to classes of pushout squares. While the unsorted sub-institution of **FOL** has Craig-Robinson interpolation for all pushout squares [22], (many sorted) **FOL** has it only for those where one component is an (*i\*\**)-morphism [25], and **HCL** and **EQL** only have Craig interpolation for pushout squares where  $\mathcal{R}$  is the class of (*iii*)-morphisms [43, 16].

<sup>6</sup>As these will not be used anywhere else in our work we prefer to leave the proofs of these properties as exercises for the interested reader.

- Using sets of sentences rather than single sentences accommodates interpolation results for equational logic [43] as well as for other institutions having Birkhoff-style axiomatizability properties [16]. However it is easy to notice that when  $E_2$  consists of a single sentence, if the institution is compact, then the interpolant can be chosen finite, and if the institution has finite conjunctions too, then the interpolant can also be chosen to be a single sentence.
- Craig-Robinson interpolation strengthen Craig interpolation by adding to the ‘primary’ premises  $E_1$  a set  $\Gamma_2$  (of  $\Sigma_2$ -sentences) as ‘secondary’ premises. Craig-Robinson interpolation plays an important role in specification language theory, see [4, 21, 22]. The name “Craig-Robinson” interpolation has been used for instances of this property in [46, 52, 22] and “strong Craig interpolation” in [21]. One can prove that in any institution which has implications and is compact, Craig-Robinson interpolation is equivalent to Craig interpolation [22].

**THEOREM 4.1.** In any semi-exact (compact) institution having Craig-Robinson  $(\mathcal{L}, \mathcal{R})$ -interpolation for classes  $\mathcal{L}$  and  $\mathcal{R}$  of signature morphisms which are stable under pushouts, any signature morphism in  $\mathcal{L} \cap \mathcal{R}$  has the (finite) definability property.

**PROOF.** Let  $(\varphi: \Sigma \rightarrow \Sigma') \in \mathcal{L} \cap \mathcal{R}$  be defined implicitly by  $E' \subseteq \text{Sen}(\Sigma')$ . We consider the pushout of  $\varphi$  with an arbitrary signature morphism  $\theta: \Sigma \rightarrow \Sigma_1$  and a  $\Sigma'_1$ -sentence  $\rho$ .

$$\begin{array}{ccc} \Sigma & \xrightarrow{\varphi} & \Sigma' \\ \theta \downarrow & & \downarrow \theta' \\ \Sigma_1 & \xrightarrow{\varphi_1} & \Sigma'_1 \end{array}$$

Now we consider the pushout of  $\varphi_1$  with itself:

$$\begin{array}{ccc} \Sigma_1 & \xrightarrow{\varphi_1} & \Sigma'_1 \\ \varphi_1 \downarrow & & \downarrow \theta_1 \\ \Sigma'_1 & \xrightarrow{\theta_2} & \Sigma'' \end{array}$$

Let us show that  $\theta_1(\theta'(E')) \cup \theta_1(\rho) \cup \theta_2(\theta'(E')) \models_{\Sigma''} \theta_2(\rho)$ . Consider a  $\Sigma''$ -model  $M'' \models \theta_1(\theta'(E')) \cup \theta_1(\rho) \cup \theta_2(\theta'(E'))$ . We have that  $(M'' \upharpoonright_{\theta_1 \upharpoonright_{\theta'}}) \upharpoonright_{\varphi} = (M'' \upharpoonright_{\theta_1 \upharpoonright_{\varphi_1}}) \upharpoonright_{\theta} = (M'' \upharpoonright_{\theta_2 \upharpoonright_{\varphi_1}}) \upharpoonright_{\theta} = (M'' \upharpoonright_{\theta_2 \upharpoonright_{\theta'}}) \upharpoonright_{\varphi}$ . By the Satisfaction Condition we have that  $(M'' \upharpoonright_{\theta_1}) \upharpoonright_{\theta'} \models E'$  and  $(M'' \upharpoonright_{\theta_2}) \upharpoonright_{\theta'} \models E'$ . By the implicit definability of  $\varphi$ , we get that  $(M'' \upharpoonright_{\theta_1}) \upharpoonright_{\theta'} = (M'' \upharpoonright_{\theta_2}) \upharpoonright_{\theta'}$ . Since we also have  $(M'' \upharpoonright_{\theta_1}) \upharpoonright_{\varphi_1} = (M'' \upharpoonright_{\theta_2}) \upharpoonright_{\varphi_1}$ , by the semi-exactness we get  $M'' \upharpoonright_{\theta_1} = M'' \upharpoonright_{\theta_2}$ . By the Satisfaction Condition  $M'' \models \theta_1(\rho)$  implies  $M'' \upharpoonright_{\theta_2} = M'' \upharpoonright_{\theta_1} \models \rho$  which further implies  $M'' \models \theta_2(\rho)$ .

Now because  $\varphi \in \mathcal{L} \cap \mathcal{R}$  and  $\mathcal{L}$  and  $\mathcal{R}$  are stable under pushouts, we have that  $\varphi_1 \in \mathcal{L} \cap \mathcal{R}$ , and by Craig-Robinson interpolation (and compactness) there exists (finite)  $E_\rho \subseteq \text{Sen}(\Sigma_1)$  such that  $\theta'(E') \cup \{\rho\} \models \varphi_1(E_\rho)$  and  $\theta'(E') \cup \varphi_1(E_\rho) \models \rho$ , which just means that  $\theta'(E') \models \rho \Leftrightarrow \varphi_1(E_\rho)$ . At this point, it follows immediately that  $E' \models (\forall \theta')(\rho \Leftrightarrow \varphi_1(E_\rho))$ .  $\dashv$

**COROLLARY 4.2.** In (many sorted) **FOL**, any  $(i^{**})$ -morphism of signatures has the finite definability property.

**PROOF.** Let  $\S$  be the class of  $(i^{**})$ -morphisms of signatures. From [8, 17, 25] we know that **FOL** has Craig  $\langle \S, \S \rangle$ -interpolation, hence it has Craig-Robinson  $\langle \S, \S \rangle$ -interpolation (because **FOL** has implications and is compact; see [22]).  $\dashv$

**REMARK 4.3.** Because tight signature morphisms in **FOL** are the  $(s^{**})$ -morphisms of signatures, it means that the equivalence between implicit and explicit definability holds in **FOL** for the  $(b^{**})$ -morphisms of signatures.

**§5. Definability via axiomatizability.** Definability Theorem 4.1 relies on Craig-Robinson interpolation, which does not hold for institutions having strong axiomatizability properties, such as **HCL** and **EQL**. In this section we develop another definability result which relies on axiomatizability properties and which can be applied to a series of actual situations when Craig-Robinson interpolation fails.

The so-called ‘Birkhoff institutions’ of [16] define an abstract concept of Birkhoff-style axiomatizability in arbitrary institutions going well beyond the classical axiomatizability results for (quasi-)varieties. They had been used in [16] as a basis for developing an institution-independent proof of Craig interpolation theorem by dependency of axiomatizability properties.

**Filtered products.** Recall that a poset (i.e., partially ordered set)  $(J, \leq)$  is *directed* when to any two elements  $i$  and  $j$  there exists an element  $k$  such that  $i \leq k$  and  $j \leq k$ . A colimit of a functor  $D: J \rightarrow \mathbb{C}$  is *directed* when  $J$  is a directed poset.

Let  $\mathbb{C}$  be a category with small products and directed colimits. Consider a family of objects  $\{A_i\}_{i \in I}$ . Each filter  $F$  over the set of indices  $I$  determines a functor  $A_F: F \rightarrow \mathbb{C}$  such that  $A_F(J \subset J') = p_{J',J}: \prod_{i \in J'} A_i \rightarrow \prod_{i \in J} A_i$  for each  $J, J' \in F$  with  $J \subset J'$ , and with  $p_{J',J}$  being the canonical projection.

Then the *filtered product of  $\{A_i\}_{i \in I}$  modulo  $F$*  is the colimit  $\mu: A_F \Rightarrow \prod_F A_i$  of the functor  $A_F$ .

$$\begin{array}{ccc}
 \prod_{i \in J'} A_i & \xrightarrow{p_{J',J}} & \prod_{i \in J} A_i \\
 \searrow \mu_{J'} & & \swarrow \mu_J \\
 & \prod_F A_i &
 \end{array}$$

If  $F$  is an ultrafilter then the filtered product modulo  $F$  is called an *ultraproduct*.

Notice that  $F$  is a directed poset, hence under our assumptions the filtered products always exist. The filtered product construction from classical model theory (see Chapter 4 of [11]) has been probably defined categorically for the first time in [34] and has been used in some abstract model theoretic works, such as [1]. The equivalence between the category theoretic and the set theoretic definitions of the filtered products is shown in [29].<sup>7</sup>

---

<sup>7</sup>However this relies upon an appropriate concept of model homomorphism avoiding the usual classical model theoretic restrictions to ‘embeddings’ (i.e., closed inclusive model homomorphisms) or even to ‘elementary embeddings’. In fact it is easy to see that the categorical filtered products makes essential use of projections, which are rather far from any concept of model ‘embedding’.

Given a class  $\mathcal{F}$  of filters, for each class  $\mathbf{K} \subseteq |\mathbb{C}|$  of objects in the category  $\mathbb{C}$  let  $\mathcal{F}\mathbf{K}$  be the class of all filtered products modulo  $F$  of models from  $\mathbf{K}$  for all filters  $F \in \mathcal{F}$ , i.e.,  $\mathcal{F}\mathbf{K} = \{\prod_F A_i \mid F \in \mathcal{F} \text{ filter over some set of indices } I \text{ and } A_i \in \mathbf{K} \text{ for each } i \in I\}$ . Notice that  $\mathcal{F}\mathbf{K}$  is the closure of  $\mathbf{K}$  under products when  $\mathcal{F} = \{\{I\} \mid I \text{ set}\}$  and it is the closure under isomorphisms when  $\mathcal{F} = \{\{\{*\}\}\}$ .

**Birkhoff institutions.** Recall from [16] that  $(\text{Sig}, \text{Sen}, \text{Mod}, \models, \mathcal{F}, \mathcal{B})$  is a *Birkhoff institution* if and only if

1.  $(\text{Sig}, \text{Sen}, \text{Mod}, \models)$  is an institution such that the category of models  $\text{Mod}(\Sigma)$  has filtered products for each signature  $\Sigma \in |\text{Sig}|$ ,
2.  $\mathcal{F}$  is a class of filters with  $\{\{*\}\} \in \mathcal{F}$ , and
3.  $\mathcal{B}_\Sigma \subseteq |\text{Mod}(\Sigma)| \times |\text{Mod}(\Sigma)|$  is a reflexive binary relation for each signature  $\Sigma \in |\text{Sig}|$

such that

$$\mathbb{M}^{**} = \mathcal{B}_\Sigma^{-1}(\mathcal{F}\mathbb{M})$$

for each signature  $\Sigma$  and each class of  $\Sigma$ -models  $\mathbb{M} \subseteq |\text{Mod}(\Sigma)|$ .

Here we slightly strenghten<sup>8</sup> the original concept of Birkhoff institution introduced in [16] by imposing that  $\mathcal{B}$  is *closed under isomorphisms*, i.e.,  $\mathcal{B}_\Sigma; \cong_\Sigma = \mathcal{B}_\Sigma = \cong_\Sigma; \mathcal{B}_\Sigma$  for each signature  $\Sigma$ .

NOTATION 5.1. Given a class  $\mathcal{H} \subseteq \mathbb{C}$  of arrows (morphisms) of the category  $\mathbb{C}$ , we define the (class) relation  $\xrightarrow{\mathcal{H}} \subseteq |\mathbb{C}| \times |\mathbb{C}|$  by  $a \xrightarrow{\mathcal{H}} b$  if and only if there exists an arrow  $h: a \rightarrow b$  with  $h \in \mathcal{H}$ . The inverse  $(\xrightarrow{\mathcal{H}})^{-1}$  is denoted as  $\xleftarrow{\mathcal{H}}$ .

EXAMPLE 5.2. The following is a rather short list of Birkhoff institutions obtained as sub-institutions of  $\mathbf{FOL}_{\infty, \omega}$  by varying the type of sentences and via various well-known axiomatizability results:

institution	$\mathcal{B}$	$\mathcal{F}$
<b>FOL</b>	$\equiv$	all ultrafilters
<b>FOL</b>	<i>ultraradicals</i> (see [42])	all ultrafilters
<b>PL</b>	$=$	all ultrafilters
universal (or quantifier-free) <b>FOL</b> -sentences	$\xrightarrow{S_c}$	all ultrafilters
universal <b>FOL</b> <sub><math>\infty, \omega</math></sub> -sentences	$\xrightarrow{S_c}$	$\{\{\{*\}\}\}$
<b>HCL</b> <sub><math>\infty</math></sub>	$\xrightarrow{S_c}$	$\{\{I\} \mid I \text{ set}\}$
<b>HCL</b>	$\xrightarrow{S_c}$	all filters
universal <b>FOL</b> -atoms	$\xleftarrow{H_r}; \xrightarrow{S_c}$	$\{\{I\} \mid I \text{ set}\}$
<b>EQL</b>	$\xleftarrow{H_r}; \xrightarrow{S_w}$	$\{\{I\} \mid I \text{ set}\}$
$\forall\forall$	$\xleftarrow{H_s}; \xrightarrow{S_c}$	all ultrafilters
$\forall\forall_{\infty}$	$\xleftarrow{H_s}; \xrightarrow{S_c}$	$\{\{\{*\}\}\}$
$\forall\exists$ (universal-existential <b>FOL</b> -sentences)	<i>sandwiches</i> (see [11])	all ultrafilters.

where a model homomorphism  $h: M \rightarrow N$  for a signature  $(S, F, P)$  is *closed* when  $M_\pi = h^{-1}(N_\pi)$ , and *strong* when  $h(M_\pi) = N_\pi$ , for each arity  $w \in S^*$  and each

<sup>8</sup>But without really narrowing the actual examples.

relation symbol  $\pi \in P_w$ , and where we let  $H_r$  denote the class of surjective,  $H_s$  the class of strong surjective,  $H_c$  the class of closed surjective,  $S_w$  the class of injective, and  $S_c$  the class of closed injective<sup>9</sup> model homomorphisms.

A complete list of **FOL**-based Birkhoff institutions can be obtained by using results from [2, 41].

EXAMPLE 5.3. A large list of **PA**-based Birkhoff institutions can be also obtained from [2, 41]; here we list only few of them:

institution	$\mathcal{B}$	$\mathcal{F}$
<b>PA</b>	$\equiv$	all ultrafilters
universal <b>PA</b> -sentences	$\xrightarrow{S_c}$	all ultrafilters
$QE(\mathbf{PA})$	$\xrightarrow{S_c}$	$\{\{I\} \mid I \text{ set}\}$
$QE^\omega(\mathbf{PA})$	$\xrightarrow{S_c}$	all filters
$QE_1(\mathbf{PA})$	$\xrightarrow{S_f}$	$\{\{I\} \mid I \text{ set}\}$
$QE_1^\omega(\mathbf{PA}) = QE^\omega(\mathbf{PA}) \cap QE_1(\mathbf{PA})$	$\xrightarrow{S_f}$	all filters
$QE_2(\mathbf{PA})$	$\xrightarrow{S_w}$	$\{\{I\} \mid I \text{ set}\}$
$QE_2^\omega(\mathbf{PA}) = QE^\omega(\mathbf{PA}) \cap QE_2(\mathbf{PA})$	$\xrightarrow{S_w}$	all filters
$E(\mathbf{PA})$ (universal existence equations)	$\xleftarrow{H_r}; \xrightarrow{S_c}$	$\{\{I\} \mid I \text{ set}\}$
$E_1(\mathbf{PA}) = E(\mathbf{PA}) \cap QE_1(\mathbf{PA})$	$\xleftarrow{H_r}; \xrightarrow{S_f}$	$\{\{I\} \mid I \text{ set}\}$
$E_2(\mathbf{PA}) = E(\mathbf{PA}) \cap QE_2(\mathbf{PA})$	$\xleftarrow{H_r}; \xrightarrow{S_w}$	$\{\{I\} \mid I \text{ set}\}$

where  $S_w$  is the class of all injective homomorphisms,  $S_c$  is the class of all *closed* injective homomorphisms  $h: A \rightarrow B$  (i.e.,  $A_\sigma(a)$  is defined if  $B_\sigma(h(a))$  is defined),  $S_f$  of all *full* injective homomorphisms  $h: A \rightarrow B$  (i.e.,  $A_\sigma(a) = a_0$  if  $B_\sigma(h(a)) = h(a_0)$  for  $a, a_0 \in A$ ), and  $H_r$  of surjective homomorphisms.

Also, the general axiomatizability results of [2] can be easily applied for obtaining Birkhoff institutions out of recent algebraic specification logics such as membership algebra [37], rewriting logic [36], multi-algebras for non-determinism [32], etc. In dependence of Birkhoff-style axiomatizability results many other examples can be developed for various institutions in algebraic specification, computing science, or logic.

The abstract Beth definability via axiomatizability relies on a ‘lifting’ condition of the signature morphism.

DEFINITION 5.4. Given a family of relations  $R = \{R_\Sigma \subseteq |\text{Mod}(\Sigma)| \times |\text{Mod}(\Sigma)|\}_{\Sigma \in |\text{Sig}|}$  indexed by the category of the signatures of an institution, a signature morphism  $\varphi: \Sigma \rightarrow \Sigma'$

- *lifts*  $R$  iff for each  $\Sigma'$ -model  $M'$  and each  $\Sigma$ -model  $N$ , if  $\langle M' \upharpoonright_\varphi, N \rangle \in R_\Sigma$  then there exists  $N'$  a  $\varphi$ -expansion of  $N$  such that  $\langle M', N' \rangle \in R_{\Sigma'}$ , and
- *lifts weakly*  $R$  iff for each  $\Sigma'$ -model  $M'$  and  $N'$ , if  $\langle M' \upharpoonright_\varphi, N' \upharpoonright_\varphi \rangle \in R_\Sigma$  then there exists  $P'$  a  $\varphi$ -expansion of  $N' \upharpoonright_\varphi$  such that  $\langle M', P' \rangle \in R_{\Sigma'}$ .

REMARK 5.5. A signature morphism lifts weakly a family of relations  $R$  whenever it lifts  $R$ .

<sup>9</sup>In [3] these are called ‘strong’ rather than ‘closed’.

The (non-weakly) lifting concept of Definition 5.4 has been defined and used in [16], however it is important to notice that Theorem 5.6 below uses the lifting condition in a reverse direction than the main result of [16], a fact which suggests that contrary to what happens in Theorem 4.1 the definability result of Theorem 5.6 below is not caused by an interpolation property.

**THEOREM 5.6.** Consider a (compact) semi-exact Birkhoff institution  $(\text{Sig}, \text{Sen}, \text{Mod}, \models, \mathcal{F}, \mathcal{B})$  and a class  $\mathcal{S} \subseteq \text{Sig}$  of signature morphisms which is stable under pushouts and such that for each  $\varphi \in \mathcal{S}$

- $\text{Mod}(\varphi)$  preserves filtered products (of models), and
- $\varphi$  lifts weakly  $\mathcal{B}^{-1}$ .

Then any signature morphism in  $\mathcal{S}$  has the (finite) definability property.

**PROOF.** Let  $\varphi \in \mathcal{S}$ . If  $\varphi: \Sigma \rightarrow \Sigma'$  is implicitly defined by  $E'$ , then we show it is (finitely) explicitly defined by  $E'$  too. Therefore consider any pushout square of signature morphisms for the span  $\Sigma_1 \xleftarrow{\theta} \Sigma \xrightarrow{\varphi} \Sigma'$

$$\begin{array}{ccc} \Sigma & \xrightarrow{\varphi} & \Sigma' \\ \theta \downarrow & & \downarrow \theta' \\ \Sigma_1 & \xrightarrow{\varphi_1} & \Sigma'_1 \end{array}$$

and any  $\rho \in \text{Sen}(\Sigma'_1)$ .

By the hypotheses on the Birkhoff institution we have that  $\varphi_1$  lifts weakly  $\mathcal{B}^{-1}$  and preserves filtered products. Let us denote  $\text{Mod}(\Sigma'_1, \theta'(E') \cup \{\rho\})$  by  $\mathbb{M}'_1$ . We define  $E_\rho$  as  $(\mathbb{M}'_1 \upharpoonright_{\varphi_1})^*$ .

We first show  $\theta'(E') \cup \{\rho\} \models \varphi_1(E_\rho)$ . Consider  $M'_1$  a model of  $\theta'(E') \cup \rho$ . This implies that  $M'_1 \upharpoonright_{\varphi_1} \in \mathbb{M}'_1 \upharpoonright_{\varphi_1}$  and because  $E_\rho$  is satisfied by all models in  $\mathbb{M}'_1 \upharpoonright_{\varphi_1}$  we have that  $M'_1 \upharpoonright_{\varphi_1} \models E_\rho$ . By the Satisfaction Condition we obtain that  $M'_1 \models \varphi_1(E_\rho)$ .

Now we show that  $\theta'(E') \cup \varphi_1(E_\rho) \models \rho$ . Consider  $M'_1$  a  $\Sigma'_1$ -model satisfying  $\theta'(E') \cup \varphi_1(E_\rho)$ . By the Satisfaction Condition we have that  $M'_1 \upharpoonright_{\varphi_1} \models E_\rho = (\mathbb{M}'_1 \upharpoonright_{\varphi_1})^*$ . Because of the conditions on our Birkhoff institution  $M'_1 \upharpoonright_{\varphi_1} \in (\mathbb{M}'_1 \upharpoonright_{\varphi_1})^{**} = \mathcal{B}_{\Sigma_1}^{-1}(\mathcal{F}(\mathbb{M}'_1 \upharpoonright_{\varphi_1}))$ . By considering the following:

- $\mathcal{F}(\mathbb{M}'_1 \upharpoonright_{\varphi_1}) = \cong_{\Sigma_1}((\mathcal{F}\mathbb{M}'_1) \upharpoonright_{\varphi_1})$  because  $\varphi_1$  preserves filtered products,
- $\cong_{\Sigma_1}; \mathcal{B}_{\Sigma_1}^{-1} = \mathcal{B}_{\Sigma_1}^{-1}$  because  $\mathcal{B}$  is closed under isomorphisms,
- $\mathcal{F}\mathbb{M}'_1 \subseteq \mathcal{B}_{\Sigma'_1}^{-1}(\mathcal{F}\mathbb{M}'_1)$  because  $\mathcal{B}$  is reflexive, and
- $\mathcal{B}_{\Sigma'_1}^{-1}(\mathcal{F}\mathbb{M}'_1) = \mathbb{M}'_1$  because  $\mathbb{M}'_1$  is elementary.

it results that

$$\begin{aligned} M'_1 \upharpoonright_{\varphi_1} \in \mathcal{B}_{\Sigma_1}^{-1}(\mathcal{F}(\mathbb{M}'_1 \upharpoonright_{\varphi_1})) &= \mathcal{B}_{\Sigma_1}^{-1}(\cong_{\Sigma_1}((\mathcal{F}\mathbb{M}'_1) \upharpoonright_{\varphi_1})) \\ &= \mathcal{B}_{\Sigma_1}^{-1}((\mathcal{F}\mathbb{M}'_1) \upharpoonright_{\varphi_1}) \subseteq \mathcal{B}_{\Sigma'_1}^{-1}(\mathbb{M}'_1 \upharpoonright_{\varphi_1}). \end{aligned}$$

This implies that there exists a  $\Sigma'_1$ -model  $N'_1$  satisfying  $\theta'(E') \cup \{\rho\}$  and such that  $\langle M'_1 \upharpoonright_{\varphi_1}, N'_1 \upharpoonright_{\varphi_1} \rangle \in \mathcal{B}_{\Sigma_1}$ . Because  $\varphi_1$  lifts  $\mathcal{B}^{-1}$  it exists a  $\Sigma'_1$ -model  $P'_1$  such that  $P'_1 \upharpoonright_{\varphi_1} = M'_1 \upharpoonright_{\varphi_1}$  and  $\langle P'_1, N'_1 \rangle \in \mathcal{B}_{\Sigma'_1}$ .

Because  $\{\{\{*\}\}\} \in \mathcal{F}$  we have that  $\mathcal{B}_{\Sigma_1}^{-1}(\mathbb{M}'_1) \subseteq \mathcal{B}_{\Sigma_1}^{-1}(\mathcal{F}\mathbb{M}'_1) = \mathbb{M}'_1$ . From  $P'_1 \in \mathcal{B}_{\Sigma_1}^{-1}(N'_1) \subseteq \mathcal{B}_{\Sigma_1}^{-1}(\mathbb{M}'_1)$  we therefore get that  $P'_1 \in \mathbb{M}'_1$  which means that  $P'_1 \models \theta'(E') \cup \{\rho\}$ .

From  $M'_1, P'_1 \models \theta'(E')$  we have that  $M'_1 \upharpoonright_{\theta'}, P'_1 \upharpoonright_{\theta'} \models E'$  and because  $\varphi$  is implicitly defined by  $E'$  and  $(M'_1 \upharpoonright_{\theta'}) \upharpoonright_{\varphi} = M'_1 \upharpoonright_{\varphi_1} \upharpoonright_{\theta} = P'_1 \upharpoonright_{\varphi_1} \upharpoonright_{\theta} = (P'_1 \upharpoonright_{\theta'}) \upharpoonright_{\varphi}$  we obtain  $M'_1 \upharpoonright_{\theta'} = P'_1 \upharpoonright_{\theta'}$ . By the semi-exactness, from  $M'_1 \upharpoonright_{\varphi_1} = P'_1 \upharpoonright_{\varphi_1}$  and  $M'_1 \upharpoonright_{\theta'} = P'_1 \upharpoonright_{\theta'}$  we get that  $M'_1 = P'_1$ . Thus  $M'_1 \models \rho$ .

We have therefore showed that  $\theta'(E') \cup \{\rho\} \models \varphi_1(E_\rho)$  and  $\theta'(E') \cup \varphi_1(E_\rho) \models \rho$ . Moreover, when the institution is compact,  $E_\rho$  can be chosen finite. Thus  $\theta'(E') \models \rho \Leftrightarrow \varphi_1(E_\rho)$ , which implies that  $E' \models (\forall \theta')(\rho \Leftrightarrow \varphi_1(E_\rho))$ .  $\dashv$

**REMARK 5.7.** This definability result relies primarily on the Birkhoff-style axiomatizability property of the institution. Secondly, it relies on the lifting condition of the Birkhoff relation, which in the actual Birkhoff institutions is the core technical condition which should be established in order to obtain the definability property. The other conditions are very mild or even trivial in the applications. The preservation of filtered products by the model reduct functor follows in general from preservation of direct products and directed colimits. Preservation of direct products of models follows from the existence of free models along signature morphisms (since right adjoint functors preserve all limits) which can be established easily even at an institution-independent level by making use of elementary diagrams [15]. Preservation of directed colimits of models is a consequence of the finiteness of the arities of the symbols of the signatures, in fact under this condition the model reduct functors create directed colimits (see [33] for the special case of general (total) algebra).

We now illustrate the applicability of Theorem 5.6 with the sub-institutions of **FOL** listed by Example 5.2 and of **PA** listed by Example 5.3.

**PROPOSITION 5.8.** In **FOL**, any  $(bbi)$ -morphism of signatures lifts  $\xrightarrow{S_w}, \xleftarrow{S_c}, \xrightarrow{H_t}$ , and  $\xrightarrow{H_s}$ , and any  $(ss^*)$ -morphism of signatures lifts weakly  $\xleftarrow{S_w}$  and  $\xleftarrow{S_c}$ .

**PROOF.** Assume  $\varphi: (S, F, P) \rightarrow (S', F', P')$  is a  $(bbi)$ -morphism.

Let  $h: N \rightarrow M' \upharpoonright_{\varphi}$  be an injective  $(S, F, P)$ -model homomorphism. We define  $N'$  to be the unique  $\varphi$ -expansion of  $N$  such that  $N'_\pi = h^{-1}(M'_\pi)$  for each  $\pi \in P' \setminus \varphi(P)$ . Then  $h': N' \rightarrow M'$ , the unique  $\varphi$ -expansion of  $h$ , is an injective  $(S', F', P')$ -model homomorphism. Moreover, if  $h: N \hookrightarrow M' \upharpoonright_{\varphi}$  is closed, then  $h': N' \hookrightarrow M'$  is closed too.

Now let  $h: M' \upharpoonright_{\varphi} \rightarrow N$  be a surjective  $(S, F, P)$ -model homomorphism. We define  $N'$  to be the unique  $\varphi$ -expansion of  $N$  such that  $N'_\pi = h(M'_\pi)$  for each  $\pi \in P' \setminus \varphi(P)$ . Then  $h': M' \rightarrow N'$ , the unique  $\varphi$ -expansion of  $h$ , is also a surjective  $(S', F', P')$ -model homomorphism. Moreover, if  $h: M' \upharpoonright_{\varphi} \rightarrow N$  is strong, then  $h': M' \rightarrow N'$  is strong too.

Now we assume  $\varphi: (S, F, P) \rightarrow (S', F', P')$  is a  $(ss^*)$ -morphism.

Let  $h: N' \upharpoonright_{\varphi} \rightarrow M' \upharpoonright_{\varphi}$  be an injective  $(S, F, P)$ -model homomorphism. We define  $Q'$  to be the unique  $(S', F', P')$ -expansion of  $N' \upharpoonright_{(S', F', \varphi(P))}$  such that  $Q'_\pi = h^{-1}(M'_\pi)$  for each  $\pi \in P' \setminus \varphi(P)$ . Then  $h': Q' \rightarrow M'$  defined by  $h'_{\varphi(s)} = h_s$  for each  $s \in S$  is well defined and is an injective  $(S', F', P')$ -model homomorphism. Moreover, if  $h: N' \upharpoonright_{\varphi} \rightarrow M' \upharpoonright_{\varphi}$  is closed, then  $h: Q' \hookrightarrow M'$  is closed too.  $\dashv$

COROLLARY 5.9. We have the following table of definability results:

institution	signature morphism	definability property
<b>HCL</b>	$ss*$	finite definability
<b>HCL</b> <sub>∞</sub>	$ss*$	definability
universal <b>FOL</b> -sentences	$ss*$	finite definability
universal <b>FOL</b> <sub>∞,ω</sub> -sentences	$ss*$	definability
universal <b>FOL</b> -atoms	$bbi$	finite definability
$\forall\forall$	$bbi$	finite definability
$\forall\forall_\infty$	$bbi$	definability

PROOF. From Example 5.2 and Theorem 5.6, because the composition of a relation lifted weakly by  $\varphi$  with a relation lifted by  $\varphi$  gets a relation lifted weakly by  $\varphi$ , because every model reduct functor preserves direct products and directed colimits, and by taking into consideration the compactness property of each institution.  $\dashv$

PROPOSITION 5.10. In **PA**, any  $(bbi)$ -morphism of signatures lifts  $\xleftarrow{S_w}$  and  $\xleftarrow{S_f}$  and any  $(ss*)$ -morphism of signatures lifts weakly  $\xleftarrow{S_w}$  and  $\xleftarrow{S_f}$ .

PROOF. Assume  $\varphi: (S, TF, PF) \rightarrow (S', TF', PF')$  is a  $(bbi)$ -morphism. Let  $h: B \rightarrow A' \upharpoonright_\varphi$  be an injective  $(S, TF, PF)$ -algebra homomorphism, i.e.,  $h \in S_w$ . We define  $B'$  to be the unique  $\varphi$ -expansion of  $B$  such that

$$B'_\sigma(b) = \begin{cases} A'_\sigma(h(b)) & \text{if } A'_\sigma(h(b)) \text{ defined and } A'_\sigma(h(b)) \in h(B), \\ \text{undefined} & \text{otherwise} \end{cases}$$

for each  $\sigma \in PF' \setminus \varphi(PF)$ . Then  $h': B' \rightarrow A'$ , the unique  $\varphi$ -expansion of  $h$ , is an injective  $(S', TF', PF')$ -algebra homomorphism. Moreover, if  $h: B \rightarrow A' \upharpoonright_\varphi$  is full, then  $h': B' \rightarrow A'$  is full too.

Now we assume  $\varphi: (S, TF, PF) \rightarrow (S', TF', PF')$  is an  $(ss*)$ -morphism.

Let  $h: B' \upharpoonright_\varphi \rightarrow A' \upharpoonright_\varphi$  be an injective  $(S, TF, PF)$ -algebra homomorphism, i.e.,  $h \in S_w$ . Let  $C'$  be the  $(S', TF', PF')$ -expansion of  $B' \upharpoonright_{(S', TF', \varphi(PF))}$  such that the graph of  $C'_\sigma$  is empty for each  $\sigma \in PF' \setminus \varphi(PF)$ .

If  $h \in S_f$ , then we define  $C'$  to be the unique  $(S', TF', PF')$ -expansion of  $B' \upharpoonright_{(S', TF', \varphi(PF))}$  such that  $h(C'_\sigma(c)) = A'_\sigma(c)$  for each  $\sigma \in PF'_{\varphi(w) \rightarrow \varphi(s)} \setminus \varphi(PF_{w \rightarrow s})$  and  $c \in C'_{\varphi(w)}$  such that  $A'_\sigma(c) \in h(C'_{\varphi(s)})$ . Then  $h': C' \rightarrow A'$  defined by  $h'_{\varphi(s)} = h_s$  for each  $s \in S$  is well defined and is an injective  $(S', TF', PF')$ -algebra homomorphism. Moreover, if  $h: B' \upharpoonright_\varphi \rightarrow A' \upharpoonright_\varphi$  is full, we get that  $h': C' \rightarrow A'$  is full too.  $\dashv$

REMARK 5.11. The  $(bbi)$ -morphisms do not lift weakly neither the closed subalgebra relation  $\xleftarrow{S_c}$  nor  $\xrightarrow{H_r}$ .

COROLLARY 5.12. We have the following table of definability results:

institution	signature morphism	definability property
$QE_1^\omega(\mathbf{PA})$	$ss*$	finite definability
$QE_2^\omega(\mathbf{PA})$	$ss*$	finite definability
$QE_1(\mathbf{PA})$	$ss*$	definability
$QE_1(\mathbf{PA})$	$ss*$	definability

PROOF. From Example 5.3 and Theorem 5.6, by the same argument as the proof of Corollary 5.9.  $\dashv$

REMARK 5.13. While  $E_1(\mathbf{PA})$ ,  $QE_1(\mathbf{PA})$  and  $QE_1^\omega(\mathbf{PA})$  have the elementary diagrams of  $\mathbf{PA}$ ,  $E_2(\mathbf{PA})$ ,  $QE_2(\mathbf{PA})$  and  $QE_2^\omega(\mathbf{PA})$  do not. This means that for  $E_2(\mathbf{PA})$ ,  $QE_2(\mathbf{PA})$  and  $QE_2^\omega(\mathbf{PA})$ , the inclusion of the explicit definability into the implicit definability cannot be established by means of Proposition 3.9. Moreover, in  $E_2(\mathbf{PA})$ ,  $QE_2(\mathbf{PA})$  and  $QE_2^\omega(\mathbf{PA})$ , the interpretation of an implicitly defined partial operation symbol is always empty.

**§6. Borrowing definability.** In this section we develop a method which establishes the definability property rather indirectly by lifting and solving the definability problem to a different institution where the definability results are better known or easier to solve. Then the result is translated back to the original institution. Similar ‘borrowing’ methods have been used frequently in institution-independent model and specification theory, most notably, but not only, in [10] and [39].

For this we have to be able to map structurally between institutions. In the literature there are several concepts of such structure preserving mappings between institutions. The original one, introduced by [26], is adequate for encoding a ‘forgetful’ operation from a ‘richer’ institution to a ‘poorer’ one. However, institution comorphisms [28], previously know as ‘plain map’ in [35] or ‘representation’ in [50, 51], and capturing the idea of embedding of a ‘poorer’ institution into a ‘richer’ one, serve best our task here.

An *institution comorphism*  $(\Phi, \alpha, \beta): \mathcal{I} \rightarrow \mathcal{I}'$  consists of

1. a functor  $\Phi: \text{Sig} \rightarrow \text{Sig}'$ ,
2. a natural transformation  $\alpha: \text{Sen} \Rightarrow \Phi; \text{Sen}'$ , and
3. a natural transformation  $\beta: \Phi^{\text{op}}; \text{Mod}' \Rightarrow \text{Mod}$

such that the following *satisfaction condition* holds

$$M' \models'_{\Phi(\Sigma)} \alpha_\Sigma(e) \text{ iff } \beta_\Sigma(M') \models_\Sigma e$$

for each signature  $\Sigma \in |\text{Sig}|$ , for each  $\Phi(\Sigma)$ -model  $M'$ , and each  $\Sigma$ -sentence  $e$ .

EXAMPLE 6.1. The canonical embedding of equational logic **EQL** into first order logic can be expressed as a comorphism  $(\Phi, \alpha, \beta): \mathbf{EQL} \rightarrow \mathbf{FOL}$  such that  $\Phi(S, F) = (S, F, \emptyset)$ ,  $\alpha$  regards any equation as a first order sentence, and  $\beta_{(S,F)}: \text{Mod}^{\mathbf{FOL}}(S, F, \emptyset) \rightarrow \text{Mod}^{\mathbf{EQL}}(S, F)$  is the trivial isomorphism which regards any  $(S, F, \emptyset)$ -model as an  $(S, F)$ -algebra.

EXAMPLE 6.2. **EQL** can be embedded into the institution **PA** of partial algebra by means of the canonical comorphism which maps an algebraic signature  $(S, F)$  to the partial algebra signature  $(S, F, \emptyset)$ .

A rather different class of examples of comorphisms expresses the encoding of a ‘richer’, more complex, institution into a simpler one. Such encoding comorphisms are meaningful for our definability borrowing method because we would like to borrow definability from a simpler institution to a more complex one.

EXAMPLE 6.3. The institution **PA** of partial algebras can be encoded into the institution  $\mathbf{FOL}^T$  of the theories of first order logic by the following comorphism:

- Each **PA** signature  $(S, TF, PF)$  gets mapped to the **FOL** theory  $((S, TF, \overline{PF}), \Gamma)$  such that  $\overline{PF}_{w \rightarrow s} = PF_{w \rightarrow s}$  for each  $w \in S^*$  and  $s \in S$ , and

$$\Gamma = \{(\forall X \uplus \{y, z\})\sigma(X, y) \wedge \sigma(X, z) \Rightarrow (y = z) \mid \sigma \in PF\}$$

- Each  $(S, TF, \overline{PF})$ -model  $M$  gets mapped to the  $(S, TF, PF)$ -algebra  $\beta(M)$  such that  $\beta(M)_x = M_x$  for each  $x \in S$  or  $x \in TF$ , and  $\beta(M)_\sigma(m) = m_0$  when  $\sigma \in PF$  and  $(m, m_0) \in M_\sigma$ . (Notice that each  $(S, TF, \overline{PF})$ -model homomorphism  $h: M \rightarrow N$  is a  $(S, TF, PF)$ -algebra homomorphism too.)
- $\alpha$  preserves the quantifications and the logical connectives, and

$$\alpha(t \stackrel{e}{=} t') = (\exists X \uplus \{x_0\})bind(t, x_0) \wedge bind(t', x_0)$$

where for each  $(S, TF, PF)$ -term  $t$  and variable  $x$ ,  $bind(t, x)$  is a (finite) conjunction of atoms defined by

$$bind(\sigma(t_1 \dots t_n), x) = \bigwedge_{1 \leq i \leq n} bind(t_i, x_i) \wedge \begin{cases} \sigma(x_1, \dots, x_n) = x & \text{when } \sigma \in TF, \\ \sigma(x_1, \dots, x_n, x) & \text{when } \sigma \in PF \end{cases}$$

and  $X$  is the set of the new constants introduced by  $bind(t, x_0)$  and  $bind(t', x_0)$ .

(The proof of the Satisfaction Condition uses the fact that  $M \models (\exists X \uplus \{x_0\})bind(t, x_0)$  if and only if  $\beta(M)_t = M'_{x_0}$  where  $M'$  is the unique expansion of  $M$  that satisfies  $bind(t, x_0)$ .)

It is interesting to notice at this point that there is another more conventional encoding comorphism  $\mathbf{PA} \rightarrow \mathbf{FOL}^T$  which maps all **PA** operation symbols (total or partial) to **FOL** operation symbols (see [39]), however that one will not be adequate for the purpose of this section.

**DEFINITION 6.4.** Let  $(\Phi, \alpha, \beta): \mathcal{S} \rightarrow \mathcal{S}'$  be an institution comorphism. A  $\mathcal{S}$ -signature morphism  $\varphi: \Sigma_1 \rightarrow \Sigma_2$  is  $(\Phi, \alpha, \beta)$ -precise whenever the function  $\text{Mod}'(\Phi(\Sigma_2)) \rightarrow \text{Mod}'(\Phi(\Sigma_1)) \times \text{Mod}(\Sigma_2)$  mapping each  $M'_2$  to  $\langle M'_2 \upharpoonright_{\Phi(\varphi)}, \beta_{\Sigma_2}(M'_2) \rangle$  is injective.

The comorphism  $(\Phi, \alpha, \beta)$  is precise when each  $\mathcal{S}$ -signature morphism is  $(\Phi, \alpha, \beta)$ -precise.

**FACT 6.5.** The canonical embedding comorphisms  $\mathbf{EQL} \rightarrow \mathbf{FOL}$  and  $\mathbf{EQL} \rightarrow \mathbf{PA}$  and the encoding comorphism  $\mathbf{PA} \rightarrow \mathbf{FOL}^T$  are trivially precise.

**PROPOSITION 6.6.** Let  $(\Phi, \alpha, \beta): \mathcal{S} \rightarrow \mathcal{S}'$  be an institution comorphism. Then for any  $(\Phi, \alpha, \beta)$ -precise signature morphism  $\varphi$  and theory  $E'$ ,  $\Phi(\varphi)$  is defined implicitly by  $\alpha(E')$  if  $\varphi$  is defined implicitly by a  $E'$ .

**PROOF.** Let  $\varphi: \Sigma \rightarrow \Sigma'$  be a  $(\Phi, \alpha, \beta)$ -precise signature morphism.

Assume  $\varphi$  is defined implicitly by a  $E'$ , and let  $M'_1, M'_2 \in |\text{Mod}'(\Phi(\Sigma'), \alpha(E'))|$  such that  $M'_1 \upharpoonright_{\Phi(\varphi)} = M'_2 \upharpoonright_{\Phi(\varphi)}$ . Because  $\varphi$  is  $(\Phi, \alpha, \beta)$ -precise, if we show that  $\beta_{\Sigma'}(M'_1) = \beta_{\Sigma'}(M'_2)$  then we can deduce that  $M'_1 = M'_2$ .

But by the Satisfaction Condition for  $(\Phi, \alpha, \beta)$ ,  $\beta_{\Sigma'}(M'_1), \beta_{\Sigma'}(M'_2) \models E'$ , and  $\beta_{\Sigma'}(M'_1) \upharpoonright_{\varphi} = \beta_{\Sigma}(M'_1 \upharpoonright_{\Phi(\varphi)}) = \beta_{\Sigma}(M'_2 \upharpoonright_{\Phi(\varphi)}) = \beta_{\Sigma'}(M'_2) \upharpoonright_{\varphi}$  by the naturality of  $\beta$ . Because  $\varphi$  is defined implicitly by  $E'$ , we obtain that  $\beta_{\Sigma'}(M'_1) = \beta_{\Sigma'}(M'_2)$ .  $\dashv$

**DEFINITION 6.7.** An institution comorphism  $(\Phi, \alpha, \beta): \mathcal{I} \rightarrow \mathcal{I}'$  is *conservative* when for each  $\mathcal{I}$ -signature  $\Sigma$ , each  $\Sigma$ -model has at least one  $\beta_\Sigma$ -expansion, i.e.,  $\beta_\Sigma$  is surjective.

**FACT 6.8.** The comorphisms  $\mathbf{EQL} \rightarrow \mathbf{FOL}$ ,  $\mathbf{EQL} \rightarrow \mathbf{PA}$  and  $\mathbf{PA} \rightarrow \mathbf{FOL}^T$  are trivially conservative.

**PROPOSITION 6.9.** Let  $(\Phi, \alpha, \beta): \mathcal{I} \rightarrow \mathcal{I}'$  be a conservative institution comorphism such that  $\Phi$  preserves pushouts and  $\alpha$  is surjective modulo the semantic equivalence  $\models$ .

Then any  $\mathcal{I}$ -signature morphism  $\varphi$  is defined (finitely) explicitly by a theory  $E'$  if  $\Phi(\varphi)$  is defined (finitely) explicitly by  $\alpha(E')$ .

**PROOF.** Assume  $\Phi(\varphi)$  is defined explicitly by  $\alpha(E')$  and let

$$\begin{array}{ccc} \Sigma & \xrightarrow{\varphi} & \Sigma' \\ \theta \downarrow & & \downarrow \theta' \\ \Sigma_1 & \xrightarrow{\varphi_1} & \Sigma'_1 \end{array}$$

be any pushout of the span  $\Sigma_1 \xleftarrow{\theta} \Sigma \xrightarrow{\varphi} \Sigma'$  of signature morphisms and let  $\rho \in \text{Sen}(\Sigma'_1)$ .

Because  $\Phi$  preserves pushouts we have that

$$\begin{array}{ccc} \Phi(\Sigma) & \xrightarrow{\Phi(\varphi)} & \Phi(\Sigma') \\ \Phi(\theta) \downarrow & & \downarrow \Phi(\theta') \\ \Phi(\Sigma_1) & \xrightarrow{\Phi(\varphi_1)} & \Phi(\Sigma'_1) \end{array}$$

is a pushout in  $\text{Sig}'$ .

Because  $\Phi(\varphi)$  is defined (finitely) explicitly by  $\alpha_{\Sigma'}(E')$ , there exists (finite)  $E_{\alpha_{\Sigma'_1}(\rho)} \subseteq \text{Sen}'(\Phi(\Sigma_1))$  such that  $\alpha_{\Sigma'}(E') \models (\forall \Phi(\theta'))(\alpha_{\Sigma'_1}(\rho) \Leftrightarrow \Phi(\varphi_1)(E_{\alpha_{\Sigma'_1}(\rho)}))$ . Notice that  $E_\rho$  is finite whenever  $E_{\alpha_{\Sigma'_1}(\rho)}$  is finite.

We show that  $E' \models (\forall \theta')(\rho \Leftrightarrow \varphi_1(E_\rho))$  where  $E_\rho$  is chosen such that  $\alpha_{\Sigma_1}(E_\rho) \models E_{\alpha_{\Sigma'_1}(\rho)}$ , which is possible because  $\alpha_{\Sigma_1}$  is surjective modulo semantical equivalence  $\models$ .

Let us first notice that because  $\alpha$  preserves  $\Leftrightarrow$  and because it is natural,  $(\alpha_{\Sigma'_1}(\rho) \Leftrightarrow \Phi(\varphi_1)(E_{\alpha_{\Sigma'_1}(\rho)})) \models \alpha_{\Sigma'_1}(\rho \Leftrightarrow \varphi_1(E_\rho))$ . Therefore it is enough to show that  $\alpha_{\Sigma'}(E') \models (\forall \Phi(\theta'))\alpha_{\Sigma'_1}(e)$  implies  $E' \models (\forall \theta')e$  for each  $\Sigma'_1$ -sentence  $e$ .

We assume  $\alpha_{\Sigma'}(E') \models (\forall \Phi(\theta'))\alpha_{\Sigma'_1}(e)$ . By the Satisfaction Condition and the definition of quantifier satisfaction, this is equivalent to  $\Phi(\theta')(\alpha_{\Sigma'}(E')) \models \alpha_{\Sigma'_1}(e)$ . By the naturality of  $\alpha$ , this is equivalent to  $\alpha_{\Sigma'_1}(\theta'(E')) \models \alpha_{\Sigma'_1}(e)$ . From the conservativity of  $\beta$  we get that  $\theta'(E') \models e$ . Again by the Satisfaction Condition and the definition of quantifier satisfaction we get that  $E' \models (\forall \theta')e$ .  $\dashv$

**COROLLARY 6.10.** Under the assumptions of Proposition 6.9, any  $(\Phi, \alpha, \beta)$ -precise signature morphism  $\varphi$  has the definability property if  $\Phi(\varphi)$  has the definability property.

FACT 6.11. A theory morphism  $\varphi: (\Sigma, E) \rightarrow (\Sigma', E')$  is defined implicitly, respectively (finitely) explicitly, by  $E''$  in the institution of theories  $\mathcal{S}^T$  if and only if  $\varphi: \Sigma \rightarrow \Sigma'$  is defined implicitly, respectively (finitely) explicitly, by  $E' \cup E''$  in the base institution  $\mathcal{S}$ .

Consequently,  $\varphi$  has the (finite) definability property in the institution of theories if and only if it has the (finite) definability property in the base institution.

The following Corollary borrows definability results from **FOL** to **PA**. Notice that the result of 2. has already been obtained by Corollary 5.12.

- COROLLARY 6.12. 1. Any  $(i^{**})$ -morphism of signatures has the finite definability property in **PA**.  
 2. Any  $(ss^*)$ -morphism of signatures has the definability property in  $QE_1(\mathbf{PA})$  and  $QE_1^\omega(\mathbf{PA})$ .

PROOF. 1. By Corollary 4.2 any **FOL** signature morphism which is sort injective has the finite definability property, and consequently in **FOL**<sup>T</sup> too (by Fact 6.11).

We apply Corollary 6.10 to the encoding comorphism  $\mathbf{PA} \rightarrow \mathbf{FOL}^T$  of Example 6.3, which is precise (Fact 6.5) and conservative (Fact 6.8). It is also easy to see that  $\Phi$  preserves pushouts.  $\alpha$  is surjective modulo  $\models$  because it preserves the quantifications and the logical connectives, and because it is surjective on the atoms ( $\alpha(t \stackrel{e}{=} t') \models (t = t')$  for each equational  $(S, TF, \overline{PF})$ -atom and  $\alpha(\sigma(t_1, \dots, t_n) \stackrel{e}{=} t) \models \sigma(t_1, \dots, t_n, t)$  for each relational  $(S, TF, \overline{PF})$ -atom.)

2. The following argument for  $QE_1^\omega(\mathbf{PA})$  can be extended easily to  $QE_1(\mathbf{PA})$  too, hence we focus only to  $QE_1^\omega(\mathbf{PA})$ .

By Corollary 5.9 any **FOL** signature morphism which is surjective on the sorts and on the total operation symbols has the finite definability property in **HCL**, and consequently in **HCL**<sup>T</sup> too (by Fact 6.11).

Let us consider the restriction of the encoding comorphism  $\mathbf{PA} \rightarrow \mathbf{FOL}^T$  to  $QE_1^\omega(\mathbf{PA})$ . Notice that  $\Phi(S, TF, PF)$  is a **HCL**-theory for each **PA** signature  $(S, TF, PF)$ .

The crucial point of this argument is that for each  $QE_1^\omega(\mathbf{PA})$  sentence  $\rho$ ,  $\alpha(\rho)$  is semantically equivalent to a set of **HCL** sentences. In order to establish this, it is enough to establish that  $\alpha(\rho)$  is preserved by all filtered products and closed injective homomorphisms (see Example 5.2). The preservation by filtered products comes immediately as a consequence of  $\beta$ 's being isomorphisms. Now let us consider a closed injective homomorphism  $h: M \rightarrow N$  such that  $N \models \alpha(\rho)$ . We have that  $\beta(h)$  is full injective homomorphism and that  $\beta(N) \models \rho$ . Because  $QE_1^\omega$  sentences are preserved by full injective homomorphisms,  $\beta(M) \models \rho$ , hence  $M \models \alpha(\rho)$ .

Finally, concerning the surjectivity modulo  $\models$  of the sentence translations, by using the surjectivity on the atoms described at 1., it is easy to see that for each **HCL** sentence  $\rho$ , there exists a  $QE_1^\omega$  sentence  $\rho'$  such that  $\alpha(\rho') \models \rho$ .  $\dashv$

REMARK 6.13. The result of 2. of Corollary 6.12 cannot be extended to  $E_1(\mathbf{PA})$  because each **PA** signature gets encoded as a Horn theory rather than as an universal atomic theory. This obstacle in applying Corollary 6.10 to varieties of partial algebras is perfectly coherent with the obstacle mentioned in Remark 5.11 (i.e., that  $H_r \rightarrow$  does not get lifted) which in this case blocks the application of Theorem 5.6.

**§7. Conclusions.** We have generalized the concept of definability from the classical definability of a symbol to the definability of signature morphisms in arbitrary institutions. After establishing a natural general and rather mild framework in which the explicit definability implies the implicit definability, our study has focused on the hard part of the definability problem, i.e., that implicit definability implies the explicit one. We have generalized Beth theorem to institutions with Craig-Robinson interpolation. We have developed a general definability theorem in institutions supporting Birkhoff style axiomatizability properties. We have seen that the main condition setting the limits in the applications of this theorem, is in some sense the opposite of the corresponding condition underlying the interpolation via axiomatizability result of [16]; this can be regarded as an indication that interpolation cannot be used for actual definability problems in this framework.

We have illustrated the power of our general definability results with a list of applications in fragments of classical model theory and partial algebra, obtaining some definability results for (quasi-)varieties of models and partial algebras which, to our knowledge, are new. The same method can be applied to many other institutions having good Birkhoff-style axiomatizability properties.

Finally, we have developed a general result which borrows definability properties via an institution comorphism satisfying certain specific properties. By illustrating this with the example of a comorphism encoding partial algebra signatures as Horn theories in **FOL**, we have lifted Beth theorem from first order logic to partial algebra, and have also recovered the definability results for quasi-varieties of partial algebras which we had obtained before by the definability via axiomatizability result.

One future research direction concerns obtaining definability results for the multitude of computing science logics by applying our general results in the style we have illustrated with our examples here. We think this would be a rather straightforward enterprise. Another research direction concerns the extension of our definability via axiomatizability result for covering examples such as definability of operation symbols in Horn logic or of total operation symbols in quasi-varieties of partial algebras.

**Acknowledgement.** We thank the anonymous referee for checking the submitted version of our paper carefully and competently, which resulted in a correction of a subtle mathematical error and improvement of the presentation.

## REFERENCES

- [1] HAJNAL ANDRÉKA and ISTVÁN NÉMETHI, *Loś lemma holds in every category*, *Studia Scientiarum Mathematicarum Hungarica*, vol. 13 (1978), pp. 361–376.
- [2] ———, *A general axiomatizability theorem formulated in terms of cone-injective subcategories*, *Universal algebra* (B. Csakany, E. Fried, and E. T. Schmidt, editors), North-Holland, 1981, *Colloquia Mathematica Societas János Bolyai*, 29, pp. 13–35.
- [3] ———, *Generalization of the concept of variety and quasivariety to partial algebras through category theory*, *Dissertationes Mathematicae*, vol. CCIV (1983).
- [4] JAN BERGSTRA, JAN HEERING, and PAUL KLINT, *Module algebra*, *Journal of the Association for Computing Machinery*, vol. 37 (1990), no. 2, pp. 335–372.

- [5] MICHEL BIDOIT and ROLF HENNICKER, *On the integration of the observability and reachability concepts*, *Proceedings of the 5th International Conference on Foundations of Software Science and Computation Structures (FOSSACS'2002)*, Lecture Notes in Computer Science, vol. 2303, 2002, pp. 21–36.
- [6] MICHEL BIDOIT and ANDRZEJ TARLECKI, *Behavioural satisfaction and equivalence in concrete model categories*, *Proceedings of the 21st Colloquium on Trees in Algebra and Programming*, Lecture Notes in Computer Science, vol. 1059, Springer Verlag, 1996, pp. 241–256.
- [7] TOMASZ BORZYSZKOWSKI, *Higher-order logic and theorem proving for structured specifications*, *Workshop on Algebraic Development Techniques 1999* (Christine Choppy, Didier Bert, and Peter Mosses, editors), Lecture Notes in Computer Science, vol. 1827, 2000, pp. 401–418.
- [8] ———, *Generalized interpolation in CASL*, *Information Processing Letters*, vol. 76 (2001), pp. 19–24.
- [9] PETER BURMEISTER, *A model theoretic oriented approach to partial algebras*, Akademie-Verlag, Berlin, 1986.
- [10] MAURA CERIOLO and JOSÉ MESEGUER, *May I borrow your logic? (transporting logical structures along maps)*, *Theoretical Computer Science*, vol. 173 (1997), pp. 311–347.
- [11] C. C. CHANG and H. J. KEISLER, *Model theory*, North Holland, Amsterdam, 1990.
- [12] CORINA CÎRSTEĂ, *Institutionalising many-sorted coalgebraic modal logic*, *Coalgebraic Methods in Computer Science 2002*, Electronic Notes in Theoretical Computer Science, 2002.
- [13] RĂZVAN DIACONESCU, *Extra theory morphisms for institutions: logical semantics for multi-paradigm languages*, *Applied Categorical Structures*, vol. 6 (1998), no. 4, pp. 427–453, a preliminary version appeared as JAIST Technical Report IS-RR-97-0032F in 1997.
- [14] ———, *Institution-independent ultraproducts*, *Fundamenta Informaticæ*, vol. 55 (2003), no. 3–4, pp. 321–348.
- [15] ———, *Elementary diagrams in institutions*, *Journal of Logic and Computation*, vol. 14 (2004), no. 5, pp. 651–674.
- [16] ———, *An institution-independent proof of Craig Interpolation Theorem*, *Studia Logica*, vol. 77 (2004), no. 1, pp. 59–79.
- [17] ———, *Interpolation in Grothendieck institutions*, *Theoretical Computer Science*, vol. 311 (2004), pp. 439–461.
- [18] ———, *Jewels of institution-independent model theory*, *Algebra, meaning and computation (Essays dedicated to Joseph A. Goguen on the occasion of his 65th birthday)*, Lecture Notes in Computer Science, vol. 4060, Springer, 2006.
- [19] ———, *Proof systems for institutional logic*, *Journal of Logic and Computation*, vol. 16 (2006), pp. 339–357.
- [20] ———, *Institution-independent model theory*, to appear, book draft. Ask author for current draft at Razvan.Diaconescu@imar.ro.
- [21] RĂZVAN DIACONESCU, JOSEPH GOGUEN, and PETROS STEFANEAS, *Logical support for modularisation*, *Logical environments* (Gerard Huet and Gordon Plotkin, editors), Cambridge, 1993, proceedings of a workshop held in Edinburgh, Scotland, May 1991, pp. 83–130.
- [22] THEODOSIS DIMITRAKOS and TOM MAIBAUM, *On a generalized modularization theorem*, *Information Processing Letters*, vol. 74 (2000), pp. 65–71.
- [23] J. L. FIADEIRO and J. F. COSTA, *Mirror, mirror in my hand: A duality between specifications and models of process behaviour*, *Mathematical Structures in Computer Science*, vol. 6 (1996), no. 4, pp. 353–373.
- [24] DANIEL GĂINĂ and ANDREI POPESCU, *An institution-independent generalization of Tarski's Elementary Chain Theorem*, *Journal of Logic and Computation*, to appear.
- [25] ———, *An institution-independent proof of Robinson consistency theorem*, *Studia Logica*, to appear.
- [26] JOSEPH GOGUEN and ROD BURSTALL, *Institutions: Abstract model theory for specification and programming*, *Journal of the Association for Computing Machinery*, vol. 39 (1992), no. 1, pp. 95–146.
- [27] JOSEPH GOGUEN and RĂZVAN DIACONESCU, *Towards an algebraic semantics for the object paradigm*, *Recent trends in data type specification* (Harmut Ehrig and Fernando Orejas, editors), Lecture Notes in Computer Science, vol. 785, Springer, 1994, pp. 1–34.
- [28] JOSEPH GOGUEN and GRIGORE ROȘU, *Institution morphisms*, *Formal Aspects of Computing*, vol. 13 (2002), pp. 274–307.
- [29] GEORGE GRÄTZER, *Universal algebra*, Springer, 1979.

- [30] WILFRID HODGES, *Model theory*, Cambridge University Press, 1993.
- [31] JOACHIM LAMBEK and PHIL SCOTT, *Introduction to higher order categorical logic*, Cambridge Studies in Advanced Mathematics, vol. 7, Cambridge, 1986.
- [32] YNGVE LAMO, *The institution of multialgebras—a general framework for algebraic software development*, PhD thesis, University of Bergen, 2003.
- [33] SAUNDERS MAC LANE, *Categories for the working mathematician*, second ed., Springer, 1998.
- [34] G. MATTHIESSEN, *Regular and strongly finitary structures over strongly algebraic categories*, *Canadian Journal of Mathematics*, vol. 30 (1978), pp. 250–261.
- [35] JOSÉ MESEGUER, *General logics*, *Logic Colloquium, 1987* (H.-D. Ebbinghaus et al., editors), North-Holland, 1989, pp. 275–329.
- [36] ———, *Conditional rewriting logic as a unified model of concurrency*, *Theoretical Computer Science*, vol. 96 (1992), no. 1, pp. 73–155.
- [37] ———, *Membership algebra as a logical framework for equational specification*, *Proceedings of the WADT'97* (F. Parisi-Presicce, editor), Lecture Notes in Computer Science, no. 1376, Springer, 1998, pp. 18–61.
- [38] TILL MOSSAKOWSKI, *Specification in an arbitrary institution with symbols*, *Recent trends in Algebraic Development Techniques, 14th International Workshop, WADT'99, Bonas, France* (C. Choppy, D. Bert, and P. Mosses, editors), Lecture Notes in Computer Science, vol. 1827, Springer-Verlag, 2000, pp. 252–270.
- [39] ———, *Relating CASL with other specification languages: the institution level*, *Theoretical Computer Science*, vol. 286 (2002), pp. 367–475.
- [40] TILL MOSSAKOWSKI, JOSEPH GOGUEN, RĂZVAN DIACONESCU, and ANDRZEJ TARLECKI, *What is a logic?*, *Logica universalis* (Jean-Yves Beziau, editor), Birkh pp. 113–133.
- [41] ISTVÁN NÉMETHI and ILDIKÓ SAIN, *Cone-implicational subcategories and some Birkhoff-type theorems*, *Universal algebra* (B. Csakany, E. Fried, and E. T. Schmidt, editors), North-Holland, 1981, *Colloquia Mathematica Societas János Bolyai*, 29, pp. 535–578.
- [42] ANDREI POPESCU, TRAIAN ȘERBĂNUȚĂ, and GRIGORE ROȘU, *A semantic approach to interpolation*, submitted.
- [43] PIETER-HENDRIK RODENBURG, *A simple algebraic proof of the equational interpolation theorem*, *Algebra Universalis*, vol. 28 (1991), pp. 48–51.
- [44] DONALD SANNELLA and ANDRZEJ TARLECKI, *Specifications in an arbitrary institution*, *Information and Control*, vol. 76 (1988), pp. 165–210, earlier version in *Proceedings, International Symposium on the Semantics of Data Types*, Lecture Notes in Computer Science, vol. 173, Springer, 1985.
- [45] LUTZ SCHRÖDER, TILL MOSSAKOWSKI, and CHRISTOPH LÜTH, *Type class polymorphism in an institutional framework*, *Recent trends in Algebraic Development Techniques, 17th International Workshop (WADT 2004)* (José Fiadeiro, editor), Lecture Notes in Computer Science, vol. 3423, Springer, Berlin, 2004, pp. 234–248.
- [46] JOSEPH SHOENFIELD, *Mathematical logic*, Addison-Wesley, 1967.
- [47] ANDRZEJ TARLECKI, *Bits and pieces of the theory of institutions*, *Proceedings, Summer Workshop on Category Theory and Computer Programming* (David Pitt, Samson Abramsky, Axel Poigné, and David Rydeheard, editors), Lecture Notes in Computer Science, vol. 240, Springer, 1986, pp. 334–360.
- [48] ———, *On the existence of free models in abstract algebraic institutions*, *Theoretical Computer Science*, vol. 37 (1986), pp. 269–304, preliminary version, University of Edinburgh, Computer Science Department, Report CSR-165-84, 1984.
- [49] ———, *Quasi-varieties in abstract algebraic institutions*, *Journal of Computer and System Sciences*, vol. 33 (1986), no. 3, pp. 333–360, original version, University of Edinburgh, Report CSR-173-84.
- [50] ———, *Moving between logical systems*, *Recent trends in data type specification* (Magne Haveraaen, Olaf Owe, and Ole-Johan Dahl, editors), Lecture Notes in Computer Science, Springer, 1996, pp. 478–502.
- [51] ———, *Towards heterogeneous specifications*, *Proceedings, International Conference on Frontiers of Combining Systems (FroCoS'98)* (D. Gabbay and M. van Rijke, editors), Research Studies Press, 2000, pp. 337–360.
- [52] PAULO VELOSO, *On pushout consistency, modularity and interpolation for logical specifications*, *Information Processing Letters*, vol. 60 (1996), no. 2, pp. 59–66.

ȘCOALA NORMALĂ SUPERIOARĂ  
CALEA GIVIȚEI 21  
BUCHAREST 010702, ROMANIA  
*Current address:* University of Edinburgh  
*E-mail:* mariuspyah@yahoo.com

INSTITUTE OF MATHEMATICS "SIMION STOILOW"  
PO BOX 1-764  
BUCHAREST 014700, ROMANIA  
*E-mail:* Razvan.Diaconescu@imar.ro