THE JOURNAL OF SYMBOLIC LOGIC
Volume 71, Number 3, Sept. 2006

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.

0022-4812/06/7103-0016/$3.70

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 π with respect to a given signature Σ, which determines a signature inclusion $\Sigma \hookrightarrow \Sigma \cup \{\pi\}$ is generalised and abstracted to any signature morphism $\varphi\colon \Sigma \to \Sigma'$ 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 $\mathbb{S}et$, and $\mathbb{CAT}$ is the category of all categories.[1] The opposite of a category $\mathbb{C}$ (obtained by reversing the arrows of $\mathbb{C}$) is denoted $\mathbb{C}^{\mathrm{op}}$.

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

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[dr, "f'"'] & B \arrow[d, "h"] \\ & B' \end{tikzcd}$$

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

$$\begin{tikzcd} \bullet \arrow[r, "u"] \arrow[d] & \bullet \arrow[d] \\ \bullet \arrow[r, "u'"'] & \bullet \end{tikzcd}$$

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

Institutions. An *institution* $\mathcal{I} = (\mathbb{S}ig^{\mathcal{I}}, \mathsf{Sen}^{\mathcal{I}}, \mathsf{Mod}^{\mathcal{I}}, \models^{\mathcal{I}})$ consists of

1. a category $\mathbb{S}ig^{\mathcal{I}}$, whose objects are called *signatures*,
2. a functor $\mathsf{Sen}^{\mathcal{I}}\colon \mathbb{S}ig^{\mathcal{I}} \to \mathbb{S}et$, giving for each signature a set whose elements are called *sentences* over that signature,

[1] Strictly speaking, this is only a quasi-category living in a higher set-theoretic universe.

3. a functor $\mathsf{Mod}^{\mathscr{I}}\colon (\mathbb{S}ig^{\mathscr{I}})^{\mathrm{op}} \to \mathbb{CAT}$ giving for each signature Σ a category whose objects are called Σ-*models*, and whose arrows are called Σ-(*model*) *morphisms*, and
4. a relation $\models^{\mathscr{I}}_{\Sigma} \subseteq |\mathsf{Mod}^{\mathscr{I}}(\Sigma)| \times \mathsf{Sen}^{\mathscr{I}}(\Sigma)$ for each $\Sigma \in |\mathbb{S}ig^{\mathscr{I}}|$, called Σ-*satisfaction*,

such that for each morphism $\varphi\colon \Sigma \to \Sigma'$ in $\mathbb{S}ig^{\mathscr{I}}$, the *satisfaction condition*

$$M' \models^{\mathscr{I}}_{\Sigma'} \mathsf{Sen}^{\mathscr{I}}(\varphi)(\rho) \text{ iff } \mathsf{Mod}^{\mathscr{I}}(\varphi)(M') \models^{\mathscr{I}}_{\Sigma} \rho$$

holds for each $M' \in |\mathsf{Mod}^{\mathscr{I}}(\Sigma')|$ and $\rho \in \mathsf{Sen}^{\mathscr{I}}(\Sigma)$. We denote the *reduct* functor $\mathsf{Mod}^{\mathscr{I}}(\sigma)$ by $\_\restriction_{\varphi}$ and the sentence translation $\mathsf{Sen}^{\mathscr{I}}(\varphi)$ by $\varphi(\_)$. When $M = M'\restriction_{\varphi}$ we say that M is a φ-*reduct of* M', and that M' is a φ-*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 $\mathbb{S}ig^{\mathscr{I}}$ may be simply denoted $\mathbb{S}ig$.

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

In any institution, a signature morphism $\varphi\colon \Sigma \to \Sigma'$ is *conservative* when each Σ-model has at least one φ-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. $F_{w\to s}$ 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 σ as a function M_{σ} from the product of the interpretations of the argument sorts to the interpretation of the result sort, and each relation symbol π as a subset M_{π} 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 φ, the reduct $M'\restriction_{\varphi}$ of a model M' is defined by $(M'\restriction_{\varphi})_x = M'_{\varphi(x)}$ for each x sort, function, or relation symbol from the domain signature of φ. 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 of 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 $\forall\vee$, its infinitary extension by $\forall\vee_{\infty}$.

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* $h \colon A \to B$ is a family of (total) functions $\{h_s \colon A_s \to B_s\}_{s\in S}$ 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 $\sigma \colon w \to s$ 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.[2] 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',[3] 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 $x, y, x \in \{i, s, b, *\}$ (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.

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

[3] 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 Σ in an institution $\mathscr{I}$, a Σ-*theory* is any set of Σ-sentences.

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

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

Two sentences, ρ_1 and ρ_2 of the same signature are *semantically equivalent*, denoted $\models\!\mid$ 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* $\varphi\colon (\Sigma, E) \to (\Sigma', E')$ is just a signature morphism $\varphi\colon \Sigma \to \Sigma'$ such that $E' \models \varphi(E)$. The *institution* $\mathscr{I}^T$ *of* $\mathscr{I}$*-theories* has the category of theories $\mathbb{T}h^{\mathscr{I}}$ of $\mathscr{I}$ as its category of signatures, $\mathsf{Sen}^{\mathscr{I}^T}(\Sigma, E) = \mathsf{Sen}^{\mathscr{I}}(\Sigma)$, and $\mathsf{Mod}^{\mathscr{I}^T}(\Sigma, E)$ is the full subcategory of $\mathsf{Mod}^{\mathscr{I}}$ consisting of the Σ-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 $\mathscr{I}$ is *exact* if and only if the model functor $\mathsf{Mod}^{\mathscr{I}}\colon (\mathbb{S}ig^{\mathscr{I}})^{\mathrm{op}} \to \mathbb{CAT}$ preserves finite limits. The institution is *semi-exact* if and only if $\mathsf{Mod}^{\mathscr{I}}$ 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[4] version of exactness suffices [13, 51, 39].

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

$$\begin{tikzcd} \Sigma \arrow[r, "\varphi_1"] \arrow[d, "\varphi_2"'] & \Sigma_1 \arrow[d, "\theta_1"] \\ \Sigma_2 \arrow[r, "\theta_2"'] & \Sigma' \end{tikzcd}$$

[4] In the sense of 'weak' universal properties [33] not requiring uniqueness.

is an *amalgamation square* if and only if for each Σ_1-model M_1 and a Σ_2-model M_2 such that $M_1\upharpoonright_{\varphi_1} = M_2\upharpoonright_{\varphi_2}$, there exists an unique Σ'-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 $\mathscr{I}$ has *elementary diagrams* [15] iff for each signature Σ and each Σ-model M, there exists a signature morphism $\imath_\Sigma(M)\colon \Sigma \to \Sigma_M$, "functorial" in Σ and M, and a set E_M of Σ_M-sentences such that $\mathsf{Mod}(\Sigma_M, E_M)$ and the comma category $M/\mathsf{Mod}(\Sigma)$ are naturally isomorphic, i.e., the following diagram commutes by the isomorphism $i_{\Sigma,M}$ "natural" in Σ and M

$$\begin{tikzcd}
\mathsf{Mod}(\Sigma_M, E_M) \arrow[r, "i_{\Sigma,M}"] \arrow[dr, "\mathsf{Mod}(\imath_\Sigma(M))"'] & (M/\mathsf{Mod}(\Sigma)) \arrow[d, "forgetful"] \\
& \mathsf{Mod}(\Sigma)
\end{tikzcd}$$

The signature morphism $\imath_\Sigma(M)\colon \Sigma \to \Sigma_M$ is called the *elementary extension of* Σ *via* M and the set E_M of Σ_M-sentences is called the *elementary diagram* of the model M. Note that $i^{-1}_{\Sigma,M}(1_M)$ is the initial model of (Σ_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 $(F_M)_{\to s} = F_{\to s} \cup M_s$, otherwise let $(F_M)_{w\to s} = F_{w\to s}$, 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 $h\colon M \to N$ 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 **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 **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 **FOL**-signature (S, F, P), a new relation symbol π is 'implicitly' defined by a theory E if and only if it is 'explicitly' defined by the same theory. π is implicitly defined when the forgetful reduct $\mathsf{Mod}^{\mathbf{FOL}}((S, F, P \uplus \{\pi\}), E) \to \mathsf{Mod}^{\mathbf{FOL}}(S, F, P)$ 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 π' of the same arity and where $E[\pi/\pi']$ is the copy of E in which π is replaced by π', while π is explicity defined if π can be 'defined' by an $(S, F \uplus X, P)$-sentence E_π, 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 π.

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 Σ in an arbitrary institution, for any Σ-sentences ρ_1 and ρ_2, a Σ-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 $\chi\colon \Sigma \to \Sigma'$ in an arbitrary institution, for any Σ'-sentence ρ and any Σ-model M, we say that M *satisfies* $(\forall\chi)\rho$, denoted by $M \models (\forall\chi)\rho$, if and only if each χ-expansion of M satisfies ρ in the institution. The institution *has universal $\mathscr{D}$-quantification* for a class $\mathscr{D}$ of signature morphisms, when for each $(\chi\colon \Sigma \to \Sigma') \in \mathscr{D}$ and each $\rho \in \mathsf{Sen}(\Sigma')$ there exists a Σ-sentence semantically equivalent to $(\forall\chi)\rho$.[5] Notice that the concept of 'internal quantification' captures ordinary quantification of the actual institutions, for example **FOL** has $\mathscr{D}$-quantification for $\mathscr{D}$ the class of signature extensions with a finite number of constants, while in the case of second order logic $\mathscr{D}$ 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 $\varphi\colon \Sigma \to \Sigma'$ be a signature morphism and E' be a Σ'-theory. Then φ

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

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

where

$$\begin{tikzcd} \Sigma \arrow[r, "\varphi"] \arrow[d, "\theta"'] & \Sigma' \arrow[d, "\theta'"] \\ \Sigma_1 \arrow[r, "\varphi_1"'] & \Sigma_1' \end{tikzcd}$$

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

REMARK 3.2. Note that E_ρ 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 θ to belong to a class $\mathscr{D}$ of signature morphisms stable under pushouts such that the institution has universal $\mathscr{D}$-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 θ *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

[5] Existential quantification is defined similarly.

symbol π we have the atom $\pi(X)$. Similarly, in **PA** for each partial operation symbol σ, we have the atom $def(\sigma(X))$. This means that explicit definability ensures a *uniform* elimination of the symbol π 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 ι, a signature morphism $\varphi\colon \Sigma \to \Sigma'$ is *ι-tight* when for all Σ'-models M' and N' with a common φ-reduct, $M' \otimes M_M \equiv N' \otimes N_N$ implies $M' = N'$ (where $M = M'\restriction_\varphi = N'\restriction_\varphi = N$).

$$\begin{tikzcd} \Sigma \arrow[r, "\varphi"] \arrow[d, "\iota_\Sigma(M)"'] & \Sigma' \arrow[d, "\theta'"] \\ \Sigma_M \arrow[r, "\varphi_1"'] & \Sigma'_1 \end{tikzcd}$$

EXAMPLE 3.5. Consider the classical situation when φ is a signature morphism in **FOL** adding one relation symbol π. Then the only possible difference between M' and N' could only be found in the difference between M'_π and N'_π. 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 φ 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 $\varphi\colon \Sigma \to \Sigma'$ be a ι-tight signature morphism in a semi-exact institution with elementary diagrams ι. Then any two Σ'-models which are isomorphic by a φ-expansion of an identity, are equal.

PROOF. Let $h\colon M' \to N'$ be a Σ'-isomorphism such that $h\restriction_\varphi$ is identity. Let $M = M'\restriction_\varphi$ and $N = N'\restriction_\varphi$. 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 φ being tight, we get that $M' = N'$. $\dashv$

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

Proof. The surjectivity on the sorts is necessary because otherwise, given a Σ'-model M' we may consider another Σ'-model N' which is like M' but interprets the sorts outside image of the tight signature morphism $\varphi\colon \Sigma \to \Sigma'$ differently but isomorphically to M'. This gives a Σ'-isomorphism expanding a Σ-identity between different Σ'-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 Σ'-models M', N' with $(M =)M'\restriction_\varphi = N'\restriction_\varphi (= N)$, then for all operation symbols $(\sigma\colon w \to s) \in \Sigma'$ 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 ι, each ι-tight signature morphism is defined implicitly whenever it is defined explicitly.

Proof. Let $\varphi\colon \Sigma \to \Sigma'$ be a tight signature morphism which is explicitly defined by $E' \subseteq \mathsf{Sen}(\Sigma')$. We show that φ is defined implicitly by E'. Let $M', N' \in |\mathsf{Mod}(\Sigma', E')|$ with $M'\restriction_\varphi = N'\restriction_\varphi$.

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

Let $M' \otimes M_M \models \rho$. Because φ is explicitly defined by E', there exists $E_\rho \subseteq \mathsf{Sen}(\Sigma_M)$ 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 θ'-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 $\iota_\Sigma(M)\colon \Sigma \to \Sigma_M$ and the existence of M_M as a 'canonical' $\iota_\Sigma(M)$-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 φ *has the (finite) definability property* iff a theory defines φ (finitely) explicitly whenever it defines φ implicitly.

Before focusing on various methods for obtaining the definability property, let us give without proof[6] 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 $\mathscr{D}$-quantification for a class $\mathscr{D}$ of signature morphisms which is stable under pushouts, for any pushout square of signature morphisms

$$\begin{tikzcd} \Sigma \arrow[r, "\varphi"] \arrow[d, "\theta"'] & \Sigma' \arrow[d, "\theta'"] \\ \Sigma_1 \arrow[r, "\varphi_1"'] & \Sigma_1' \end{tikzcd}$$

such that $\theta \in \mathscr{D}$ and is conservative, φ has the definability property with respect to E' whenever φ_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 $\mathscr{L}$ and $\mathscr{R}$ of signature morphisms in an institution $\mathscr{I}$, the institution has the *Craig-Robinson* $\langle\mathscr{L}, \mathscr{R}\rangle$*-interpolation property*, if for any pushout in $\mathbb{S}ig$ such that $\varphi_1 \in \mathscr{L}$ and $\varphi_2 \in \mathscr{R}$, any set of Σ_1-sentences E_1 and any sets of Σ_2-sentences E_2 and Γ_2 with $\theta_1(E_1) \cup \theta_2(\Gamma_2) \models \theta_2(E_2)$ there exists a set of Σ-sentences E (called the *interpolant*) such that $E_1 \models \varphi_1(E)$ and $\varphi_2(E) \cup \Gamma_2 \models E_2$.

$$\begin{tikzcd} \Sigma \arrow[r, "\varphi_1"] \arrow[d, "\varphi_2"'] & \Sigma_1 \arrow[d, "\theta_1"] \\ \Sigma_2 \arrow[r, "\theta_2"'] & \Sigma' \end{tikzcd}$$

The restriction given by Γ_2 being empty is called *Craig* $\langle\mathscr{L}, \mathscr{R}\rangle$*-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 $\mathscr{R}$ is the class of (iii)-morphisms [43, 16].

[6]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 Γ_2 (of Σ_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 $(\mathscr{L}, \mathscr{R})$-interpolation for classes $\mathscr{L}$ and $\mathscr{R}$ of signature morphisms which are stable under pushouts, any signature morphism in $\mathscr{L} \cap \mathscr{R}$ has the (finite) definability property.

Proof. Let $(\varphi\colon \Sigma \to \Sigma') \in \mathscr{L} \cap \mathscr{R}$ be defined implicitly by $E' \subseteq \mathsf{Sen}(\Sigma')$. We consider the pushout of φ with an arbitrary signature morphism $\theta\colon \Sigma \to \Sigma_1$ and a Σ'_1-sentence ρ.

$$\begin{tikzcd} \Sigma \arrow[r, "\varphi"] \arrow[d, "\theta"'] & \Sigma' \arrow[d, "\theta'"] \\ \Sigma_1 \arrow[r, "\varphi_1"'] & \Sigma'_1 \end{tikzcd}$$

Now we consider the pushout of φ_1 with itself:

$$\begin{tikzcd} \Sigma_1 \arrow[r, "\varphi_1"] \arrow[d, "\varphi_1"'] & \Sigma'_1 \arrow[d, "\theta_1"] \\ \Sigma'_1 \arrow[r, "\theta_2"'] & \Sigma'' \end{tikzcd}$$

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

Now because $\varphi \in \mathscr{L} \cap \mathscr{R}$ and $\mathscr{L}$ and $\mathscr{R}$ are stable under pushouts, we have that $\varphi_1 \in \mathscr{L} \cap \mathscr{R}$, and by Craig-Robinson interpolation (and compactness) there exists (finite) $E_\rho \subseteq \mathsf{Sen}(\Sigma_1)$ 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 § be the class of $(i{*}{*})$-morphisms of signatures. From [8, 17, 25] we know that **FOL** has Craig $\langle §, §\rangle$-interpolation, hence it has Craig-Robinson $\langle §, §\rangle$-interpolation (because **FOL** has implications and is compact; see [22]). ⊣

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 $D\colon J \to \mathbb{C}$ 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 $A_F\colon F \to \mathbb{C}$ such that $A_F(J \subset J') = p_{J',J}\colon \prod_{i\in J'} A_i \to \prod_{i\in J} A_i$ 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 $\mu\colon A_F \Rightarrow \prod_F A_i$ of the functor A_F.

$$\begin{tikzcd} \prod_{i\in J'} A_i \arrow[rr, "p_{J',J}"] \arrow[dr, "\mu_{J'}"'] & & \prod_{i\in J} A_i \arrow[dl, "\mu_J"] \\ & \prod_F A_i & \end{tikzcd}$$

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].[7]

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

Birkhoff institutions. Recall from [16] that $(\mathbb{S}ig, \mathsf{Sen}, \mathsf{Mod}, \models, \mathscr{F}, \mathscr{B})$ is a *Birkhoff institution* if and only if

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

such that

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

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

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

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

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	$\mathscr{B}$	$\mathscr{F}$
FOL	$\equiv$	all ultrafilters
FOL	*ultraradicals* (see [42])	all ultrafilters
PL	$=$	all ultrafilters
universal (or quantifier-free) **FOL**-sentences	$\xrightarrow{S_c}$	all ultrafilters
universal $\mathbf{FOL}_{\infty,\omega}$-sentences	$\xrightarrow{S_c}$	$\{\{\{*\}\}\}$
$\mathbf{HCL}_\infty$	$\xrightarrow{S_c}$	$\{\{I\} \mid I \text{ set}\}$
HCL	$\xrightarrow{S_c}$	all filters
universal **FOL**-atoms	$\xleftarrow{H_r}; \xrightarrow{S_c}$	$\{\{I\} \mid I \text{ set}\}$
EQL	$\xleftarrow{H_r}; \xrightarrow{S_w}$	$\{\{I\} \mid I \text{ set}\}$
$\forall\vee$	$\xleftarrow{H_s}; \xrightarrow{S_c}$	all ultrafilters
$\forall\vee_\infty$	$\xleftarrow{H_s}; \xrightarrow{S_c}$	$\{\{\{*\}\}\}$
$\forall\exists$ (universal-existential **FOL**-sentences)	*sandwiches* (see [11])	all ultrafilters.

where a model homomorphism $h \colon M \to N$ 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

[8] 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[9] 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	$\mathscr{B}$	$\mathscr{F}$
PA	$\equiv$	all ultrafilters
universal **PA**-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 $h\colon A \to B$ (i.e., $A_\sigma(a)$ is defined if $B_\sigma(h(a))$ is defined), S_f of all *full* injective homomorphisms $h\colon A \to B$ (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 $R = \{R_\Sigma \subseteq |\mathsf{Mod}(\Sigma)| \times |\mathsf{Mod}(\Sigma)|\}_{\Sigma \in |\mathbb{S}ig|}$ indexed by the category of the signatures of an institution, a signature morphism $\varphi\colon \Sigma \to \Sigma'$

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

[9] 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 $(\mathbb{S}ig, \mathsf{Sen}, \mathsf{Mod}, \models, \mathscr{F}, \mathscr{B})$ and a class $\mathcal{S} \subseteq \mathbb{S}ig$ of signature morphisms which is stable under pushouts and such that for each $\varphi \in \mathcal{S}$

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

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

PROOF. Let $\varphi \in \mathcal{S}$. If $\varphi\colon \Sigma \to \Sigma'$ 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 $\Sigma_1 \xleftarrow{\theta} \Sigma \xrightarrow{\varphi} \Sigma'$

$$\begin{tikzcd} \Sigma \arrow[r, "\varphi"] \arrow[d, "\theta"'] & \Sigma' \arrow[d, "\theta'"] \\ \Sigma_1 \arrow[r, "\varphi_1"'] & \Sigma_1' \end{tikzcd}$$

and any $\rho \in \mathsf{Sen}(\Sigma_1')$.

By the hypotheses on the Birkhoff institution we have that φ_1 lifts weakly $\mathscr{B}^{-1}$ and preserves filtered products. Let us denote $\mathsf{Mod}(\Sigma_1', \theta'(E') \cup \{\rho\})$ by $\mathbb{M}_1'$. We define E_ρ as $(\mathbb{M}_1'\restriction_{\varphi_1})^*$.

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 $M_1'\restriction_{\varphi_1} \in \mathbb{M}_1'\restriction_{\varphi_1}$ and because E_ρ is satisfied by all models in $\mathbb{M}_1'\restriction_{\varphi_1}$ we have that $M_1'\restriction_{\varphi_1} \models E_\rho$. 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 Σ_1'-model satisfying $\theta'(E') \cup \varphi_1(E_\rho)$. By the Satisfaction Condition we have that $M_1'\restriction_{\varphi_1} \models E_\rho = (\mathbb{M}_1'\restriction_{\varphi_1})^*$. Because of the conditions on our Birkhoff institution $M_1'\restriction_{\varphi_1} \in (\mathbb{M}_1'\restriction_{\varphi_1})^{**} = \mathscr{B}_{\Sigma_1}^{-1}(\mathscr{F}(\mathbb{M}_1'\restriction_{\varphi_1}))$. By considering the following:

- $\mathscr{F}(\mathbb{M}_1'\restriction_{\varphi_1}) = \cong_{\Sigma_1}((\mathscr{F}\mathbb{M}_1')\restriction_{\varphi_1})$ because φ_1 preserves filtered products,
- $\cong_{\Sigma_1}; \mathscr{B}_{\Sigma_1}^{-1} = \mathscr{B}_{\Sigma_1}^{-1}$ because $\mathscr{B}$ is closed under isomorphisms,
- $\mathscr{F}\mathbb{M}_1' \subseteq \mathscr{B}_{\Sigma_1'}^{-1}(\mathscr{F}\mathbb{M}_1')$ because $\mathscr{B}$ is reflexive, and
- $\mathscr{B}_{\Sigma_1'}^{-1}(\mathscr{F}\mathbb{M}_1') = \mathbb{M}_1'$ because $\mathbb{M}_1'$ is elementary.

it results that

$$\begin{aligned} M_1'\restriction_{\varphi_1} \in \mathscr{B}_{\Sigma_1}^{-1}(\mathscr{F}(\mathbb{M}_1'\restriction_{\varphi_1})) &= \mathscr{B}_{\Sigma_1}^{-1}(\cong_{\Sigma_1}((\mathscr{F}\mathbb{M}_1')\restriction_{\varphi_1})) \\ &= \mathscr{B}_{\Sigma_1}^{-1}((\mathscr{F}\mathbb{M}_1')\restriction_{\varphi_1}) \subseteq \mathscr{B}_{\Sigma_1}^{-1}(\mathbb{M}_1'\restriction_{\varphi_1}). \end{aligned}$$

This implies that there exists a Σ_1'-model N_1' satisfying $\theta'(E') \cup \{\rho\}$ and such that $\langle M_1'\restriction_{\varphi_1}, N_1'\restriction_{\varphi_1}\rangle \in \mathscr{B}_{\Sigma_1}$. Because φ_1 lifts $\mathscr{B}^{-1}$ it exists a Σ_1'-model P_1' such that $P_1'\restriction_{\varphi_1} = M_1'\restriction_{\varphi_1}$ and $\langle P_1', N_1'\rangle \in \mathscr{B}_{\Sigma_1'}$.

Because $\{\{\{*\}\}\} \in \mathscr{F}$ we have that $\mathscr{B}_{\Sigma_1'}^{-1}(\mathbb{M}_1') \subseteq \mathscr{B}_{\Sigma_1'}^{-1}(\mathscr{F}\mathbb{M}_1') = \mathbb{M}_1'$. From $P_1' \in \mathscr{B}_{\Sigma_1'}^{-1}(N_1') \subseteq \mathscr{B}_{\Sigma_1'}^{-1}(\mathbb{M}_1')$ 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 φ 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_ρ can be chosen finite. Thus $\theta'(E') \models \rho \Leftrightarrow \varphi_1(E_\rho)$, which implies that $E' \models (\forall\theta')(\rho \Leftrightarrow \varphi_1(E_\rho))$. ⊣

Remark 5.7. This definability result relies primarily on the Birkhoff-style axiomatizability property of the institution. Secondarily, 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 $\xleftarrow{S_w}$, $\xleftarrow{S_c}$, $\xrightarrow{H_r}$, and $\xrightarrow{H_s}$, and any $(ss*)$-morphism of signatures lifts weakly $\xleftarrow{S_w}$ and $\xleftarrow{S_c}$.

Proof. Assume $\varphi\colon (S, F, P) \to (S', F', P')$ is a (bbi)-morphism.

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

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

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

Let $h\colon N'\upharpoonright_{\varphi} \to M'\upharpoonright_{\varphi}$ 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 $h'\colon Q' \to M'$ 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 $h\colon N'\upharpoonright_{\varphi} \to M'\upharpoonright_{\varphi}$ is closed, then $h\colon Q' \hookrightarrow M'$ is closed too. ⊣

Corollary 5.9. We have the foolowing table of definability results:

institution	signature morphism	definability property
HCL	$ss*$	finite definability
$\mathbf{HCL}_\infty$	$ss*$	definability
universal **FOL**-sentences	$ss*$	finite definability
universal $\mathbf{FOL}_{\infty,\omega}$-sentences	$ss*$	definability
universal **FOL**-atoms	bbi	finite definability
$\forall\vee$	bbi	finite definability
$\forall\vee_\infty$	bbi	definability

Proof. From Example 5.2 and Theorem 5.6, because the composition of a relation lifted weakly by φ with a relation lifted by φ gets a relation lifted weakly by φ, 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 $\stackrel{S_w}{\leftarrow}$ and $\stackrel{S_f}{\leftarrow}$ and any $(ss*)$-morphism of signatures lifts weakly $\stackrel{S_w}{\leftarrow}$ and $\stackrel{S_f}{\leftarrow}$.

Proof. Assume $\varphi\colon (S, TF, PF) \to (S', TF', PF')$ is a (bbi)-morphism. Let $h\colon B \to A'\upharpoonright_\varphi$ be an injective (S, TF, PF)-algebra homomorphism, i.e., $h \in S_w$. We define B' to be the unique φ-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 $h'\colon B' \to A'$, the unique φ-expansion of h, is an injective (S', TF', PF')-algebra homomorphism. Moreover, if $h\colon B \to A'\upharpoonright_\varphi$ is full, then $h'\colon B' \to A'$ is full too.

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

Let $h\colon B'\upharpoonright_\varphi \to A'\upharpoonright_\varphi$ 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'_σ 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 $\sigma \in PF'_{\varphi(w)\to\varphi(s)} \setminus \varphi(PF_{w\to s})$ and $c \in C'_{\varphi(w)}$ such that $A'_\sigma(c) \in h(C'_{\varphi(s)})$. Then $h'\colon C' \to A'$ 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 $h\colon B'\upharpoonright_\varphi \to A'\upharpoonright_\varphi$ is full, we get that $h'\colon C' \to A'$ is full too. $\dashv$

Remark 5.11. The (bbi)-morphisms do not lift weakly neither the closed subalgebra relation $\stackrel{S_c}{\leftarrow}$ nor $\stackrel{H_r}{\rightarrow}$.

Corollary 5.12. We have the foolowing 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. ⊣

REMARK 5.13. While $E_1(\mathbf{PA})$, $QE_1(\mathbf{PA})$ and $QE_1^{\omega}(\mathbf{PA})$ have the elementary diagrams of **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. Howvever, 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* $(\Phi, \alpha, \beta)\colon \mathscr{I} \to \mathscr{I}'$ consists of

1. a functor $\Phi\colon \mathbb{S}ig \to \mathbb{S}ig'$,
2. a natural transformation $\alpha\colon \mathsf{Sen} \Rightarrow \Phi; \mathsf{Sen}'$, and
3. a natural transformation $\beta\colon \Phi^{\mathrm{op}}; \mathsf{Mod}' \Rightarrow \mathsf{Mod}$

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 $\Sigma \in |\mathbb{S}ig|$, for each $\Phi(\Sigma)$-model M', and each Σ-sentence e.

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

EXAMPLE 6.2. **EQL** can 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 $\overline{PF}_{w \to s} = PF_{w \to s}$ 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 $h\colon M \to N$ is a (S, TF, PF)-algebra homomorphism too.)
- α 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
$$bind(\sigma(t_1 \dots t_n), x) = \bigwedge_{1 \le i \le n} bind(t_i, x_i) \wedge \begin{cases} \sigma(x_1, \dots, x_n) = x & \text{when } \sigma \in TF, \\ \sigma(x_1, \dots, x_n, x) & \text{when } \sigma \in PF \end{cases}$$
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 $\mathbf{PA} \to \mathbf{FOL}^T$ 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 $(\Phi, \alpha, \beta)\colon \mathscr{I} \to \mathscr{I}'$ be an institution comorphism. A $\mathscr{I}$-signature morphism $\varphi\colon \Sigma_1 \to \Sigma_2$ is (Φ, α, β)-*precise* whenever the function $\mathsf{Mod}'(\Phi(\Sigma_2)) \to \mathsf{Mod}'(\Phi(\Sigma_1)) \times \mathsf{Mod}(\Sigma_2)$ mapping each M'_2 to $\langle M'_2\restriction_{\Phi(\varphi)}, \beta_{\Sigma_2}(M'_2)\rangle$ is injective.

The comorphism (Φ, α, β) is *precise* when each $\mathscr{I}$-signature morphism is (Φ, α, β)-precise.

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

Proposition 6.6. Let $(\Phi, \alpha, \beta)\colon \mathscr{I} \to \mathscr{I}'$ be an institution comorphism. Then for any (Φ, α, β)-precise signature morphism φ and theory E', $\Phi(\varphi)$ is defined implicitly by $\alpha(E')$ if φ is defined implicitly by a E'.

Proof. Let $\varphi\colon \Sigma \to \Sigma'$ be a (Φ, α, β)-precise signature morphism.

Assume φ is defined implicitly by a E', and let $M'_1, M'_2 \in |\mathsf{Mod}'(\Phi(\Sigma'), \alpha(E'))|$ such that $M'_1\restriction_{\Phi(\varphi)} = M'_2\restriction_{\Phi(\varphi)}$. Because φ is (Φ, α, β)-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 (Φ, α, β), $\beta_{\Sigma'}(M'_1), \beta_{\Sigma'}(M'_2) \models E'$, and $\beta_{\Sigma'}(M'_1)\restriction_\varphi = \beta_\Sigma(M'_1\restriction_{\Phi(\varphi)}) = \beta_\Sigma(M'_2\restriction_{\Phi(\varphi)}) = \beta_{\Sigma'}(M'_2)\restriction_\varphi$ by the naturality of β. Because φ is defined implicitly by E', we obtain that $\beta_{\Sigma'}(M'_1) = \beta_{\Sigma'}(M'_2)$. ⊣

Definition 6.7. An institution comorphism $(\Phi, \alpha, \beta)\colon \mathscr{I} \to \mathscr{I}'$ is *conservative* when for each $\mathscr{I}$-signature Σ, each Σ-model has at least one β_Σ-expansion, i.e., β_Σ is surjective.

Fact 6.8. The comorphisms $\mathbf{EQL} \to \mathbf{FOL}$, $\mathbf{EQL} \to \mathbf{PA}$ and $\mathbf{PA} \to \mathbf{FOL}^T$ are trivially conservative.

Proposition 6.9. Let $(\Phi, \alpha, \beta)\colon \mathscr{I} \to \mathscr{I}'$ be a conservative institution comorphism such that Φ preserves pushouts and α is surjective modulo the semantic equivalence $\models\!\mid$.

Then any $\mathscr{I}$-signature morphism φ 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

$$\begin{tikzcd} \Sigma \arrow[r, "\varphi"] \arrow[d, "\theta"'] & \Sigma' \arrow[d, "\theta'"] \\ \Sigma_1 \arrow[r, "\varphi_1"'] & \Sigma_1' \end{tikzcd}$$

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

Because Φ preserves pushouts we have that

$$\begin{tikzcd} \Phi(\Sigma) \arrow[r, "\Phi(\varphi)"] \arrow[d, "\Phi(\theta)"'] & \Phi(\Sigma') \arrow[d, "\Phi(\theta')"] \\ \Phi(\Sigma_1) \arrow[r, "\Phi(\varphi_1)"'] & \Phi(\Sigma_1') \end{tikzcd}$$

is a pushout in $\mathbb{S}ig'$.

Because $\Phi(\varphi)$ is defined (finitely) explicitly by $\alpha_{\Sigma'}(E')$, there exists (finite) $E_{\alpha_{\Sigma_1'}(\rho)} \subseteq \mathsf{Sen}'(\Phi(\Sigma_1))$ 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_ρ 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_ρ is chosen such that $\alpha_{\Sigma_1}(E_\rho) \models\!\mid E_{\alpha_{\Sigma_1'}(\rho)}$, which is possible because α_{Σ_1} is surjective modulo semantical equivalence $\models\!\mid$.

Let us first notice that because α preserves $\Leftrightarrow$ and because it is natural, $(\alpha_{\Sigma_1'}(\rho) \Leftrightarrow \Phi(\varphi_1)(E_{\alpha_{\Sigma_1'}(\rho)})) \models\!\mid \alpha_{\Sigma_1'}(\rho \Leftrightarrow \varphi_1(E_\rho))$. 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 Σ_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 α, this is equivalent to $\alpha_{\Sigma_1'}(\theta'(E')) \models \alpha_{\Sigma_1'}(e)$. From the conservativity of β we get that $\theta'(E') \models e$. Again by the Satisfaction Condition and the definiton of quantifier satisfaction we get that $E' \models (\forall\theta')e$. $\dashv$

Corollary 6.10. Under the assumptions of Proposition 6.9, any (Φ, α, β)-precise signature morphism φ has the definability property if $\Phi(\varphi)$ has the definability property.

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

Consequently, φ 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 $\mathbf{FOL}^T$ too (by Fact 6.11).

We apply Corollary 6.10 to the encoding comorphism $\mathbf{PA} \to \mathbf{FOL}^T$ of Example 6.3, which is precise (Fact 6.5) and conservative (Fact 6.8). It is also easy to see that Φ preserves pushouts. α is surjective modulo $\models\mid$ because it preserves the quantifications and the logical connectives, and because it is surjective on the atoms ($\alpha(t \stackrel{e}{=} t') \models\mid (t = t')$ for each equational $(S, TF, \overline{PF})$-atom and $\alpha(\sigma(t_1, \dots, t_n) \stackrel{e}{=} t) \models\mid \sigma(t_1, \dots, t_n, t)$ 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 $\mathbf{HCL}^T$ too (by Fact 6.11).

Let us consider the restriction of the encoding comorphism $\mathbf{PA} \to \mathbf{FOL}^T$ 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 ρ, $\alpha(\rho)$ 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 β's being isomorphisms. Now let us consider a closed injective homomorphism $h\colon M \to N$ such that $N \models \alpha(\rho)$. We have that $\beta(h)$ is full injective homomorphism and that $\beta(N) \models \rho$. Because QE_1^{ω} sentences are preserved by full injective homomorphisms, $\beta(M) \models \rho$, hence $M \models \alpha(\rho)$.

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

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 $\stackrel{H_r}{\to}$ 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 Németi, *Łoś 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 Mathematics 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 CERIOLI 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 CÎRSTEA, *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 ROŞU, *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 algebroidal 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-Pressice, 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 Németi 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 Mathematics Societas János Bolyai, 29***, pp. 535–578.

[42] Andrei Popescu, Traian Şerbănuţă, and Grigore Roşu, *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.

ŞCOALA NORMALĂ SUPERIOARĂ
CALEA GIVIŢEI 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

Date: Thu, 8 Mar 2007 19:35:21 -0800 (PST)
From: Andrei Popescu <uuomul@yahoo.com>
Subject: Fwd: FW: LOGCOM PROOF exl006 - URGENT
To: Logcom@oxfordjournals.org
Cc: daniel gaina <gaina_daniel@yahoo.com>
MIME-Version: 1.0
Content-Transfer-Encoding: 8bit
Message-ID: <634253.83574.qm@web56110.mail.re3.yahoo.com>

[1 <multipart/alternative>]

[1.1 <text/plain; iso-8859-1 (8bit)>]
Dear Publisher,

I have noticed an omission of a very important affiliation information from the publisher-processed version of our paper, which has unfortunately escaped our vigilance, especially due to the unlikelihood of this type of important omissions: our *primary* affiliation, which read "Scoala Normala Superioara, Bucuresti" (that is a Romanian institution name, translated as "Super Normal School, Bucharest") was completely deleted and only our *secondary* affiliation, "Dept. of Computer Science, Univ. Bucharest etc." was kept.

Since I see that the paper was not published yet as part of a volume accessible to subscribers, I kindly (but strongly) ask you to change the affiliation of the authors into the following:

Primary affiliation: Scoala Normala Superioara, Bucuresti
(correct diacritics spelling, in Latex: \c{S}coala Normal\u{a} Superioar\u{a}, Bucure\c{s}ti)
(or the English version of the name (though the Romanian version is preferred): Super Normal School, Bucharest)
Secondary affiliation: Department of Fundamentals of Computer Science, Fac. of Mathematics, Univ. of Bucharest.

(please see the final author-processed version of our accepted paper, forwarded in the other message that I send you, where the affiliation was listed correctly)

This matter is of a particular importance for correctly identifying the intellectual environment where the paper was conceived (under the *direct* influence of some classes taught in that institution), and also for promoting this young institution which has to face some serious obstacles (bureaucratic and of other kinds) in its efforts to create an elite computer science school in our home country, Romania.

Respectfully,
Andrei Popescu
(CCed the coauthor, Daniel Gaina)

Daniel Găină
Andrei Popescu

An Institution-Independent Proof of the Robinson Consistency Theorem

Abstract. We prove an institutional version of A. Robinson's *Consistency Theorem*. This result is then applied to the institution of many-sorted first-order predicate logic and to two of its variations, infinitary and partial, obtaining very general syntactic criteria sufficient for a signature square in order to satisfy the Robinson consistency and Craig interpolation properties.

Keywords: institution, Robinson consistency, Craig interpolation, elementary diagram, many-sorted first-order logic.

1. Introduction

The many-sorted, rather than unsorted, versions of logical systems (such as equational logic, first-order logic, etc.) are acknowledged as being particularly suitable for applications to computer science, in areas like semantics of programming languages and formal specifications. However, in pure mathematical logic, many-sorted logics tend to be classified as "inessential variations" [33] of their unsorted forms. While this might be true w.r.t. some classical logical aspects such as *compactness*, *completeness*, *Löwenheim properties*, or *axiomatizability*, there is at least one important class of properties that become significantly more intricate when passing from the unsorted to the many-sorted case: those involving the concept of *translation* between languages (signatures), also known as *signature morphism*. Although classical logic, dealing usually just with the very simple case of unsorted language *inclusions*, very rarely cared about these problems, nevertheless any kind of study aiming at providing logical support for diverse areas of theoretical computer science has to consider them, due to the crucial importance of translation between languages in the latter field.

In order to point out the difference between unsorted and many-sorted w.r.t. signature morphisms, we consider two examples in first-order logic. As noticed in [20], the functor *Mod*, taking signatures into their corresponding classes of models and signature morphisms into corresponding "forgetful"

Presented by **Robert Goldblatt**; *Received* September 21, 2004

Studia Logica (2007) 85: 41–73
DOI: 10.1007/s11225-007-9022-4

functors, preserves arbitrary colimits in the many-sorted case, but only some colimits, such as pushouts, in the unsorted case. Another example regards the Craig interpolation property [13], abbreviated CIP, which is classically stated as follows: if $e_1 \vdash e_2$ for two first-order sentences e_1 and e_2, then there exists a sentence e, called the *interpolant* of e_1 and e_2, that uses only logical symbols which appear both in e_1 and e_2 and such that $e_1 \vdash e \vdash e_2$. An equivalent expression of the above property assumes $e_1 \vdash e_2$ in the *union* language $L_1 \cup L_2$ and asks from e to be in the *intersection language* $L_1 \cap L_2$, where L_i is the language of e_i. If, following an approach originating in [49], we naturally generalize the inclusion square

$$\begin{tikzcd}
 & L_1 \arrow[dr, hook] & \\
L_1 \cap L_2 \arrow[ur, hook] \arrow[dr, hook] & & L_1 \cup L_2 \\
 & L_2 \arrow[ur, hook] &
\end{tikzcd}$$

to a pushout of language translations (signature morphisms)

$$\begin{tikzcd}
 & L_1 \arrow[dr, "\varphi'_1"] & \\
L \arrow[ur, "\varphi_1"] \arrow[dr, "\varphi_2"'] & & L' \\
 & L_2 \arrow[ur, "\varphi'_2"'] &
\end{tikzcd}$$

and replace sentences e_1, e_2, e with *sets of sentences* E_1, E_2, E we obtain the following form of CIP: If $\varphi'_1(E_1) \vdash \varphi'_2(E_2)$, then there exists a set E of Σ-sentences such that $E_1 \vdash \varphi_1(E)$ and $\varphi_2(E) \vdash E_2$. Now, the question of which pushout squares have CIP has a definite answer in the unsorted case: *all* of them; this is probably folklore, but also follows from a many-sorted result in [6]. On the other hand, the problem of characterizing the pushout squares which have CIP is still open for the many-sorted case.[1]

An equivalent formulation of CIP in classical logic, with a more model-theoretical flavor, is the Robinson consistency property [42], abbreviated RCP, which states that, if two theories are joint-consistent in their common

[1] Although in the context of the so-called *abstract model-theoretic logics* [3], many-sortedness has a significantly more important status than in classical logic, the issue of interpolation is still treated there only w.r.t. language inclusions.

language, then they are so in their union language. More precisely, for any theories (i.e., sets of sentences closed under deduction) T_1 and T_2 over languages L_1 and L_2 respectively, if $\{e \in T_1 \cup T_2 \mid e \text{ is a sentence in } L_1 \cap L_2\}$ has a model in $L_1 \cap L_2$, then $T_1 \cup T_2$ also has a model in $L_1 \cup L_2$. This paper builds on the generalization of RCP to the abstract level of institutions.

Some Motivation

Finding criteria as general as possible for such a significant property as RCP to hold in a logic is an interesting problem in itself, from the abstract model theory point of view. However, there are reasons why such a study might be useful in theoretical computer science too, reasons given by the tight relationship between RCP and CIP, which goes beyond classical first-order logic; indeed, inside any compact logic with enough expressive power, RCP and CIP are equivalent [49]. In fact, all our RCP results, since they will be based on conditions that make RCP and CIP equivalent, are also results regarding CIP.

CIP is a very useful and broadly studied property in mathematical logic and theoretical computer science - see especially [5, 20, 2], but also [7, 18, 21, 6] for some discussion on the usefulness of this property. Applications of CIP mostly deal with *combining and decomposing theories* and involve areas like algebraic specifications [4, 20, 48, 21], theorem proving and symbolic model checking [38, 39, 30, 31, 52], or algebraic logic [2, 45, 27].[2]

In what follows, we shall offer some motivation for the study of CIP along the lines of our generalization, in the context of structured specifications. A good methodology in specifying hardware or software systems is the *modular* approach, which prescribes building large specifications out of small and easily analyzable pieces. As argued in [20], this allows the verification of many properties *at a very early stage*, at the level of specification rather than that of implementation, thus improving reliability of the systems. The mentioned approach combines specifications stated in different languages (signatures) into larger specifications, using the notion of *language translation*, i.e., *signature morphism.* In many settings for algebraic specification [4, 20, 48, 54], two main operations on modules are considered: that of *reusing text* in a meaningful and model-consistent way, which might involve some *renaming*, and that of *hiding information.* Both these operations, fundamentally differ-

[2]In algebraic logic, CIP is studied in connection to its algebraic counterpart, the amalgamation property; note that the latter property, stated on models and embeddings in the quasi-variety attached to the considered propositional logical system, has nothing to do with the (weak) amalgamation property that we consider later on signature morphisms.

ent in nature, are carried *along* signature morphisms; hence the distinction between two classes of signature morphisms:

- the class of *hiding morphisms*, used for hiding some of the symbols, let it be $\mathcal{H}$, and
- the class of *translating morphisms* used for renaming and/or adding some symbols, let it be $\mathcal{T}$.

A very desirable property is the existence of a (sound and) *complete proof system* for reasoning about structured specifications.[3] It was proved in [10] (for the case of first-order logic) and in [7] for the general case of *institutions* [9, 24] that, in order for such a complete proof system to exist, one needs some good properties of $\mathcal{H}$ and $\mathcal{T}$ w.r.t. each other, among which the most crucial one is $(\mathcal{H}, \mathcal{D})$-*interpolation*, stating that any pushout of signature morphisms
$(\Sigma_2 \stackrel{\varphi_2}{\leftarrow} \Sigma \stackrel{\varphi_1}{\rightarrow} \Sigma_1, \Sigma_2 \stackrel{\varphi_2'}{\rightarrow} \Sigma' \stackrel{\varphi_1'}{\leftarrow} \Sigma_1)$ with $\varphi_1, \varphi_2' \in \mathcal{H}$ and $\varphi_2, \varphi_1' \in \mathcal{T}$ has CIP. It is not clear which types of morphisms are appropriate for hiding and which for translating. But of course, for expressivity reasons, one would like to allow these types to be as general as possible, while keeping the $(\mathcal{H}, \mathcal{D})$-interpolation property. Hence the problem of finding general conditions under which a pushout of signature morphisms has CIP seems to be an important one. Our paper provides such general conditions in the abstract framework of institutions, obtaining in particular the strongest syntactic condition that we are aware of from the literature for a pushout square to have CIP in many-sorted first-order logic (*FOPL*) and in its partial-operation and infinitary-conjunction variations, *PFOPL* and *IFOPL*. Applied to algebraic specification theory, our results give more flexibility to a specification language based on first-order logic such as CASL [12]: one is allowed, for instance, to use signature morphisms that are injective on sorts for hiding purposes and *arbitrary morphisms* for translation purposes, and still have a complete proof system.

The Structure of the Paper

After a preliminary section, recalling some categorical and institutional definitions and notations, in Section 3 we state CIP and different versions of RCP in institutions and show the connections between them. In Section 4, we prove an institutional form of Robinson Consistency Theorem. The framework is that of an institution with elementary diagrams which has

[3] That is, provided that one already has such a proof system for flat specifications.

sufficient expressive power: admits negations and certain quantifications; this loses sight of the equational logics, but concentrates on more expressive first-order-like logics. Section 5 is dedicated to the application of our previous results to many-sorted first-order logic and two variations, infinitary and partial. We obtain a sufficient syntactic criterion for a signature square to be a Craig interpolation square and a Robinson square - this criterion does not assume injectivity on sorts, and covers the case when *one* of the morphisms is injective on sorts. Some concluding remarks and discussion of related work end the paper.

2. Preliminaries

Categories

We assume that the reader is familiar with basic categorical notions like functor, natural transformation, colimit, comma category, etc. A standard textbook on the topic is [26]. We are going to use the terminology from there, with a few exceptions that we point out below. We use both the terms "morphism" and "arrow" to refer morphisms of a category. Composition of morphisms and functors is denoted using the symbol ";" and is considered in diagrammatic order.

Let $\mathcal{C}$ and $\mathcal{C}'$ be two categories. Given an object $A \in |\mathcal{C}|$, the *comma category of objects in* $\mathcal{C}$ *under* A is denoted $A/\mathcal{C}$. Recall that the objects of this category are pairs (h, B), where $B \in |\mathcal{C}|$ and $A \xrightarrow{h} B$ is a morphism in $\mathcal{C}$. Throughout the paper, we might let either $(A \xrightarrow{h} B, B)$, or (h, B), or even h, indicate objects in $A/\mathcal{C}$. A morphism in $A/\mathcal{C}$ between two objects (h, B) and (g, D) is just a morphism $B \xrightarrow{f} D$ in $\mathcal{C}$ such that $h; f = g$ in $\mathcal{C}$. Thus a morphism $A \xrightarrow{h} B$ can be seen in $A/\mathcal{C}$ both as an object and as a morphism between $(1_A, A)$ and (h, B) - this "duplicity" will often appear throughout the paper, so the reader should consider herself warned! There exists a canonical forgetful functor between $A/\mathcal{C}$ and $\mathcal{C}$, mapping each (h, B) to B and each $f : (h, B) \to (g, D)$ to $f : B \to D$. Also, if $F : \mathcal{C}' \to \mathcal{C}$ is a functor, $A \in |\mathcal{C}|$, $A' \in |\mathcal{C}'|$, and $A \xrightarrow{u} F(A')$ is in $\mathcal{C}$, then there exists a canonical functor $u/F : A'/\mathcal{C}' \to A/\mathcal{C}$ mapping each $(A' \xrightarrow{h} B, B)$ to $(u; F(h), F(B))$ and each $f : (h, B) \to (g, D)$ to $F(f) : (u; F(h), F(B)) \to (u; F(g), F(D))$. If $\mathcal{C} = \mathcal{C}'$ and F is the identity functor $1_{\mathcal{C}}$, we write $u/\mathcal{C}$ instead of u/F.

Let $\mathcal{C}$ and $\mathcal{S}$ be two categories such that $\mathcal{S}$ is small. If $D : \mathcal{S} \to \mathcal{C}$ is a functor (also called a *diagram*), then a *cocone* of D is a natural transformation $\mu : D \Longrightarrow V$ between the functor D and [the constant functor mapping all objects to V and all morphisms to 1_V]; V is an object in $\mathcal{C}$, the *vertex* of the colimit, and the components of μ are the *structural morphisms* of the colimit. A diagram defined on the ordered set of natural numbers (regarded as a category) shall be called *ω-diagram*, and a colimit of such a diagram *ω-colimit*. We sometimes identify a diagram $D : J \to \mathcal{C}$ with its image in $\mathcal{C}$, $D(J)$.

Institutions

Institutions were introduced in [9] with the original goal of providing an abstract, logic-independent framework for algebraic specifications of computer science systems. However, by isolating the essence of a logical system in the abstract *satisfaction relation*, these structures also turned out to be appropriate for the development of abstract model theory, as shown by a whole series of (old and new) papers: [49, 50, 51, 46, 47, 15, 16, 18, 17, 23, 40]. See also [34] for an up-to-date discussion on institutions as abstract logics.

An institution [9, 24] consists of:

1. a category $Sign$, whose objects are called *signatures*.
2. a functor $Sen : Sign \to Set$, providing for each signature a set whose elements are called *(Σ-)sentences*.
3. a functor $Mod : Sign \to Cat^{op}$, providing for each signature Σ a category whose objects are called *(Σ-)models* and whose arrows are called *(Σ-)morphisms*.
4. a relation $\models_\Sigma \subseteq |Mod(\Sigma)| \times Sen(\Sigma)$ for each $\Sigma \in |Sign|$, called *(Σ-) satisfaction*, such that for each morphism $\varphi : \Sigma \to \Sigma'$ in $Sign$, the *satisfaction condition*

$$M' \models_{\Sigma'} Sen(\varphi)(e) \text{ iff } Mod(\varphi)(M') \models_\Sigma e$$

holds for all $M' \in |Mod(\Sigma')|$ and $e \in Sen(\Sigma)$. Following the usual notational conventions, we sometimes let $\_\upharpoonright_\varphi$ denote the reduct functor $Mod(\varphi)$ and let φ denote the sentence translation $Sen(\varphi)$. When $M = M'\upharpoonright_\varphi$ we say that M' is a *φ-expansion* of M, and that M is the *φ-reduct* of M'; and similarly for model morphisms.

For all the following concepts related to institutions that we recall below, the reader is referred to [24] unless some other place is explicitly indicated.

Let Σ be a signature. Then,

- for each $E \subseteq Sen(\Sigma)$, let $E^* = \{M \in |Mod(\Sigma)| \mid M \models_\Sigma e \text{ for all } e \in E\}$.

- for each class $\mathcal{M}$ of Σ-models, let $\mathcal{M}^* = \{e \in Sen(\Sigma) \mid M \models_\Sigma e$ for all $M \in \mathcal{M}\}$.

With no danger of confusion, we let $\bullet$ denote any of the two compositions $**$ of the two operators $*$. Each of the two bullets is a closure operator. When E and E' are sets of sentences of the same signature Σ, we let $E \models_\Sigma E'$ denote the fact that $E^* \subseteq E'^*$. The relation $\models_\Sigma$ between sets of sentences is called the *(Σ-)semantic consequence relation* (notice that it is written just like the satisfaction relation). If $E' = \{e'\}$, we might write $E \models_\Sigma e'$. In order to simplify notation, we usually write $\models$ instead of $\models_\Sigma$, for both the satisfaction relation and the semantic consequence relation. Two sentences e and e' are called *equivalent*, denoted $e \equiv e'$, if $\{e\}^* = \{e'\}^*$. Dually, two models M and M' are called *elementary equivalent*, denoted $M \equiv M'$, if $\{M\}^* = \{M'\}^*$. The fact that two models M and M' are isomorphic is indicated by $M \simeq M'$.

A signature morphism $\varphi : \Sigma \to \Sigma'$ is called *conservative* if every Σ-model has a φ-expansion. A *presentation* is a pair (Σ, E), where $E \subseteq Sen(\Sigma)$. A *theory* is a presentation (Σ, E) with E closed, i.e., with $E^\bullet = E$. One usually calls "presentation" or "theory" only the set E, and not the whole pair (Σ, E). A *presentation morphism* $\varphi : (\Sigma, E) \to (\Sigma', E')$ is a signature morphism $\varphi : \Sigma \to \Sigma'$ such that $\varphi(E) \subseteq E'^\bullet$. A presentation morphism between theories is called *theory morphism*. For a presentation (Σ, E), we let $Mod(\Sigma, E)$ denote the category of all Σ-models A such that $A \models E$. A presentation is called *consistent* if it has at least one model; otherwise it is called *inconsistent*.

An institution is called *compact* [20] if, for each signature Σ, the closure operator $\bullet$ on $Sen(\Sigma)$ is compact; in other words, if, for each $E \cup \{e\} \subseteq Sen(\Sigma)$ such that $E \models e$, there exists a finite subset F of E such that $F \models e$. An institution is called *semi-exact* [32] if the model functor $Mod : Sign \to Cat^{op}$ preserves pushouts. A property weaker than semi-exactness that we shall consider is the following. An institution is called *weakly model-semi-exact* if for any pushout of signature morphisms $(\Sigma_2 \stackrel{\varphi_2}{\leftarrow} \Sigma \stackrel{\varphi_1}{\rightarrow} \Sigma_1, \Sigma_2 \stackrel{\varphi_2'}{\rightarrow} \Sigma' \stackrel{\varphi_1'}{\leftarrow} \Sigma_1)$, for any $M_1 \in |Mod(\Sigma_1)|$, $M_2 \in |Mod(\Sigma_2)|$ such that $M_1\restriction_{\varphi_1} = M_2\restriction_{\varphi_2}$, there exists a model $M' \in |Mod(\Sigma')|$ such that $M'\restriction_{\varphi_1'} = M_1$ and $M'\restriction_{\varphi_2'} = M_2$.

The following institutional notions dealing with logical connectives and quantifiers were defined in [49]. Let $\Sigma \in |Sign|$, $e, e_1, e_2 \in Sen(\Sigma)$, $E \subseteq Sen(\Sigma)$, $e' \in Sen(\Sigma')$, and $\varphi : \Sigma \to \Sigma'$.

- a Σ-sentence $\neg e$ is a *negation* of e when $M \models \neg e$ iff $M \not\models e$ for each $M \in |Mod(\Sigma)|$;

- a Σ-sentence $e_1 \wedge e_2$ is a *conjunction* of e_1 and e_2 when $M \models e_1 \wedge e_2$ iff $[M \models e_1$ and $M \models e_2]$ for each $M \in |Mod(\Sigma)|$;
- a Σ-sentence $\bigwedge E$ is a *conjunction* of the set of sentences E when $[M \models \bigwedge E$ iff there exists $f \in E$ such that $M \models f]$ for each $M \in |Mod(\Sigma)|$;
- a Σ-sentence $(\forall\varphi)e'$ is a *universal quantification* of e' over φ when $[M \models (\forall\varphi)e'$ iff there exists $M' \in |Mod(\Sigma')|$ such that $M'\upharpoonright_{\varphi}= M$ and $M' \models e']$ for each $M \in |Mod(\Sigma)|$.

The signature morphisms commute with the logical connectives [49], i.e., using the above notations,

- $\varphi(\neg e)$ is a negation of $\varphi(e)$,
- $\varphi(e_1 \wedge e_2)$ is a conjunction of $\varphi(e_1)$ and $\varphi(e_2)$,
- $\varphi(\bigwedge E)$ is a conjunction of the set of sentences $\varphi(E)$.

An institution is said to *admit*:

- *negations*, if every sentence has a negation;
- *(finite) conjunctions*, if every two sentences have a conjunction;
- *arbitrary conjunctions*, if every set of sentences has a conjunction;
- *universal quantifications* over a given signature morphism $\varphi : \Sigma \to \Sigma'$ if every Σ'-sentence has a universal quantification over φ;

A theory (Σ, T) is called *complete* [49] if it is maximally consistent, i.e., T is consistent and any strict superset T' of it is inconsistent. If the institution admits negations, then a theory T is complete iff there exists a Σ-model A such that $\{A\}^* = T$. We next give two easy, but very useful lemmas.

Lemma 1. [14] (The Institution-Independent Theorem of Constants) *Let $\varphi : \Sigma \to \Sigma'$ be a signature morphism, $E \subseteq Sen(\Sigma)$, $e' \in Sen(\Sigma')$ and $(\forall\varphi)e' \in Sen(\Sigma)$ (so we assume the existence of a universal quantification of e' over φ). Then $\varphi(E) \models e'$ if and only if $E \models (\forall\varphi)e'$.*

Lemma 2. *Assume that the institution is weakly model-semi-exact and let $(\Sigma_2 \xleftarrow{\varphi_2} \Sigma \xrightarrow{\varphi_1} \Sigma_1, \Sigma_2 \xrightarrow{\varphi_2'} \Sigma' \xleftarrow{\varphi_1'} \Sigma_1)$ be a pushout of signature morphisms. Then the following hold:*

1. *If a sentence $e_1 \in Sen(\Sigma_1)$ has a universal quantification over φ_1, then $\varphi_1'(e_1)$ has a universal quantification over φ_2'.*

2. *For each sentence e_1 having a universal quantification over φ_1, it holds that $M_2 \models (\forall\varphi_2')\varphi_1'(e_1)$ iff $M_2\upharpoonright_{\varphi_2} \models (\forall\varphi_1)e_1$ iff $M_2 \models \varphi_2((\forall\varphi_1)e_1)$ for all $M_2 \in |Mod(\Sigma_2)|$.*

PROOF. (1): Let $e_1 \in Sen(\Sigma_1)$ having a universal quantification over φ_1, $(\forall\varphi_1)e_1$. We claim that $\varphi_2((\forall\varphi_1)e_1)$ is a universal quantification of $\varphi_1'(e_1)$ over φ_2'. Indeed, let $M_2 \in |Mod(\Sigma_2)|$.

- Assume $M_2 \models \varphi_2((\forall\varphi_1)e_1)$. Let M' be a φ_2'-expansion of M_2. We need to show $M' \models \varphi_1'(e_1)$, that is, $M'\upharpoonright_{\varphi_1'} \models e_1$. But the last is true, because $M'\upharpoonright_{\varphi_1'}$ is a φ_1-expansion of $M_2\upharpoonright_{\varphi_2}$ and $M_2\upharpoonright_{\varphi_2} \models (\forall\varphi_1)e_1$.
- Conversely, assume that each φ_2'-expansion of M_2 satisfies $\varphi_1'(e_1)$. In order to show $M_2 \models \varphi_2((\forall\varphi_1)e_1)$, i.e., $M_2\upharpoonright_{\varphi_2} \models (\forall\varphi_1)e_1$, let M_1 be a φ_1-expansion of $M_2\upharpoonright_{\varphi_2}$. By weak model-semi-exactness, there exists $M' \in |Mod(\Sigma')|$ such that $M'\upharpoonright_{\varphi_1'} = M_1$ and $M'\upharpoonright_{\varphi_2'} = M_2$. Then $M' \models \varphi_1'(e_1)$, that is, $M_1 \models e_1$.

(2): Immediate by the proof of (1). ■

Elementary Diagrams

Diagrams are an important concept and proof tool in classical model theory [11]. They were first generalized to the institutional framework in [50, 51]; there it is defined the concept of *abstract algebraic institution*, which is an institution subject to some additional natural requirements (like finite-exactness, existence of direct products of models etc.) and enriched with a system of diagrams. The reason for introducing diagrams there was making all algebras *accessible*, for specification purposes. Our proof of the Robinson Consistency Theorem will make heavy use of a more recent institutional notion of elementary diagram, defined in [16].

An institution $\mathcal{I} = (Sign, Sen, Mod, \models)$ is said to have *elementary diagrams* [16] if

1. for each signature Σ and Σ-model A there exists a signature morphism $\iota_\Sigma(A) : \Sigma \to \Sigma_A$ (called the *elementary extension* of Σ via A) and a set E_A of Σ_A-sentences (called the *elementary diagram* of A) such that $Mod(\Sigma_A, E_A)$ and $A/Mod(\Sigma)$ are isomorphic by an isomorphism $i_{\Sigma,A}$ making the following diagram commutative:

$$\begin{tikzcd} Mod(\Sigma_A, E_A) \arrow[rr, "i_{\Sigma,A}"] \arrow[dr, "-\upharpoonright_{\iota_\Sigma(A)}"'] & & A/Mod(\Sigma) \arrow[dl, "\text{forgetful}"] \\ & Mod(\Sigma) & \end{tikzcd}$$

2. ι is "functorial", i.e., for each signature morphism $\varphi : \Sigma \to \Sigma'$, each $A \in |Mod(\Sigma)|$, $A' \in |Mod(\Sigma')|$ and $h : A \to A'\upharpoonright_\varphi$ in $Mod(\Sigma)$, there exists a presentation morphism $\iota_\varphi(h) : (\Sigma_A, E_A) \to (\Sigma'_{A'}, E_{A'})$ making the following diagram commutative:

$$\begin{tikzcd} \Sigma \arrow[r, "\iota_\Sigma(A)"] \arrow[d, "\varphi"'] & \Sigma_A \arrow[d, "\iota_\varphi(h)"] \\ \Sigma' \arrow[r, "\iota_{\Sigma'}(A')"'] & \Sigma'_{A'} \end{tikzcd}$$

3. i is natural, i.e., for each signature morphism $\varphi : \Sigma \to \Sigma'$, each $A \in |Mod(\Sigma)|$, $A' \in |Mod(\Sigma')|$ and $h : A \to A'\upharpoonright_\varphi$ in $Mod(\Sigma)$, the following diagram is commutative:

$$\begin{tikzcd} Mod(\Sigma_A, E_A) \arrow[r, "i_{\Sigma,A}"] & A/Mod(\Sigma) \\ Mod(\Sigma'_{A'}, E_{A'}) \arrow[r, "i_{\Sigma',A'}"] \arrow[u, "-\upharpoonright_{\iota_\varphi(h)}"] & A'/Mod(\Sigma') \arrow[u, "h/Mod(\varphi)"'] \end{tikzcd}$$

In classical model theory, Σ_A is the signature Σ enriched with all the elements of A as constants, $\iota_\Sigma(A) : \Sigma \to \Sigma_A$ is the inclusion of signatures, and E_A is a set of parameterized sentences which hold in A, depending on the considered type of arrow in the categories of models (yielding "elementary diagram" for elementary embeddings, "positive diagram" for arbitrary model homomorphisms, or "diagram" for model embeddings - see [11]). All the three ingredients $\Sigma_A, E_A, \iota_\Sigma(A)$ are also present at the abstract algebraic institutions in [50, 51], where it is also required the natural and potentially very useful fact that A_A be accessible. The important additions of the definition in [16] that we use here are the "functoriality" and naturality conditions, which postulate smooth communication between diagrams along signature morphisms, taking real advantage of the categorical structure of institutions.

The above definition of elementary diagrams may seem, at a first sight, to be adding a great deal of complicated extra structure to institutions. However, it has several advantages:

- looks extremely natural and self-explanatory to anyone familiar with diagrams from classical logic;
- it is so general, that almost all meaningful institutions have elementary diagrams;

- it really provides a "method" for proving logical properties, as we exemplify in this paper.

Here are some notational conventions that we hope will make the reader's life easier. Let $\varphi : \Sigma \to \Sigma'$ be a signature morphism, $A' \in |Mod(\Sigma')|$, and $h : A \to B$ in $Mod(\Sigma)$. We write $\iota_\Sigma(h)$ instead of $\iota_{1_\Sigma}(h)$ and $\iota_\varphi(A'\upharpoonright_\varphi)$ instead of $\iota_\varphi(1_{(A'\upharpoonright_\varphi)})$. Let A be a fixed object in $Mod(\Sigma)$ and let $B, C \in |Mod(\Sigma)|$ and $f : A \to B$, $g : A \to C$, $u : B \to C$ morphisms in $Mod(\Sigma)$ such that $f;u = g$. Then (f, B) and (g, C) are objects in $A/Mod(\Sigma)$ and u is also a morphism in $A/Mod(\Sigma)$ between (f, B) and (g, C). We further establish the following notations: $B_f = i^{-1}_{\Sigma,A}(f, B)$ (and, similarly, $C_g = i^{-1}_{\Sigma,A}(g, C)$), $u_{f,g} = i^{-1}_{\Sigma,A}((f, B) \xrightarrow{u} (g, C))$. Thus, for instance, let $f : A \to B$ be a Σ-model morphism. Then $f_{1_A,f}$ is the image through $i^{-1}_{\Sigma,A}$ of the morphism $f : (1_A, A) \to (f, B)$ in $A/Mod(\Sigma)$, and has source $A_{(1_A)}$ and target B_f. We shall usually write A_A instead of $A_{(1_A)}$ and $f_{A,f}$ instead of $f_{1_A,f}$.

In [16], there are given some examples of institutions with elementary diagrams. Most institutions that were defined on "working" logical systems tend to have elementary diagrams. For the purposes of this paper, we only point out three examples, with their elementary variations.

1. *FOPL* - the institution of many-sorted first-order predicate logic (with equality). The signatures are triplets (S, F, P), where S is the set of sorts, $F = \{F_{w,s}\}_{w \in S^*, s \in S}$ is the ($S^*{\times}S$ -indexed) set of operation symbols, and $P = \{P_w\}_{w \in S^*}$ is the (S^*-indexed) set of relation symbols. By a slight notational abuse, we let F and P also denote and $\bigcup_{(w,s) \in S^* \times S} F_{w,s}$ and $\bigcup_{w \in S^*} P_w$ respectively. A signature morphism between (S, F, P) and (S', F', P') is a triplet $\varphi = (\varphi^{sort}, \varphi^{op}, \varphi^{rel})$, where $\varphi^{sort} : S \to S'$, $\varphi^{op} : F \to F'$, $\varphi^{rel} : P \to P'$ such that $\varphi^{op}(F_{w,s}) \subseteq F'_{\varphi^{sort}(w), \varphi^{sort}(s)}$ and $\varphi^{rel}(P_w) \subseteq P'_{\varphi^{sort}(w)}$ for all $(w, s) \in S^* \times S$. When there is no danger of confusion, we may let φ denote each of φ^{sort}, φ^{rel} and φ^{op}. Given a signature $\Sigma = (S, F, P)$, a Σ-model A is a triplet $A = (\{A_s\}_{s \in S}, \{A^w_s(\sigma)\}_{(w,s) \in S^* \times S, \sigma \in F_{w,s}}, \{A^w(R)\}_{w \in S^*, R \in P_w})$ interpreting each sort s as a set A_s, each operation symbol $\sigma \in F_{w,s}$ as a function $A^w_s(\sigma) : A_w \to A_s$ (where A_w stands for $A_{s_1} \times \ldots \times A_{s_n}$ if $w = s_1 \ldots s_n$), and each relation symbol $R \in P_w$ as a relation $A^w(R) \subseteq A_w$. When there is no danger of confusion we may let A_σ and A_R denote $A^w_s(\sigma)$ and $A^w(R)$ respectively. Morphisms between models are the usual Σ-homomorphisms, i.e., S-sorted functions that preserve the structure. The Σ-sentences are obtained from *atoms*, i.e., equality atoms $t_1 = t_2$, where $t_1, t_2 \in (T_F)_s$,[4] or relational atoms $R(t_1, \ldots, t_n)$, where $R \in P_{s_1 \ldots s_n}$ and

[4] T_F is the ground term algebra over F.

$t_i \in (T_F)_{s_i}$ for each $i \in \{1, \ldots, n\}$, by applying for a finite number of times:

- negation, conjunction, disjunction;
- universal or existential quantification over finite sets of constants.

Satisfaction is the usual first-order satisfaction and is defined using the natural interpretations of ground terms t as elements A_t in models A. The definitions of functors *Sen* and *Mod* on morphisms are the natural ones: for any signature morphism $\varphi : \Sigma \to \Sigma'$, $Sen(\varphi) : Sen(\Sigma) \to Sen(\Sigma')$ translates sentences symbol-wise, and $Mod(\varphi) : Mod(\Sigma') \to Mod(\Sigma)$ is the forgetful functor.

As shown in [16], *FOPL* has elementary diagrams in the institutional sense. However, we shall be interested in what is called "elementary diagram" according to the classical model theory terminology [11]; the latter are in fact the diagrams of a remarkable subinstitution of *FOPL*, which has the same signatures, sentences and models, but restricts the class of model morphisms to *elementary embeddings* only. We shall be more precise below.

Let $\Sigma = (S, F, P)$ be a signature in *FOPL* and let $A \xrightarrow{h} B$ be a model morphism in $Mod(\Sigma)$. Let $\Sigma_A = (S, F_A, P)$, where F_A extends F by adding, for each $s \in S$, all elements in A_s as constants of sort s; also, let A_A be the expansion of A to Σ_A which interprets each constant $a \in A_s$ as itself, for all $s \in S$. The signature Σ_B and the Σ_B-model B_B are defined similarly. Define $\iota_\Sigma(h) : Sen(\Sigma_A) \to Sen(\Sigma_B)$ to be the following: if $e \in Sen(\Sigma_A)$, then $\iota_\Sigma(h)(e)$ is obtained from e by symbol-wise translation, mapping:

- for all $s \in S$, each $a \in A_s$ into $h_s(a)$,
- each other symbol u that appears in e into u.

A morphism $A \xrightarrow{h} B$ in $Mod(\Sigma)$ is said to be an *elementary embedding* if, for each $e \in Sen(\Sigma_A)$, $A_A \models e$ iff $B_B \models \iota_\Sigma(h)(e)$. The term "embedding" is appropriate, since all the elementary embeddings are injective morphisms. It is well known, and can be easily seen, that the elementary embeddings form a broad subcategory of $Mod(\Sigma)$ and are preserved by reduct functors. Thus we have an "elementary" subinstitution of *FOPL*, denoted *ElFOPL*, which has all the structure identical to *FOPL*, just that the model morphisms are restricted to be elementary embeddings. We now define some elementary diagrams for *ElFOPL*:

Let $\Sigma = (S, F, P)$ be a *FOPL* signature and $A \in |Mod(\Sigma)|$. Then:

- $\Sigma_A = (S, F_A, P)$ and A_A were already indicated above;
- $E_A = (A_A)^* = \{e \in Sen(\Sigma_A) \mid A_A \models e\}$;

- $\Sigma \stackrel{\iota_\Sigma(A)}{\to} \Sigma_A$ is the signature inclusion;
- The functor $i_{\Sigma,A} : Mod(\Sigma_A, E_A) \to A/Mod(\Sigma)$ is defined: on objects, by $i_{\Sigma,A}(N') = (A \stackrel{h}{\to} N, N)$, where $N = N'\upharpoonright_{\iota_\Sigma(A)}$ and, for each $s \in S$ and $a \in A_s$, $h_s(a) = N'_a$; on morphisms, by $i_{\Sigma,A}(f) = f$.

Let $\varphi : \Sigma = (S, F, P) \to \Sigma' = (S', F', P')$ be a *FOPL* signature morphism, $A \in |Mod(\Sigma)|$, $C \in |Mod(\Sigma')|$, and $h : A \to C\upharpoonright_\varphi$ an elementary morphism in $Mod(\Sigma)$. Then the natural presentation morphism $\iota_\varphi(h) : (\Sigma_A, E_A) \to (\Sigma_C, E_C)$ from the definition of elementary diagrams is the following: if $e \in Sen(\Sigma_A)$, then $\iota_\varphi(h)(e)$ is obtained from e by symbol-wise translation, mapping:

- each $f \in F$ into $\varphi^{op}(f)$,
- each $R \in P$ into $\varphi^{rel}(R)$,
- for all $s \in S$, each $a \in A_s$ into $h_s(a) \in C_{\varphi^{sort}(s)}$,
- for all $s \in S$, each variable $x{:}s$ of sort s into a variable $x{:}\varphi^{sort}(s)$ of sort $\varphi^{sort}(s)$,
- each other symbol u that appears in e (e.g., logical connectives and quantifiers) into u.

It is routine to check that *ElFOPL*, together with the above structure, is an institution with elementary diagrams.

2. *PFOPL* - the institution of partial first-order predicate logic, an extension of *FOPL* whose signatures $\Sigma = (S, F, F', P)$ contain, besides relation and (total) operation symbols (in F and P), *partial* operation symbols too, in F'. Models of course interpret the symbols in F' as partial operations of appropriate ranks. Σ-model morphisms $h : A \to B$ are S-sorted functions which commute with the total operations and relations in the usual way, and with the partial operations $\sigma \in F'_{s_1 \ldots s_n, s}$ in the following way: for each $(a_1, \ldots, a_n) \in A_{s_1 \ldots s_n}$, if $A_\sigma(a_1, \ldots, a_n)$ is defined, then so is $B_\sigma(h_{s_1}(a_1), \ldots, h_{s_n}(a_n))$, and in this case the latter is equal to $h_s(A_\sigma(a_1, \ldots, a_n))$. Signature morphisms are allowed to map partial operation symbols to total operation symbols, but not *vice versa.* There exist three kinds of atoms: *relational* atoms just like at *FOPL*, *undefinedness* atoms $t\uparrow$, and *(strong) equality* atoms $t = t'$. A relational atom $R(t_1, \ldots, t_n)$ holds in a model A when all terms t_i are defined and their interpretations A_{t_i} stay in relation A_R. The undefinedness $t\uparrow$ of a term t holds in a model A when the corresponding interpretation A_t of the term is undefined. The equality $t = t'$

holds when both terms are undefined or both terms are defined and equal. The sentences are obtained from atoms just like in the case of *FOPL*. Partial algebras (i.e., *PFOPL*-models over signatures with no relation symbols) and their applications were extensively studied in [41] and [8].

3. *IFOPL* - the institution of infinitary first-order logic, an infinitary extension of *FOPL*, which allows conjunctions on arbitrary sets of sentences. This logical system is known under the name $L_{\infty,\omega}$ [29, 28][5] and plays an important role in categorical logic.

The corresponding "elementary" subinstitutions of *IFOPL* and *PFOPL*, denoted *ElIFOPL* and *ElPFOPL*, as well as their diagrams, are defined similarly to the case of *FOPL*. For *ElIFOPL*, the definitions are identical, while for *PFOPL* they have to be incremented in the obvious way to consider the partial operation symbols too.

Notice that models in the above institutions are not required to have non-empty carriers on sorts. There are subtle issues in algebraic specifications (like the unconditional existence of free models) that plead for this approach, which departs from the (unsorted) algebraic tradition of assuming non-emptiness of carrier sets. However, it seems to be a habit taking the non-emptiness assumption when considering Craig interpolation, sometimes even within algebraic specification frameworks [44, 43, 6]. In what follows, we shall take the trouble of distinguishing between the two apparently very similar situations, and shall point out some differences w.r.t. RCP and CIP (see Corollaries 9 and 10) that give a technical explanation for the above mentioned habit. Let *ElFOPL'*, *ElIFOPL'*, *ElPFOPL'*, *FOPL'*, *IFOPL'*, *PFOPL'* denote the variations of *ElFOPL*, *ElIFOPL*, *ElPFOPL*, *FOPL*, *IFOPL*, *PFOPL* with the additional requirement that models have non-empty carriers on all sorts. Many relevant properties of the original institutions, like semi-exactness (hence weak model-semi-exactness), compactness etc., hold for their non-empty-carrier versions too. Also, for our future discussions about elementary chains, the non-emptiness assumption is irrelevant (see also the proof of Corollary 9). The only moment when important technical differences will come into the picture is occasioned by quantifications over signature morphisms.

[5] Actually, the mentioned books allow a more general form of signature, with infinitary operation- and relation- symbols too. The results of this paper cover the cases of such signatures too, as an interested reader could easily check.

3. Institutional Formulation of the Robinson Consistency Property

We fix an institution $\mathcal{I}$. Next, we state some logical properties regarding language translation, following a generalization originating in [49], on *arbitrary* squares of signature morphisms rather than inclusion squares.

DEFINITION 3. Let $\mathcal{S}$ be a commutative signature square

$$\begin{tikzcd}
 & \Sigma_1 \arrow[dr, "\varphi_1'"] & \\
\Sigma \arrow[ur, "\varphi_1"] \arrow[dr, "\varphi_2"'] & & \Sigma' \\
 & \Sigma_2 \arrow[ur, "\varphi_2'"'] &
\end{tikzcd}$$

$\mathcal{S}$ is said to be:

1. a *weak amalgamation square (w.a. square)*, if every two models $A_1 \in |Mod(\Sigma_1)|$ and $A_2 \in |Mod(\Sigma_2)|$ having the same reduct (i.e., such that $A_1 \upharpoonright_{\varphi_1} = A_2 \upharpoonright_{\varphi_2}$), have a common expansion (i.e., there exists $A' \in |Mod(\Sigma')|$ such that $A' \upharpoonright_{\varphi_1'} = A_1$ and $A' \upharpoonright_{\varphi_2'} = A_2$);
2. a *Craig interpolation square (CI square)*, if for every $E_1 \subseteq Sen(\Sigma_1)$ and $E_2 \subseteq Sen(\Sigma_2)$ such that $\varphi_1'(E_1) \models \varphi_2'(E_2)$, there exists $E \subseteq Sen(\Sigma)$ such that $E_1 \models \varphi_1(E)$ and $\varphi_2(E) \models E_2$.

Note that if the institution is weakly model-semi-exact, then any pushout of signatures is a weak amalgamation square. The CI property from above was defined [49] on arbitrary *pushout squares*. However we shall prefer to work, in the style of [18], under the slightly more general hypothesis of w.a. square. We next provide three candidates for the notion of Robinson square, two of them already defined in the literature:

DEFINITION 4. A commutative square as in the figure of Definition 3 is said to be:

1. a 1-*Robinson square*, if for every consistent theories $T_2 \subseteq Sen(\Sigma_2)$, $T_1 \subseteq Sen(\Sigma_1)$ and complete theory $T \subseteq Sen(\Sigma)$ such that φ_1, φ_2 are theory morphisms, it holds that $\varphi_1'(T_1) \cup \varphi_2'(T_2)$ is consistent;
2. a 2-*Robinson square*, if for every two models $A_1 \in |Mod(\Sigma_1)|$ and $A_2 \in |Mod(\Sigma_2)|$ such that $A_1 \upharpoonright_{\varphi_1} \equiv A_2 \upharpoonright_{\varphi_2}$, there exists $A' \in |Mod(\Sigma')|$ such that $A' \upharpoonright_{\varphi_1'} \equiv A_1$ and $A' \upharpoonright_{\varphi_2'} \equiv A_2$;

3. a 3-*Robinson square*, if for every two consistent theories $T_1 \subseteq Sen(\Sigma_1)$ and $T_2 \subseteq Sen(\Sigma_2)$ such that $\varphi_1^{-1}(T_1) \cup \varphi_2^{-1}(T_2)$ is consistent, it holds that $\varphi_1'(T_1) \cup \varphi_2'(T_2)$ is consistent;

Remark 5. *1. The converses of the [2 and 3]-Robinson properties in Definition 4 are always true:*

- *if for two models $A_1 \in |Mod(\Sigma_1)|$ and $A_2 \in |Mod(\Sigma_2)|$ there exists $A' \in |Mod(\Sigma')|$ such that $A'\upharpoonright_{\varphi_1'} \equiv A_1$ and $A'\upharpoonright_{\varphi_2'} \equiv A_2$, then $A_1\upharpoonright_{\varphi_1} \equiv A_2\upharpoonright_{\varphi_2}$;*
- *if $\varphi_1'(T_1) \cup \varphi_2'(T_2)$ is consistent, then so is $\varphi_1^{-1}(T_1) \cup \varphi_2^{-1}(T_2)$.*

2. If the institution is compact and admits negations and finite conjunctions, then the definition of 3-Robinson square can be rewritten as follows (with the notations of Definition 4.(3)): if $\varphi_1'(T_1) \cup \varphi_2'(T_2)$ is not consistent then there exists $e \in \varphi_1^{-1}(T_1)$ such that $\neg e \in \varphi_2^{-1}(T_2)$.

The 1-Robinson property was defined in [49] following a variant of the corresponding classical property in unsorted first-order logic. The 3-Robinson property follows the other equally used classical definition [53]. On the other hand, the differently looking 2-Robinson property, introduced in [47] for *preinstitutions* following an idea from [36, 37], was also called, for obvious reasons, the *elementary amalgamation* property [18]. In many institutions, the three Robinson properties, as well as the CI-property, are all equivalent.

Proposition 6. *Assume that $\mathcal{I}$ has negations and finite conjunctions and is compact. Then the following are equivalent for a commutative square $\mathcal{S}$:*

1. *$\mathcal{S}$ is a 1-Robinson square;*
2. *$\mathcal{S}$ is a 2-Robinson square;*
3. *$\mathcal{S}$ is a 3-Robinson square;*
4. *$\mathcal{S}$ is a CI square.*

Proof. Let $\mathcal{S}$ be a commutative square as in the figure of Definition 3.
(1) implies (2): Take $T = \{A_1\upharpoonright_{\varphi_1}\}^* = \{A_2\upharpoonright_{\varphi_2}\}^*$, $T_i = \{A_i\}^*$, $i \in \{1,2\}$. Let $i \in \{1,2\}$. If $e \in T$, then $A_i\upharpoonright_{\varphi_i} \models e$, thus $A_i \models \varphi_i(e)$, thus $\varphi_i(e) \in T_i$. Hence φ_i is a theory morphism. Moreover, T, T_1, T_2 are complete, thus there exists a Σ-model $A' \models \varphi_1'(T_1) \cup \varphi_2'(T_2)$. But $A\upharpoonright_{\varphi_i'} \models T_i$, hence $A\upharpoonright_{\varphi_i'} \equiv A_i$.
(2) implies (1): Since T is complete, there exists a Σ-model A such that $\{A\}^* = T$. Let also $A_i \models T_i$, $i \in \{1,2\}$. Since φ_i is a theory morphism, $A_i\upharpoonright_{\varphi_i} \models T$, thus $A_i\upharpoonright \varphi_i \equiv A$, $i \in \{1,2\}$. Then there exists a Σ-model A' such that $A\upharpoonright \varphi_i' \equiv A_i$, hence $A\upharpoonright_{\varphi_i'} \models T_i$, hence $A \models \varphi_i'(T_i)$, $i \in \{1,2\}$. Thus $\varphi_1'(T_1) \cup \varphi_2'(T_2)$ is consistent.

(1) equivalent to (4): Was proved in [49], Corollary 3.1.
(3) implies (4): First notice that in Definition 4.(3) the property of being a 3-Robinson square can be equivalently expressed not assuming T_1 and T_2 to be theories (but just sets of sentences), and considering $\varphi_1^{-1}(T_1^\bullet) \cup \varphi_2^{-1}(T_2^\bullet)$ instead of $\varphi_1^{-1}(T_1) \cup \varphi_2^{-1}(T_2)$.

Let now $E_1 \subseteq Sen(\Sigma_1)$ and $E_2 \subseteq Sen(\Sigma_2)$ such that $\varphi_1'(E_1) \models \varphi_2'(E_2)$. Fix $e_2 \in E_2$. We have $\varphi_1'(E_1) \models \varphi_2'(e_2)$, so $\varphi_1'(E_1) \cup \{\varphi_2'(\neg e_2)\}$ is inconsistent. Applying the 3-Robinson square property we obtain that $\varphi_1^{-1}(E_1^\bullet) \cup \varphi_2^{-1}(\{\neg e_2\}^\bullet)$ is also inconsistent, which implies, by compactness and finite conjunctions, the existence of a sentence $e \in Sen(\Sigma)$ such that $\varphi_1^{-1}(E_1^\bullet) \models e$ and $\varphi_2^{-1}(\{\neg e_2\}^\bullet) \models \neg e$. But $\varphi_1^{-1}(E_1^\bullet)$ and $\varphi_1^{-1}(\{e_2\}^\bullet)$ are closed, so $e \in \varphi_1^{-1}(E_1^\bullet)$ and $\neg e \in \varphi_2^{-1}(\{\neg e_2\}^\bullet)$, i.e. $E_1 \models \varphi_1(e)$ and $\neg e_2 \models \varphi_2(\neg e)$, the last equality being equivalent to $\neg e_2 \models \neg\varphi_2(e)$, and further to $\varphi_2(e) \models e_2$. Thus, for any $e_2 \in E_2$ we found an $e \in Sen(\Sigma)$ such that $E_1 \models \varphi_1(e)$ and $\varphi_2(e) \models e_2$. Let $E \subseteq Sen(\Sigma)$ be the set of all such e, for each $e_2 \in E_2$. Then $E_1 \models \varphi_1(E)$ and $\varphi_2(E) \models E_2$.
(4) implies (3): Let $T_1 \subseteq Sen(\Sigma_1)$ and $T_2 \subseteq Sen(\Sigma_2)$ be two theories such that $\varphi_1'(T_1) \cup \varphi_2'(T_2)$ is inconsistent. Using finite conjunctions and compactness, we find $\gamma_2 \in T_2$, such that $\varphi_1'(T_1) \cup \{\varphi_2'(\gamma_2)\}$ is inconsistent. Since $\mathcal{I}$ has negations, it follows that $\varphi_1'(T_1) \models \neg\varphi_2'(\gamma_2)$, that is, $\varphi_1'(T_1) \models \varphi_2'(\neg\gamma_2)$. By Craig interpolation, there exists $E \subseteq Sen(\Sigma)$ such that $T_1 \models \varphi_1(E)$ and $\varphi_2(E) \models \neg\gamma_2$. Hence, by compactness and finite conjunctions, $\varphi_2(e) \models \neg\gamma_2$, for some $e \in E^\bullet$; this means $\gamma_2 \models \neg\varphi_2(e)$, i.e., $\gamma_2 \models \varphi_2(\neg e)$. Furthermore, $\varphi_1(E^\bullet) \subseteq \varphi_1(E)^\bullet \subseteq T_1^\bullet = T_1$ so $T_1 \models \varphi_1(e)$. We have obtained $T_1 \models \varphi_1(e)$ and $T_2 \models \varphi_2(\neg e)$. Since T_1 and T_2 are theories, it holds that $e \in \varphi_1^{-1}(T_1)$ and $\neg e \in \varphi_2^{-1}(T_2)$, making $\varphi_1^{-1}(T_1) \cup \varphi_2^{-1}(T_2)$ inconsistent. ∎

Since we shall only deal with institutions satisfying the hypotheses in Proposition 6, we can safely say *Robinson square* instead of *i-Robinson square*. However, we are going to use the property of 3-Robinson square.

We introduce a final technical concept. The following notion of *lifting isomorphisms* generalizes a similar one in [18], from signature morphisms, to signature squares. The intuition is that φ_1 and φ_2 *together* lift isomorphisms. Notice that the below definition does not use the morphisms φ_1' and φ_2'; we keep the "square" terminology just for uniformity.

Definition 7. A commutative square as in the figure of Definition 3 is said to *lift isomorphisms* if, for each $A_1 \in |Mod(\Sigma_1)|$ and $A_2 \in |Mod(\Sigma_2)|$ such that $A_1\upharpoonright_{\varphi_1}$ is isomorphic to $A_2\upharpoonright_{\varphi_2}$, there exist $B_1 \in |Mod(\Sigma_1)|$ and $B_2 \in |Mod(\Sigma_2)|$ such that:

- B_1 is isomorphic to A_1;
- B_2 is isomorphic to A_2;
- $B_1\upharpoonright_{\varphi_1}= B_2\upharpoonright_{\varphi_2}$.

4. The Robinson Consistency Theorem

Theorem 8. (The Consistency Theorem) *We assume that the institution $\mathcal{I}$:*

- *has all the model morphisms preserving satisfaction, i.e., for each signature Σ'' and $A \to B$ in $Mod(\Sigma'')$, it holds that $\{A\}^* \subseteq \{B\}^*$,*
- *has elementary diagrams,*
- *has pushouts of signatures and is weakly model-semi-exact,*
- *has ω-colimits of models preserved by the reduct functors,*
- *admits (finite) conjunctions and negations,*
- *is compact.*[6]

Then any w.a. square (and in particular any pushout square) as in the figure of Definition 3, which lifts isomorphisms and, in addition, has the property:

- *the institution admits universal quantifications over morphisms of the forms $\iota_\Sigma(h)$ and $\iota_\Sigma(A)$ for each Σ-model morphism $A \xrightarrow{h} B$* [7] *(with the notations of elementary diagrams introduced in Section 2),*

is a Robinson square (hence a CI square).

Proof. Let $\mathcal{S}$ be a w.a. square as in the figure of Definition 3 and $T_1 \subseteq Sen(\Sigma_1)$, $T_2 \subseteq Sen(\Sigma_2)$ be two theories. Denote $\Gamma_1 = \varphi_1^{-1}(T_1)$ and $\Gamma_2 = \varphi_2^{-1}(T_2)$. Γ_1 and Γ_2 are also theories. We assume that $\Gamma_1 \cup \Gamma_2$ is consistent and want to prove that $\varphi_1'(T_1) \cup \varphi_2'(T_2)$ is consistent. It suffices to find two models $M_1 \models T_1$ and $M_2 \models T_2$ such that $M_1\upharpoonright_{\varphi_1}= M_2\upharpoonright_{\varphi_2}$ (and then apply weak amalgamation to find the desired model M' of $\varphi_2'(T_1) \cup \varphi_2'(T_2)$). We first construct inductively two chains of models, as indicated below.

(1) We find a model $A_1 \models T_1$ such that $A_1\upharpoonright_{\varphi_1}\models \Gamma_2$. If such a model didn't exist, then $T_1 \cup \varphi_1(\Gamma_2)$ would be inconsistent, so, by compactness and the existence of finite conjunctions, $T_1 \cup \{\varphi_1(\gamma_2)\}$ would be inconsistent, for some $\gamma_2 \in \Gamma_2$. By the existence of negations, this would imply $T_1 \models \neg\varphi_1(\gamma_2)$, that is, $T_1 \models \varphi_1(\neg\gamma_2)$, so $\neg\gamma_2 \in \Gamma_1$, making $\Gamma_1 \cup \Gamma_2$ inconsistent, a contradiction.

[6] In particular, any institution which admits arbitrary conjunctions, such as *IFOPL*, is compact.

[7] Notice that this last condition is a local one, involving the fixed signature Σ of the considered square.

(2) We find $A_2 \models T_2$ and $A_1\upharpoonright_{\varphi_1} \xrightarrow{h} A_2\upharpoonright_{\varphi_2}$ in $Mod(\Sigma)$. Using the elementary diagrams, it suffices to find $B \models E_{(A_1\upharpoonright_{\varphi_1})}$ and $A_2 \models T_2$ such that $B\upharpoonright_{\iota_\Sigma(A_1\upharpoonright_{\varphi_1})} = A_2\upharpoonright_{\varphi_2}$. Moreover, it suffices to consider the pushout of signatures

$$\begin{tikzcd}
 & \Sigma_{(A_1\upharpoonright_{\varphi_1})} \arrow[dr, "u"] & \\
\Sigma \arrow[ur, "\iota_\Sigma(A_1\upharpoonright_{\varphi_1})"] \arrow[dr, "\varphi_2"'] & & \Sigma' \\
 & \Sigma_2 \arrow[ur, "v"'] &
\end{tikzcd}$$

and find, in $Mod(\Sigma')$, a model of $u(E_{(A_1\upharpoonright_{\varphi_1})}) \cup v(T_2)$.[8] If such a model didn't exist, making $u(E_{(A_1\upharpoonright_{\varphi_1})}) \cup v(T_2)$ inconsistent, then, using negations, finite conjunctions and compactness, we would find $e \in (E_{(A_1\upharpoonright_{\varphi_1})})^\bullet$ such that $v(T_2) \models u(\neg e)$. By Lemma 1, we would have $T_2 \models (\forall v)u(\neg e)$ (notice that the sentence $(\forall v)u(\neg e)$ exists in our institution according to Lemma 2.(1)). Furthermore, by Lemma 2.(2), $\varphi_2((\forall \iota_\Sigma(A_1\upharpoonright_{\varphi_1}))\neg e) \equiv (\forall v)u(\neg e)$, thus $\varphi_2((\forall \iota_\Sigma(A_1\upharpoonright_{\varphi_1}))\neg e) \in T_2^\bullet = T_2$, which means $(\forall \iota_\Sigma(A_1\upharpoonright_{\varphi_1}))\neg e \in \Gamma_2$. But $A_1\upharpoonright_{\varphi_1} \models \Gamma_2$, so $A_1\upharpoonright_{\varphi_1} \models (\forall \iota_\Sigma(A_1\upharpoonright_{\varphi_1}))\neg e$, contradicting the fact that $(A_1\upharpoonright_{\varphi_1})_{(A_1\upharpoonright_{\varphi_1})} \models e$.

(3) We find $B_1 \in |Mod(\Sigma_1)|$, $A_1 \xrightarrow{g} B_1$ in $Mod(\Sigma_1)$, and $A_2\upharpoonright_{\varphi_2} \xrightarrow{f} B_1\upharpoonright_{\varphi_1}$ in $Mod(\Sigma)$ such that $h; f = g\upharpoonright_{\varphi_1}$. It suffices to find $D_1 \in |Mod(\Sigma_{1_{A_1}}, E_{A_1})|$ and $D_2 \in |Mod(\Sigma_{(A_2\upharpoonright_{\varphi_2})}, E_{(A_2\upharpoonright_{\varphi_2})})|$ such that $D_1\upharpoonright_{\iota_{\varphi_1}(A_1\upharpoonright_{\varphi_1})} = D_2\upharpoonright_{\iota_\Sigma(h)}$. Indeed, let us first assume that we found such models D_1 and D_2. Then g would be $i^{-1}_{\Sigma_1, A_1}(D_1)$ and f would be $i^{-1}_{\Sigma,(A_2\upharpoonright_{\varphi_2})}(D_2)$. In order to prove that $h; f = g\upharpoonright_{\varphi_1}$, we apply the "functoriality" of ι and obtain that the below diagram is commutative:

$$\begin{tikzcd}
 & \Sigma_{1_{A_1}} & & \\
\Sigma_1 \arrow[ur, "\iota_{\Sigma_1}(A_1)"] & & \Sigma_{(A_1\upharpoonright_{\varphi_1})} \arrow[ul, "\iota_{\varphi_1}(A_1\upharpoonright_{\varphi_1})"'] \arrow[dr, "\iota_\Sigma(h)"] & \\
 & \Sigma \arrow[ul, "\varphi_1"] \arrow[ur, "\iota_\Sigma(A_1\upharpoonright_{\varphi_1})"'] \arrow[rr, "\iota_\Sigma(A_2\upharpoonright_{\varphi_2})"'] & & \Sigma_{(A_2\upharpoonright_{\varphi_2})}
\end{tikzcd}$$

[8] We actually applied here the converse of Robinson Consistency Property for the theories $E_{(A_1\upharpoonright_{\varphi_1})}$ and T_2; see Remark 5(1).

We now apply the naturality of i to get that the below diagram is commutative:

$$\begin{tikzcd}
 & Mod(\Sigma 1_{A_1}, E_{A_1}) \arrow[dl, "i_{\Sigma_1, A_1}"'] \arrow[dr, "-\upharpoonright_{\iota_{\varphi_1}(A_1\upharpoonright_{\varphi_1})}"] & & & Mod(\Sigma_{(A_2\upharpoonright_{\varphi_2})}, E_{(A_2\upharpoonright_{\varphi_2})}) \arrow[dll, "-\upharpoonright_{\iota_\Sigma(h)}"'] \arrow[ddl, "i_{\Sigma,(A_2\upharpoonright_{\varphi_2})}"] \\
A_1/Mod(\Sigma_1) \arrow[dr, "-\upharpoonright_{\varphi_1}"'] & & Mod(\Sigma_{(A_1\upharpoonright_{\varphi_1})}, E_{(A_1\upharpoonright_{\varphi_1})}) \arrow[dl, "i_{\Sigma,(A_1\upharpoonright_{\varphi_1})}"] & & \\
 & A_1\upharpoonright_{\varphi_1} /Mod(\Sigma) & & A_2\upharpoonright_{\varphi_2} /Mod(\Sigma) \arrow[ll, "h/Mod(\Sigma)"] &
\end{tikzcd}$$

Then, since $D_1\upharpoonright_{\iota_{\varphi_1}(A_1\upharpoonright_{\varphi_1})}= D_2\upharpoonright_{\iota_\Sigma(h)}$, we have $i_{\Sigma,(A_1\upharpoonright_{\varphi_1})}(D_1\upharpoonright_{\iota_{\varphi_1}(A_1\upharpoonright_{\varphi_1})})$ $= i_{\Sigma,(A_1\upharpoonright_{\varphi_1})}(D_2\upharpoonright_{\iota_\Sigma(h)})$, so $(i_{\Sigma_1,A_1}(D_1))\upharpoonright_{\varphi_1}= h; i_{\Sigma,(A_2\upharpoonright_{\varphi_2})}(D_2)$, that is, $g\upharpoonright_{\varphi_1}= h; f$.

Now let us come back: we need to prove the existence of two models $D_1 \in |Mod(\Sigma 1_{A_1}, E_{A_1})|$ and $D_2 \in |Mod(\Sigma_{(A_2\upharpoonright_{\varphi_2})}, E_{(A_2\upharpoonright_{\varphi_2})})|$ with a common reduct to $\Sigma_{A_1\upharpoonright_{\varphi_1}}$, or, sufficiently, with a common expansion to Σ_0, where

$$\begin{tikzcd}
 & \Sigma 1_{A_1} \arrow[dr, "u"] & \\
\Sigma_{A_1\upharpoonright_{\varphi_1}} \arrow[ur, "\iota_{\varphi_1}(A_1\upharpoonright_{\varphi_1})"] \arrow[dr, "\iota_\Sigma(h)"'] & & \Sigma_0 \\
 & \Sigma_{(A_2\upharpoonright_{\varphi_2})} \arrow[ur, "v"'] &
\end{tikzcd}$$

is a pushout of signatures. Let us assume that there are no such models, i.e., that $u(E_{A_1}) \cup v(E_{(A_2\upharpoonright_{\varphi_2})})$ is not consistent. We again invoke negations, conjunctions and compactness to find $e \in (E_{(A_2\upharpoonright_{\varphi_2})})^\bullet$ such that $u(E_{A_1}) \models v(\neg e)$. Similarly to step 2, we apply Lemma 1 to get $E_{A_1} \models (\forall u)v(\neg e)$. This implies $A 1_{A_1} \models (\forall u)v(\neg e)$. By Lemma 2.(1), $(\forall u)v(\neg e) \equiv \iota_{\varphi_1}(A_1)((\forall \iota_\Sigma(h))\neg e)$, hence $A 1_{A_1} \models \iota_{\varphi_1}(A_1)((\forall \iota_\Sigma(h))\neg e)$, hence $A 1_{A_1}\upharpoonright_{\iota_{\varphi_1}(A_1)} \models (\forall \iota_\Sigma(h))\neg e$. Because of the naturality of ι, we have that $A 1_{A_1}\upharpoonright_{\iota_{\varphi_1}(A_1)}= (A_1\upharpoonright_{\varphi_1})_{(A_1\upharpoonright_{\varphi_1})}$. We obtain $(A_1\upharpoonright_{\varphi_1})_{(A_1\upharpoonright_{\varphi_1})} \models (\forall \iota_\Sigma(h))\neg e$. Since, like any model morphism, $(A_1\upharpoonright_{\varphi_1})_{(A_1\upharpoonright_{\varphi_1})} \stackrel{h_{(A_1\upharpoonright_{\varphi_1}),h}}{\rightarrow} (A_2\upharpoonright_{\varphi_2})_h$ preserves satisfaction, we have $(A_2\upharpoonright_{\varphi_2})_h \models$

$(\forall \iota_\Sigma(h))\neg e$, contradicting the fact that $(A_2\restriction_{\varphi_2})_{(A_2\restriction_{\varphi_2})}$, a $\iota_\Sigma(h)$-expansion of $(A_2\restriction_{\varphi_2})_h$, satisfies e (remember that $e \in (E_{(A_2\restriction_{\varphi_2})})^\bullet$).

(4) We reuse the technique of step 3 in order to find $B_2 \in |Mod(\Sigma_2)|$, $A_2 \xrightarrow{s} B_2$ in $Mod(\Sigma_2)$, and $B_1\restriction_{\varphi_1} \xrightarrow{p} B_2\restriction_{\varphi_2}$ in $Mod(\Sigma)$ such that $f;p = s\restriction_{\varphi_2}$. Applying this a countable number of times we obtain two ω-diagrams Ch_1 and Ch_2:

$$A_1^0 \xrightarrow{f_1^0} A_1^1 \xrightarrow{f_1^1} A_1^2 \xrightarrow{f_1^2} A_1^3 \ldots \quad \text{in} \quad Mod(\Sigma_1)$$

$$A_2^0 \xrightarrow{f_2^0} A_2^1 \xrightarrow{f_2^1} A_2^2 \xrightarrow{f_2^2} A_2^3 \ldots \quad \text{in} \quad Mod(\Sigma_2)$$

and the following infinite commutative diagram Dg in $Mod(\Sigma)$ (which is in fact an ω-diagram too):

$$\begin{tikzcd}
A_1^0\restriction_{\varphi_1} \arrow[rr, "f_1^0\restriction_{\varphi_1}"] \arrow[dr, "h_0"'] & & A_1^1\restriction_{\varphi_1} \arrow[rr, "f_1^1\restriction_{\varphi_1}"] \arrow[dr, "h_1"] & & A_1^2\restriction_{\varphi_2} \arrow[dr, "h_2"] & & \cdots \\
& A_2^0\restriction_{\varphi_1} \arrow[ur, "g_0"] \arrow[rr, "f_2^0\restriction_{\varphi_2}"] & & A_2^1\restriction_{\varphi_2} \arrow[ur, "g_1"] \arrow[rr, "f_2^1\restriction_{\varphi_2}"] & & A_2^2\restriction_{\varphi_2} &
\end{tikzcd}$$

where $A_1^0 = A_1$, $A_2^0 = A_2$, $A_1^1 = B_1$, $A_2^1 = B_2$, $f_0^1 = g$, $f_0^2 = s$, $h_0 = h$, $g_0 = f$, $h_1 = p$, $\ldots$

Because the reduct functors preserve ω-colimits, the colimits of Ch_1 and Ch_2 in $Mod(\Sigma_1)$ and $Mod(\Sigma_2)$, with vertexes denoted N_1 and N_2, are mapped by $Mod(\varphi_1)$ and $Mod(\varphi_2)$ into colimits in $Mod(\Sigma)$ of the ω-diagrams $Mod(\varphi_1)(Ch_1)$ and $Mod(\varphi_2)(Ch_2)$. But $Mod(\varphi_1)(Ch_1)$ and $Mod(\varphi_2)(Ch_2)$ are final segments of the ω-diagram Dg, so $N_1\restriction_{\varphi_1}$ and $N_2\restriction_{\varphi_2}$ are, both, colimit vertexes of Dg in $Mod(\Sigma)$. Hence $N_1\restriction_{\varphi_1}$ and $N_2\restriction_{\varphi_2}$ are isomorphic. On the other hand, since model morphisms preserve satisfaction and $A_1^0 \models T_1$, $A_2^0 \models T_2$, it follows that $N_1 \models T_1$ and $N_2 \models T_2$. Because $\mathcal{S}$ lifts isomorphisms, we find two models M_1 and M_2 such that $M_1 \simeq N_1$ and $M_2 \simeq N_2$ (thus $M_1 \models T_1$ and $M_2 \models T_2$) and $M_1\restriction_{\varphi_1} = M_2\restriction_{\varphi_2}$. ■

Among the hypotheses in the above Consistency Theorem, all look quite natural, except for two of them:

- that of satisfaction preservation by the model morphisms and
- that of the square lifting isomorphisms.

While the second is just a technical assumption, the first one is rather interesting; since $\mathcal{I}$ has negations, it implies that any two models connected through a morphism are elementary equivalent. This seems like a harsh thing to ask; but this requirement is normal if the considered model morphisms are something like elementary embeddings, thus preserving satisfaction of all "parameterized sentences", in particular of all "plain" ones. Institutions tend to have "elementary" subinstitutions; and those who have, can import the Consistency Theorem from there.

Corollary 9. *In each of the institutions ElFOPL', ElIFOPL', ElPFOPL', FOPL', IFOPL', PFOPL', any weak amalgamation square which lifts isomorphisms is a Robinson square (hence also a CI square).*

Proof. We first claim that the institutions *ElFOPL'*, *ElIFOPL'*, and *ElPFOPL'* satisfy the conditions in Theorem 8. Let us check these conditions for *ElFOPL'*.

The elementary diagrams for *ElFOPL* were discussed in Section 2. It is straightforward to see that the same construction works for *ElFOPL'* too. The existence of pushout of signatures and compactness are well-known for *ElFOPL*, and are immediately inherited by *ElFOPL'*. Weak model-semi-exactness holds for *FOPL* because this institution is actually semi-exact (and even exact); and since the property only refers to models, common to *FOPL* and *ElFOPL*, it follows that *ElFOPL* is also weakly model-semi-exact; moreover, it is easy to see that, given a pushout of *FOPL* signatures $(\Sigma_2 \overset{\varphi_2}{\leftarrow} \Sigma \overset{\varphi_1}{\rightarrow} \Sigma_1, \Sigma_2 \overset{\varphi_2'}{\rightarrow} \Sigma' \overset{\varphi_1'}{\leftarrow} \Sigma_1)$ and two models $M_1 \in |Mod(\Sigma_1)|$, $M_2 \in |Mod(\Sigma_2)|$ such that $M_1\upharpoonright_{\varphi_1} = M_2\upharpoonright_{\varphi_2}$, if M_1 and M_2 have non-empty carriers on all sorts, then their common expansion $M' \in |Mod(\Sigma')|$ can be chosen with non-empty carriers on all sorts too; hence *ElFOPL'* is also weakly model-semi-exact.

The next discussion about elementary chains is valid for *ElFOPL'* (as a routine generalization of the unsorted case, with the carrier non-emptiness assumption imported from there), but also for *ElFOPL*. The existence of ω-colimits of signatures follows from Tarski's Elementary Chain Theorem (ECT) [11]. Let Σ be a signature and $(h_{i,j} : A_i \to A_j)_{i,j \in I\!N}$, denoted Dg, a diagram in $Mod_{ElFOPL'}(\Sigma)$, that is, with all morphisms $h_{i,j}$ being elementary embeddings. We can take the colimit of Dg in the category of model embeddings, let it be $(h_i : A_i \to A)_{i \in I\!N}$. According to ECT, all the h_i's are elementary. Moreover, $(h_i : A_i \to A)_{i \in I\!N}$ is actually the colimit of Dg in $Mod_{ElFOPL'}(\Sigma)$ too. Indeed, let $(g_i : A_i \to B)_{i \in I\!N}$ be another cocone of Dg in $Mod_{ElFOPL'}(\Sigma)$. According to the definition of Dg, there exists a unique embedding $f : A \to B$ such that $h_i; f = g_i$ for all $i \in I\!N$; but f is also elementary, since each parameterized sentence $e(a_1, \ldots, a_n)$ with parameters in A has all the parameters a_j, $j \in \{1, \ldots, n\}$, of the forms $h_i(b_j)$ for some large enough $i \in I\!N$; hence $A_A \models e(a_1, \ldots, a_n)$ iff $A_{iA_i} \models e(b_1, \ldots, b_n)$ iff $B_B \models e(g_i(b_1), \ldots, g_i(b_n))$ iff $B_B \models e(f(a_1), \ldots, f(a_n))$. Now, reduct functors along signature morphisms preserve elementary embeddings, as one can easily check; hence, because reduct functors preserve ω-colimits of embeddings in *FOPL'*, applying again ECT, we find that reduct functors preserve ω-colimits of models even when only elementary embeddings are taken into consideration as morphisms between models.

We shall finally check the existence of quantifications over $\iota_\Sigma(h)$, where $\Sigma = (S, F, P)$ is a signature and $A \xrightarrow{h} B$ is an elementary embedding. This time, our discussion is valid only for *ElFOPL'*. Let e be a sentence in $Sen(\Sigma_B)$. In order to show that $(\forall \iota_\Sigma(h))e$ is (equivalent to) a first-order sentence, let Σ' be the signature which

- includes the image $\iota_\Sigma(h)(\Sigma_A)$ of $\iota_\Sigma(h)$ (which is a copy of Σ_A included in Σ_B)
- and contains, for each $s \in S$, as extra constants of sort s all the elements in B_s that are not in the image of h_s and appear in e.

Since e is finitary, the extra constants are in finite number, and thus, if we consider the natural injective signature morphisms $\Sigma_A \xrightarrow{j} \Sigma' \xrightarrow{u} \Sigma_B$, where $j; u = \iota_\Sigma(h)$, we have the following:
(1) u is an inclusion of signatures, thus $Sen(\Sigma') \subseteq Sen(\Sigma_B)$, and $e \in Sen(\Sigma')$; moreover, like any signature inclusion, u is conservative (remember that all models are assumed to have non-empty carriers on each sort);
(2) $(\forall j)e$ is (equivalent to) a first-order sentence, because j is an injective signature morphism adding only a finite number of constants, all of which appearing in e;
(3) $(\forall \iota_\Sigma(h))e$ is equivalent to $(\forall j)e$. Indeed, "$(\forall j)e$ implies $(\forall \iota_\Sigma(h))e$" obviously holds. Conversely, assume $M \models (\forall \iota_\Sigma(h))e$ and let M' be a j-expansion of M. By the conservativeness of u, there exists M'' a u-expansion of M'. M'' is also a $\iota_\Sigma(h)$-expansion of M and $M'' \models e$. Thus $M'' \models u(e)$, i.e., $M' \models e$. Hence $M \models (\forall j)e$.

A very similar argument as the one above, but simpler, can be used to show the existence of quantifications over signature morphisms of the form $\iota_\Sigma(A)$.

The above arguments can be easily adapted to *ElIFOPL'* and *ElPFOPL'*. (For a proof of the Elementary Chain Theorem which can be adapted to *IFOPL'*, see [25], and for an institutional proof, which covers the cases of *IFOPL'* and *PFOPL'*, see [23].)

The conclusion of Theorem 8 involves only items (signature morphisms) which are the same in *FOPL'*, *IFOPL'*, and *PFOPL'* as in their "elementary" subinstitutions - so *FOPL'*, *IFOPL'*, and *PFOPL'* enjoy this property too. ∎

COROLLARY 10. *Let $\mathcal{I}$ be one of the institutions FOPL, IFOPL, PFOPL, ElFOPL, ElIFOPL, ElPFOPL, and let $\mathcal{S}$ be a w.a. square in $\mathcal{I}$ as in the figure of Definition 3. Then $\mathcal{S}$ is a Robinson square if one of the following conditions holds:*

1. *$\mathcal{I}$ is one of FOPL, ElFOPL and the set $\{s \in S \mid T_{F_s} = \emptyset\}$ is finite,*[9] *where $\Sigma = (S, F, P)$.*
2. *$\mathcal{I}$ is one of PFOPL, ElPFOPL and the set $\{s \in S \mid T_{F_s} = \emptyset\}$ is finite, F being the set of* total *operation symbols of Σ;*
3. *$\mathcal{I}$ is one of IFOPL, ElIFOPL.*

PROOF. The only delicate issue, different from the situation in Corollary 9, is in each case the existence of universal quantifications over $\iota_\Sigma(h)$ and $\iota_\Sigma(A)$. (1): Recall statements (1)-(3) from the proof of the fact that the institution *FOPL'* admits universal quantifications over $\iota_\Sigma(h)$ in Corollary 9. Using the same notations, but working in *FOPL* instead of *FOPL'*, we get that u is still conservative because: it is injective, all items outside its image are constants, and these are on sorts where some constants already existed (since, by the elementarity of h, for each sort s, A_s is empty iff B_s is empty); the rest of the argument for $\iota_\Sigma(h)$ is just like at Corollary 9.

The only problem left is the existence of universal quantification over $\iota_\Sigma(A)$. Let $e \in Sen(\Sigma_A)$. Similarly as before, we factor $\iota_\Sigma(A)$ as $u'; u$, where: $u' : \Sigma \to \Sigma'$ and $u : \Sigma' \to \Sigma_A$ are inclusions of signatures, and Σ' has only finitely many constants outside the image of u'. Then $(\forall u')e$ is equivalent to a first-order sentence, denoted e'. Define $\overline{S} = \{s \in S \mid A_s \neq \emptyset \text{ and } (T_F)_s = \emptyset\}$. $(\forall \iota_\Sigma(A))e$ is then equivalent to the first-order sentence $[\bigvee_{s \in \overline{S}} \neg(\exists x : s)x = x] \vee e'$, denoted e''. Indeed, let M be a Σ-model. We have two cases:
Case 1: There exists $s \in \overline{S}$ such that $M_s = \emptyset$. Then $M \models e''$ and, since M does not have any $\iota_\Sigma(A)$-expansion, M vacuously satisfies $(\forall \iota_\Sigma(A))e$.
Case 2: For each $s \in \overline{S}$, $M_s \neq \emptyset$. Assume first that $M \models e''$; then $M \not\models [\bigvee_{s \in \overline{S}} \neg(\exists x : s)x = x]$, so $M \models (\forall u')e$; let M'' be a $\iota_\Sigma(A)$-expansion of M; then $M'' \models e$ because $M''\upharpoonright_u$, as a u'-expansion of M, satisfies e. Conversely, assume that $M \models (\forall \iota_\Sigma(A))e$ and let M' be a u'-expansion of M; because

[9] This covers the cases of S being finite and of T_{F_s} being non-empty for each $s \in S$.

$M'_s \neq \emptyset$ for each $s \in S$, M' has a u-expansion M''; but $M'' \models e$, so $M' \models e$; thus, $M' \models (\forall u')e$, which implies $M \models e''$.
(2): Similar to (1).
(3): Identical to (1). Note that here we do not need finiteness of $\overline{S}$ in order to take the disjunction $\bigvee_{s \in \overline{S}} \neg(\exists x : s)x = x$. ■

5. A Syntactic Criterion for *FOPL* Robinson Consistency

We are going to use Corollaries 9 and 10 in order to prove a very general syntactic criterion for a *FOPL* or *FOPL'* signature square to be a Robinson square. By a "syntactic criterion" we mean one which uses only the structure of signature and signature morphisms, not involving the semantic concept of a model. Since in practice one usually deals with finite signatures, a syntactic criterion is easily checkable in an automatic fashion.

Let us consider either a *FOPL*-, or a *FOPL'*-, weak amalgamation square $\mathcal{S}$ as in the figure of Definition 3, with $\Sigma = (S, F, P)$, $\Sigma_1 = (S_1, F_1, P_1)$, $\Sigma_2 = (S_2, F_2, P_2)$, $\Sigma' = (S', F', P')$. If we take it to be a *FOPL* square, we also assume that $\{s \in S \mid T_{F s} = \emptyset\}$ is finite.

Without loss of generality, we assume that, within each signature, the sets of operation and relation symbols of different ranks are disjoint. That is, for each $w, w' \in S^*$ and $s, s' \in S$

- $(w, s) \neq (w', s')$ implies $F_{w,s} \cap F'_{w',s'} = \emptyset$;
- $w \neq w'$ implies $P_w \cap P'_{w'} = \emptyset$;
- $F_{w,s} \cap P_{w'} = \emptyset$;
- and similarly for Σ_1, Σ_2, Σ'.

We let φ_1 denote the extension of $S \xrightarrow{\varphi_1} S_1$ to $S^* \to S_1^*$; also we let $\varphi_1(w, s)$ denote the pair $(\varphi_1(w), \varphi_1(s))$ for each $(w, s) \in S^* \times S$; and similarly for φ_2.

Proposition 11. *$\mathcal{S}$ is a Robinson square (and also a Craig square) if the following four conditions hold:*
(C_1) For each $w, w' \in S^*$, $s, s' \in S$, $\sigma \in F_{w,s}$, $\sigma' \in F_{w',s'}$ such that $(w, s) \neq (w', s')$,
$[\varphi_1(w, s) = \varphi_1(w', s')$ and $\varphi_1(\sigma) = \varphi_1(\sigma')]$ implies $[\varphi_2(w, s) = \varphi_2(w', s')$ and $\varphi_2(\sigma) = \varphi_2(\sigma')]$.
(C_2) For each $w, w' \in S^*$, $s, s' \in S$, $\sigma \in F_{w,s}$ such that $(w, s) \neq (w', s')$,
$\varphi_1(w, s) = \varphi_1(w', s')$ implies the existence of $\sigma' \in F_{w',s'}$ such that $\varphi_1(\sigma) = \varphi_1(\sigma')$.
(C'_1) For each $w, w' \in S^*$, $R \in P_w$, $R' \in P_{w'}$ such that $w \neq w'$,

$[\varphi_1(w) = \varphi_1(w')$ and $\varphi_1(R) = \varphi_1(R')]$ implies $[\varphi_2(w) = \varphi_2(w')$ and $\varphi_2(R) = \varphi_2(R')]$.
(C_2') For each $w, w' \in S^*$, $R \in P_w$, such that $w \neq w'$,
$\varphi_1(w) = \varphi_1(w')$ implies the existence of $R' \in P_{w'}$ such that $\varphi_1(R) = \varphi_1(R')$.

Proof. By Corollaries 9 and 10, it is sufficient to prove that $\mathcal{S}$ lifts isomorphisms, that is: if $A_1 \in |Mod(\Sigma_1)|$, $D_2 \in |Mod(\Sigma_2)|$ such that $A_1\upharpoonright_{\varphi_1} \simeq D_2\upharpoonright_{\varphi_2}$, then there exist $B_1 \in |Mod(\Sigma_1)|$, $B_2 \in |Mod(\Sigma_2)|$ such that $A_1 \simeq B_1$, $D_2 \simeq B_2$, and $B_1\upharpoonright_{\varphi_1} = B_2\upharpoonright_{\varphi_2}$.

We are going to construct two models B_1 and B_2 as above. In our construction and throughout the proof, we shall totally ignore the relational part of the signatures, concentrating on operations. On the relational part, the situation is perfectly similar, using conditions (C_1') and (C_2') instead of (C_1) and (C_2).

We first take B_2 to be isomorphic to D_2 such that $card(B_{2s}) = card(B_{2s'})$ implies $B_{2s} = B_{2s'}$ for all $s, s' \in S_1$. Denote $A = A_1\upharpoonright_{\varphi_1}$ and $B = B_2\upharpoonright_{\varphi_2}$. Since $A \simeq D$ and $B_2 \simeq D_2$, we have $A \simeq B$. Let $A \xrightarrow{g} B$ be an isomorphism between A and B. By the construction of B_2, whenever $s, s' \in S$ with $\varphi_1(s) = \varphi_1(s')$, we have $B_s = B_{s'}$ (because, if $\varphi_1(s) = \varphi_1(s')$, then $card(B_s) = card(A_s) = card(A_{\varphi_1(s)}) = card(A_{s'}) = card(B_{s'})$, hence $card(B_{2\varphi_2(s)}) = card(B_{2\varphi_2(s')})$, hence $B_{2\varphi_2(s)} = B_{2\varphi_2(s')}$, hence $B_s = B_{s'}$).

We now define B_1.

1. Let the S_1-sorted set B_1 be:
   - $B_{1\varphi_1(s)} = B_s$, for each $s \in S$ (according to the above discussion, the definition of $B_{s'}$ with $\varphi_1(s) = s'$ does not depend on the choice of s)
   - $B_{1s'} = A_{1s'}$, for each $s' \in S_1 - \varphi_1(S)$
2. Fix $\theta : \varphi_1(S) \to S$ a "choice" function such that, for each $s' \in \varphi_1(S)$, $\varphi_1(\theta(s')) = s'$.
3. Let the S-sorted function $h : A_1 \to B_1$ be:
   - $h_{s'} : A_{1s'} \to B_{1s'}$, $h_{s'} = 1_{A_{1s'}}$, for each $s' \in S_1 - \varphi_1(S)$;
   - $h_{s'} : A_{1s'} \to B_{1s'}$, $h_{s'} = g_{\theta(s')}$, for each $s' \in \varphi_1(S)$ (notice that, if $s \in S$, $h_{\varphi_1(s)} = g_{\theta(\varphi_1(s))}$). h is obviously an S-sorted bijection.
4. Define a Σ_1-structure on B_1 by copying it through h from A_1: for each $w \in S_1^*$, $s \in S_1$, $\sigma \in F_{1w,s}$, $z \in B_{1w}$, let $B_{1\sigma}(z) = h_s^{-1}(A_{1\sigma}(h_w(z)))$.

Obviously, B_1 is a Σ_1-model and $A_1 \xrightarrow{h} B_1$ is a Σ_1-isomorphism. All we need to show is that $B_1\upharpoonright_{\varphi_1} = B$.

1. On sorts: if $s \in S$, then $B_{1\varphi_1(s)} = B_s$ by definition.
2. On operations: let $w' \in S^*$, $s' \in S$, $\sigma' \in F_{w',s'}$. Let $w = \theta(\varphi_1(w'))$ and $s = \theta(\varphi_1(s'))$, where $\theta : \varphi_1(S)^* \to S^*$ is the symbol-wise extension of $\theta : \varphi_1(S) \to S$. Because $\varphi_1(w) = \varphi_1(w')$ and $\varphi_1(s) = \varphi_1(s')$, we have $B_{1\varphi_1(w')} = B_w = B_{w'}$ and $B_{1\varphi_1(s')} = B_s = B_{s'}$, so the operations $B_{1\varphi_1(\sigma')}$ and $B_{\sigma'}$ (which we want to prove equal) have the same domain and codomain. There are two cases:
   Case 1: $(w,s) = (w',s')$. Then, by definition, $B_{1\varphi_1(\sigma')}$ is the copy through $(h_{\varphi_1(w)}, h_{\varphi_1(s)})$, that is, through (g_w, g_s), of $A_{1\varphi_1(\sigma')}$; but, since g is a Σ-isomorphism, $B_{\sigma'}$ is also the copy through (g_w, g_s) of $A_{\sigma'} = A_{1\varphi_1(\sigma')}$. Hence $B_{\sigma'} = B_{1\varphi_1(\sigma')}$.
   Case 2: $(w',s') \neq (w,s)$. Since $\varphi_1(w',s') = \varphi_1(w,s)$ and $\sigma' \in F_{w',s'}$, we apply (C_2) to get $\sigma \in F_{w,s}$ such that $\varphi_1(\sigma) = \varphi_1(\sigma')$. Moreover, by (C_1), $\varphi_2(w,s) = \varphi_2(w',s')$ and $\varphi_2(\sigma) = \varphi_2(\sigma')$. Thus $A_\sigma = A_{\sigma'}$ and $B_\sigma = B_{\sigma'}$. $B_{1\varphi_1(\sigma')} : B_w \to B_s$ is, by definition the copy through (g_w, g_s) of $A_{1\varphi_1(\sigma')} = A_{\sigma'} = A_\sigma$; so $B_{1\varphi_1(\sigma')} = B_\sigma = B_{\sigma'}$. ■

Remark 12. *1. A consequence of $(C_1) + (C_2)$ is a conditional kernel inclusion between φ_1 and φ_2: if $(w,s),(w',s') \in S^* \times S$ are such that $F_{w,s} \cup F_{w',s'} \neq \emptyset$, then $\varphi_1(w,s) = \varphi_1(w',s')$ implies $\varphi_2(w,s) = \varphi_2(w',s')$.*

2. The criterion from Proposition 11 is indeed syntactical, because the property of being a weak amalgamation square is syntactically describable: a square is a w.a. square iff it is a composition between a pushout square and a conservative signature morphism; furthermore, a signature morphism $\varphi : (S,F,P) \to (S',F',P')$ is conservative iff it is injective on [sort, operations and relation] symbols and, for each $s \in S$, $(T_F)_s = \emptyset$ iff $(T_{F'})_{\varphi(s)} = \emptyset$; and pushout squares are also syntactically describable.

Corollary 13. *If either φ_1 or φ_2 in $\mathcal{S}$ is injective on sorts, then $\mathcal{S}$ is a Robinson (and also Craig) square.*

Proof. If φ_1 is injective on sorts, then all the conditions in Proposition 11 are trivially true, since it is never the case that $[(w,s) \neq (w',s')$ and $\varphi_1(w,s) = \varphi_1(w',s')]$, or $[w \neq w'$ and $\varphi_1(w) = \varphi_1(w')]$.

The case of φ_2 injective on sorts is perfectly symmetric to the previous one, thus the result follows from the symmetry of each of the two properties "w.a. square" and "Robinson square". ■

Note that Corollary 13 also has a direct proof from Corollaries 9 and 10, since $\mathcal{S}$ lifts isomorphisms whenever φ_1 or φ_2 is injective on sorts.

Example 14. Let $\mathcal{S}$ be the commutative *FOPL*-square as in the figure of Definition 3, defined as follows: $\Sigma = (\{s_1, s_2\}, \{d_1 :\to s_1, d_2 :\to s_2\})$, $\Sigma_1 = (\{s\}, \{d_1, d_2 :\to s\})$, $\Sigma_2 = (\{s\}, \{d :\to s\})$, $\Sigma' = (\{s\}, \{d :\to s\})$, all the morphisms mapping all sorts into s, φ_1 mapping d_1 and d_2 into themselves, and all the other morphisms mapping all the operation symbols into d. In [6], it is shown that $\mathcal{S}$ is not a CI square. To see this, let $E_1 = \{\neg(d_1 = d_2)\}$ and $E_2 = \{\neg(d = d)\}$. Then obviously $\varphi'_1(E_1) \models \varphi'_2(E_2)$, but E_1 and E_2 have no Σ-interpolant. Indeed, assume that there exists a set E of Σ-sentences such that $E_1 \models \varphi_1(E)$ and $\varphi_2(E) \models E_2$; let A be the Σ_1-model with $A_s = \{0, 1\}$, such that $A_{d_1} = 0$ and $A_{d_2} = 1$. Let B denote $A{\upharpoonright}_{\varphi_1}$. We have that $B_{s_1} = B_{s_2} = \{0, 1\}$, $B_{d_1} = 0$, $B_{d_2} = 1$. Because $A \models E_1$ and $E_1 \models \varphi_1(E)$, it holds that $B \models E$. Define the Σ-model C similarly to B, just that one takes $C_{d_1} = C_{d_2} = 0$. Now, C and B are isomorphic (notice that a and b are constants of *different* sorts in Σ), so $C \models E$; but C admits a φ_2-expansion D, and, because $\varphi_2(E) \models E_2$, $D \models E_2$, which is a contradiction, since no Σ_2-model can satisfy $\neg(d = d)$. According to Proposition 6, $\mathcal{S}$ is not a Robinson square either. The problem with this square, as depicted in [6], is that it has signature morphisms which are non-injective on sorts. In the light of Corollary 13, we can be more precise: the problem is that *none* of φ_1 and φ_2 is injective on sorts. Even more precisely, according to Corollary 10, $\mathcal{S}$ does not lift isomorphisms; in particular, it does not lift the unique isomorphism between $B = A{\upharpoonright}_{\varphi_1}$ and $C = D{\upharpoonright}_{\varphi_2}$ above.

Remark 15. *Proposition 11 (and hence Corollary 13 too) holds for IFOPL and IFOPL' with the same proof. Also, if we duplicate in Proposition 11 the conditions (C1) and (C2) to account separately for total and partial operation symbols, we obtain a similar criterion for PFOPL and PFOPL', with an almost identical proof.*

6. Related Work and Concluding Remarks

On Robinson Consistency

Robinson Consistency is broadly studied in connection with CIP, compactness, and other logical properties in a series of papers among which [36, 37, 35], in the framework offered by *abstract model-theoretic logics* [3]. An interesting phenomenon discovered there is that RCP implies compactness, which does not seem to hold for the more abstract case of institutions. No proof of RCP "from scratch" (i.e., without assuming CIP) is given there; moreover, as it is the case of all works within abstract model-theoretic logics, only the situation of language inclusions is considered. The paper [49] formulates for the first time an institutional version of RCP and proves its equivalence to CIP in compact institutions admitting negations and finite conjunctions. Another formulation of RCP, inspired by the one from [36, 37], is given in [47] in the context of preinstitutions, where there is also proved the equivalence of RCP with CIP assuming, instead of compactness and finite conjunctions, a rather strong property called *elementary expansion.* Recent work in [1] states RCP and gives some equivalent formulations for it using a syntactic notion of consistency of a theory, by not requiring the theory to have a model, but to not entail every sentence - this definition has the advantage that makes non-trivial sense for equational logics too.

Our Theorem 8 seems to be the first generalization of the Robinson Consistency Theorem to a fairly abstract logical framework. However, our result is "RCP-specific" only w.r.t. its proof technique, and not to its content, since it assumes some hypotheses under which RCP is equivalent to CIP.

On Craig Interpolation

After the original formulation and proof given in [13] for the unsorted first-order logic, several generalizations occur in the literature, among which that of [22] for many-sorted first-order logic, in the case of union and intersection of languages. However, the conclusion of studying various model-theoretic logics that extend first-order logic was that "interpolation is indeed [a] rare [property in logical systems]" ([3], page 68). The paper [44] proves CIP for many-sorted equational logic, stating interpolation on *sets of sentences* instead of sentences. The first institutional formulation of CIP appears in [49] and uses arbitrary pushout squares of signatures. In [43], some general axiomatizability-based criteria are provided for a pullback of categories in order to satisfy a property which generalizes CIP when the categories are instanciated to classes of models over some signatures; this result covers the cases of many versions of equational logic. Another general result, proving CIP about institutions admitting Birkhoff-style axiomatizability and covering cases beyond equational logic, can be found in [18]. Moreover, [47] and [19] provide means of transporting CIP across translations of logics (institution morphisms and comorphisms). A stronger but more specialized result, concerned only with the many-sorted first-order logic and its partial-operation variant (which are the underlying logics in many specification languages, including CASL [12]), can be found in [6], where it is proved that if both starting pushout morphisms are *injective on sort names*, then CIP holds for the considered pushout square.

The present paper brings the following contributions from the CIP point of view:

1. Our Consistency Theorem 8 solves an open problem raised in [49], showing that CIP holds in institutions with additional requirements very similar to those of *abstract algebraic institutions*. This result complements the one in [18], which derives CIP from Birkhoff-like axiomatizability properties assumed on the classes of models and hence works particularly well for logics with strong axiomatizability properties; instead, our result is suitable for sufficiently expressive logics, not requiring axiomatizability, but needing appropriate machinery for the method of diagrams.
2. Our Corollary 13 (see also Remark 15) improves the result in [6] (the strongest known so far for *FOPL*), by showing that *only one* of the pushout morphisms needs to be injective on sorts in order for CIP to hold.
3. In fact, our Proposition 11, significantly more general than Corollary 13, pushes the syntactic criterion for CIP to a form which we think is close to the limit (i.e., to an "iff" criterion).
4. Finally, our interpolation results for the infinitary logical system $L_{\infty,\omega}$ seem to be new and of potential interest in categorical logic.

Acknowledgements. This paper grew out of some lectures given by Răzvan Diaconescu on the beautiful topic of institutional abstract model theory, while the authors were students at Şcoala Normală Superioară, Romania. Răzvan Diaconescu also encouraged us and gave us feedback that significantly improved the paper. We also thank Traian Şerbănuţă, Marius Petria and Cristian Cucu for their insightful suggestions. The anonymous referee's remarks were also of valuable help.

References

[1] Aiguier, M., and F. Barbier, 'On a generalization of some results of model theory: Robinson, Craig and Beth theorems'. Submitted. (Ask authors for current version at {aiguier,barbier}lami.univ-evry.fr).

[2] Andréka, H., I. Németi, and I. Sain, 'Craig property of a logic and decomposability of theories', in P. Dekker and M. Stokhof (eds.), *Proceedings of the Ninth Amsterdam Colloquium*, vol. 3229 of *Lecture Notes in Computer Science*, 1993, pp. 87–93.

[3] Barwise, J., and J. Feferman, *Model-Theoretic Logics*, Springer, 1985.

[4] Bergstra, J., J. Heering, and P. Klint, 'Module algebra', *Journal of the Association for Computing Machinery*, 37(2):335–372, 1990.

[5] BICARREGUI, D. G. J., T. DIMITRAKOS and T. MAIBAUM, 'Interpolation in practical formal development', *Logic Journal of the IGPL*, 9(1):231–243, 2001.

[6] BORZYSZKOWSKI, T., 'Generalized interpolation in CASL, *Inf. Process. Lett.*, 76(1-2):19–24, 2000.

[7] BORZYSZKOWSKI, T., 'Logical systems for structured specifications', *Theor. Comput. Sci.*, 286(2):197–245, 2002.

[8] BURMEISTER, P., *A Model Theoretic Oriented Appraoch to Partial Algebras*, Akademie-Verlag Berlin, 1986.

[9] BURSTALL, R., and J. GOGUEN, 'Semantics of Clear', Unpublished notes handed out at the 1978 Symposium on Algebra and Applications. Stefan Banach Center, Warsaw, Poland, 1978.

[10] CENGARLE, M. V., *Formal Specifications with High-Order Parametrization*, PhD thesis, Institute for Informatics, Ludwig-Maximilians University, Munich, 1994.

[11] CHANG, C. C., and H. J. KEISLER, *Model Theory*, North Holland, Amsterdam, 1973.

[12] CoFI task group on semantics, CASL — The Common Algebraic Specification Language, Semantics, www.brics.dk/Projects/CoFI/Documents/CASL, July 1999.

[13] CRAIG, W., 'Linear reasoning. A new form of the Herbrand-Gentzen Theorem', *Journal of Symbolic Logic*, 22:250–268, 1957.

[14] DIACONESCU, R., *Institution-independent Model Theory*, Book draft. (Ask author for current version at Razvan.Diaconescu@imar.ro).

[15] DIACONESCU, R., 'Institution-independent ultraproducts', *Fundamenta Informaticae*, 55(3-4):321–348, 2003.

[16] DIACONESCU, R., 'Elementary diagrams in institutions'. *Journal of Logic and Computation*, 14(5):651–674, 2004.

[17] DIACONESCU, R., 'Herbrand theorems in arbitrary institutions', *Inf. Process. Lett.*, 90:29–37, 2004.

[18] DIACONESCU, R., 'An institution-independent proof of Craig interpolation theorem', *Studia Logica*, 77(1):59–79, 2004.

[19] DIACONESCU, R., 'Interpolation in Grothendieck institutions', *Theor. Comput. Sci.*, 311:439–461, 2004.

[20] DIACONESCU, R., J. GOGUEN, and P. STEFANEAS, 'Logical support for modularization', in G. Huet and G. Plotkin (eds.), *Logical Environments*, Cambridge, 1993, pp. 83–130.

[21] DIMITRAKOS, T., and T. MAIBAUM, 'On a generalised modularization theorem', *Inf. Process. Lett.*, 74(2):65–71, 2000.

[22] FEFERMAN, S., 'Lectures on proof theory', in M. Lob (ed.), *Proceedings of the Summer School in Logic, Lecture Notes in Mathematics*, vol. 70, pp. 1–107, 1968.

[23] GĂINĂ, D., and A. POPESCU, 'An institution-independent generalization of Tarski's elementary chain theorem', *Journal of Logic and Computation*, To appear.

[24] GOGUEN, J., and R. BURSTALL, 'Institutions: Abstract model theory for specification and programming', *Journal of the Association for Computing Machinery*, 39(1):95–146, January 1992.

[25] KEISLER, H. J., *Model Theory for Infinitary Logic*, North-Holland, 1971.

[26] MAC LANE, S., *Categories for the Working Mathematician*, Springer, 1971.

[27] Madarász, J. X., 'Interpolation and amalgamation; Pushing the limits. Part ii', *Studia Logica*, 62(1):1–19, 1999.

[28] Makkai, M., and R. Paré, *Accessible Categories: The Foundations of Categorical Model Theory*, Providence, 1989.

[29] Makkai, M., and G. Reyes, 'First order categorical logic', *Lecture Notes in Mathematics*, (611), 1977.

[30] McMillan, K. L., 'Interpolation and sat-based model checking', in *Computed Aided Verification 2003*, vol. 2725 of *Lecture Notes in Computer Science*, Springer-Verlag, 2003, pp. 1–13.

[31] McMillan, K. L., 'An interpolating theorem prover', in *TACAS*, vol. 2988, 2004, pp. 16–30.

[32] Meseguer, J., 'General logics'. in H.-D. Ebbinghaus et al. (eds.), *Proceedings, Logic Colloquium 1987*, North-Holland, 1989, pp. 275–329.

[33] Monk, J. D., *Mathematical Logic*, Springer-Verlag, 1976.

[34] Mossakowski, T., J. Goguen, R. Diaconescu, and A. Tarlecki, 'What is a logic?', in J.-Y. Beziau (ed.), *Logica Universalis*, Birkhauser, 2005, pp. 113–133.

[35] Mundici, D., 'Compactness + Craig interpolation = Robinson consistency in any logic', Manuscript, University of Florence, 1979.

[36] Mundici, D., 'Robinson consistency theorem in soft model theory', *Atti Accad. Nat. Lincei. Redingoti*, 67:383–386, 1979.

[37] Mundici, D., 'Robinson's consistency theorem in soft model theory', *Trans. of the AMS*, 263:231–241, 1981.

[38] Nelson, G., and D. C. Oppen, 'Simplification by cooperating decision procedures', *ACM Trans. Program. Lang. Syst.*, 1(2):245–257, 1979.

[39] Oppen, D. C., 'Complexity, convexity and combinations of theories', *Theoretical Computer Science*, 12:291–302, 1980.

[40] Petria, M., and R. Diaconescu, 'Abstract Beth definability institutionally', Submitted. (Ask authors for current version at Razvan.Diaconescu@imar.ro).

[41] Reichel, H., *Structural Induction on Partial Algebras*, Akademie-Verlag Berlin, 1984.

[42] Robinson, A., 'A result on consistency and its application to the theory of definition', vol. 59 of *Konikl. Ned. Akad. Wetenschap. (Amsterdam) Proc* $\{A\}$, 1956, pp. 47–58.

[43] Roşu, G., and J. Goguen, 'On equational Craig interpolation', *Journal of Universal Computer Science*, 6(1):194–200, 2000.

[44] Rodenburg, P. H., 'A simple algebraic proof of the equational interpolation theorem', *Algebra Universalis*, 28:48–51, 1991.

[45] Sain, I., 'Beth's and Craig's properties via epimorphisms and amalgamation in algebraic logic', in *Algebraic Logic and Universal Algebra in Computer Science*, 1988, pp. 209–225.

[46] Salibra, A., and G. Scollo, 'A soft stairway to institutions', in M. Bidoit and C. Choppy (eds.), *Recent Trends in Data Type Specification*, vol. 655 of *Lecture Notes in Computer Science*, Springer-Verlag, 1992, pp. 310–329.

[47] Salibra, A., and G. Scollo, 'Interpolation and compactness in categories of pre-institutions', *Math. Struct. in Comp. Science*, 6:261–286, 1996.

[48] SANNELLA, D., and A. TARLECKI, 'Specifications in an arbitrary institution', *Information and Control*, 76:165–210, 1988.

[49] TARLECKI, A., 'Bits and pieces of the theory of institutions', in D. Pitt, S. Abramsky, A. Poigné, and D. Rydeheard (eds.), *Proceedings, Summer Workshop on Category Theory and Computer Programming, Lecture Notes in Computer Science*, vol. 240, Springer, 1986, pp. 334–360.

[50] TARLECKI, A., 'On the existence of free models in abstract algebraic institutions', *Theoretical Computer Science*, 37:269–304, 1986.

[51] TARLECKI, A., 'Quasi-varieties in abstract algebraic institutions', *Journal of Computer and System Sciences*, 33(3):333–360, 1986.

[52] TINELLI, C., and C. ZARBA, 'Combining decision procedures for sorted theories', in J. J. Alferes and J. A. Leite (eds.), *Logics in Artificial Intelligence*, vol. 3229 of *Lecture Notes in Computer Science*, Springer, 2004, pp. 641–653.

[53] WEISS, W. A. R., and C. D'MELLO, 'Fundamentals of model theory', Lecture Notes, University of Toronto, 1997.

[54] WIRSING, M., 'Structured specifications: Syntax, semantics and proof calculus', in F. Bauer, W. Bauer, and H. Schwichtenberg (eds.), *Logic and Algebra of Specification*, vol. 94 of *NATO ASI Series F: Computer and System Sciences*, Springer-Verlag, 1991, pp. 411–442.

D. GĂINĂ AND A. POPESCU
Şcoala Normală Superioară, Romania
Also: Dept. of Fundamentals of Computer Science
Faculty of Mathematics and Informatics
University of Bucharest
Emails: {gaina_daniel,uuomul}@yahoo.com

Journal of Logic and Computation Advance Access published August 12, 2006

An Institution-independent Generalization of Tarski's Elementary Chain Theorem

DANIEL GĂINĂ and ANDREI POPESCU, *Department of Fundamentals of Computer Science, Faculty of Mathematics, University of Bucharest.*
Email: gaina_daniel@yahoo.com; uuomul@yahoo.com

Abstract

We prove an institutional version of Tarski's *elementary chain theorem* applicable to a whole plethora of 'first-order-accessible' logics, which are, roughly speaking, logics whose sentences can be constructed from atomic formulae by means of classical first-order connectives and quantifiers. These include the *unconditional equational*, *positive*, $(\Pi \cup \Sigma)^0_n$ and *full* first-order logics, as well as less conventional logics, used in computer science, such as *hidden* or *rewriting logic*.

Keywords: Institution, elementary morphism, elementary chain property.

1 Introduction

The notion of elementary embedding is an important one in classical first-order model theory [4]. Elementary chains (i.e. chains of elementary embeddings) are known to be a fundamental proof tool for results regarding preservation of satisfaction, axiomatizability, Robinson consistency, Craig interpolation [4], saturated models, stability, categoricity in power [30] and many others. The extension of elementary embeddings to infinitary logics [18, 21, 17] reveals the need in mathematical logic for accommodating this notion in other logical systems too. And the monograph [18] actually shows that this is a natural and very fruitful thing to do.

The present paper introduces and studies abstract notions of elementary embedding and elementary chain, in the framework of institutions [14], and points out many particular cases. Two aspects motivate and justify our study:

- The mentioned importance of elementary embeddings in model theory; and
- The logic-independent status of our concepts and results.

Besides its intrinsic abstract model-theoretic contribution, our study might be of interest for the theory of formal specifications, where a logic-independent view is desirable for as long as possible in the specification process and where structural properties usually approachable by means of elementary chains, such as Craig interpolation or axiomatizability, are crucial.

Institutions are abstract logical frameworks that provide a category of signatures (languages) and signature morphisms (language translations), and, for each signature, a set of sentences, a category of models and a satisfaction relation. Sentences have *translations*, and models have *reducts*, along signature morphisms; the translations and reducts express the sentence and model modifications under *change of notation* from one language to another. Satisfaction is required to be *invariant* under change of notation. More abstract than the *general logics* of [1], institutions were introduced as frameworks for building model theory for computer science, in a logic-independent way. Thus, general institutional results were applicable to the diversity of logical systems used in computer science. Besides their

doi:10.1093/logcom/exl006

great generality, another important feature of institutions, not present, or poorly present, in other abstract frameworks, is the flexible support for language translations. This feature, particularly useful in formal specification and the semantics of programming languages, is also interesting from a logical point of view. As shown in [34, 35, 7, 10, 8, 6] and other places, signature morphisms (language translations) turn out to be a very insightful tool for finding *concrete structure* in the core of abstract logic, for example, any institution hides inside a 'first-order logic', which can be uncovered by means of basic sentences, logical connectives, and quantifications over some signature morphisms (see Sections 2 and 3).[1]

Our institutional notion of elementary embedding (that we call 'elementary morphism') also uses signature morphisms in an essential way, by defining elementarity as *preservation of satisfaction in expansions along certain signature morphisms*. Recall that, classically, an elementary embedding between two models A and B of the same language is a model-embedding $A \xrightarrow{h} B$ such that for each formula $e(x_1, \ldots, x_n)$ and each sequence $a_1, \ldots, a_n \in A$, $A \models e(a_1, \ldots, a_n)$ iff $B \models e(h(a_1), \ldots, h(a_n))$. Because of the existence of negations, the 'iff' in the preceding can be replaced by 'implies'. In order to abstract away this concept, we follow an idea originating in [34] that treats (non-closed) Σ-formulae as sentences in signature 'extensions' $\varphi : \Sigma \to \Sigma'$. And we use quasi-representable signature morphisms [6] to capture the requirement that the 'extension' only adds 'constants'. (φ is quasi-representable if, for each $A' \in |Mod(\Sigma')|$, the canonical functor $A'/Mod(\varphi) : A'/Mod(\Sigma'), \to A'\upharpoonright_{\varphi}/Mod(\Sigma)$ is an isomorphism of categories; hence, giving a Σ'-morphism of source A' is equivalent to giving a Σ-morphism of source $A'\upharpoonright_{\varphi}$. This situation has the following intuitive explanation: all the 'extra items' of Σ' with respect to Σ being constant symbols, a Σ-morphism $A \xrightarrow{h} B$ can have only one Σ'-expansion of given source A'.) Quasi-representability of signature morphisms is a weakening of the concept of representability introduced in [10]. By fixing a class $\mathcal{Q}$ of quasi-representable signature morphisms, we call a Σ-morphism $A \xrightarrow{h} B$ $\mathcal{Q}$-*elementary* if for each signature morphism $\varphi : \Sigma \to \Sigma'$ in $\mathcal{Q}$, each φ-expansion $A' \xrightarrow{h'} B'$ of h and each Σ-sentence e', $A' \models e'$ implies $B' \models e'$. (See Section 4 for a detailed motivation of this definition.)

One can alternatively define elementary morphisms by elementary diagrams. Classically, the *elementary diagram* [4] $EDg(A)$ of a model A is the set of all sentences in $\Sigma(A)$ (the language Σ of A extended with all elements of A as constants) that are true in A. Then a model inclusion $A \xrightarrow{h} B$ is elementary iff $EDg(A) \subseteq EDg(B)$. Thanks to a recent concept of institutional diagram [8], we can also define institutionally *elementary morphisms by diagrams* (we abbreviate these as *d-elementary morphisms*). The diagrams of [8] provide, for each signature Σ and Σ-model A, a parameterized signature extension $\iota_\Sigma(A) : \Sigma \to \Sigma_A$ and a self-parameterized $\iota_\Sigma(A)$-expansion of A. In addition, the diagrams are 'functorial', i.e. they have corresponding structure for signature and model morphisms; in particular, any Σ-morphism $A \xrightarrow{h} B$ yields a signature morphism $\iota_\Sigma(h) : \Sigma_A \to \Sigma_B$ such that $\iota_\Sigma(A); \iota_\Sigma(h) = \iota_\Sigma(B)$. We call h d-elementary if, for all Σ_A-sentences e', $A_A \models e'$ implies $B_B \models \iota_\Sigma(h)(e')$. d-elementarity is expressible more compactly than $\mathcal{Q}$-elementarity, but requires an amount of rather evolved extra structure on institutions.

Here is the structure of this article. Section 2 recalls some notions regarding categories, institutions and diagrams. Section 3 discusses and exemplifies the concepts, central in this

[1] Compare this natural appearance of a first-order sublogic with the need to explicitly postulate the existence of such a sublogic in the context of general logics of [1].

article, of (finitely) representable and (finitely) quasi-representable signature morphism. Section 4 introduces elementary (model) morphisms with respect to a class of quasi-representable signature morphisms. Section 5 proves an institutional version of Tarski's elementary chain theorem, in the following slightly stronger form: elementary morphisms are closed under directed colimits. Section 6 introduces an alternative, diagrammatic version of elementary morphism and shows its equivalence to the previous notion under certain mild conditions (diagrams being in the considered class of quasi-representable morphisms and another condition that we call normality). Some concluding remarks end the article.

2 Preliminaries

2.1 Categories

We assume that the reader is familiar with basic categorical notions like functor, natural transformation, colimit, comma category, etc. A standard textbook on the topic is [19]. We are going to use the terminology from there, with a few exceptions that we point out in the following text. We use both the terms 'morphism' and 'arrow' to refer morphisms of a category. Composition of morphisms and functors is denoted using the symbol ';' and is considered in diagrammatic order.

Let $\mathcal{C}$ and $\mathcal{C}'$ be two categories. Given an object $A \in |\mathcal{C}|$, the *comma category of objects in $\mathcal{C}$ under A* is denoted $A/\mathcal{C}$. Recall that the objects of this category are pairs (h, B), where $B \in |\mathcal{C}|$ and $A \xrightarrow{h} B$ is a morphism in $\mathcal{C}$. Throughout the article we might let either $(A \xrightarrow{h} B, B)$ or (h, B) indicate objects in $A/\mathcal{C}$. A morphism in $A/\mathcal{C}$ between two objects (h, B) and (g, D) is just a morphism $B \xrightarrow{f} D$ in $\mathcal{C}$ such that $h; f = g$ in $\mathcal{C}$. There exists a canonical forgetful functor between $A/\mathcal{C}$ and $\mathcal{C}$, mapping each (h, B) to B and each $f : (h, B) \to (g, D)$ to $f : B \to D$. Also, if $F : \mathcal{C}' \to \mathcal{C}$ is a functor, $A \in |\mathcal{C}|$, $A' \in |\mathcal{C}'|$, and $A \xrightarrow{u} F(A')$ is in $\mathcal{C}$, then there exists a canonical functor $u/F : A'/\mathcal{C}' \to A/\mathcal{C}$ mapping each $(A' \xrightarrow{h} B, B)$ to $(u; F(h), F(B))$ and each $f : (h, B) \to (g, D)$ to $F(f) : (u; F(h), F(B)) \to (u; F(g), F(D))$. If $\mathcal{C} = \mathcal{C}'$ and F is the identity functor $1_\mathcal{C}$, we write $u/\mathcal{C}$ instead of u/F; and if $F(A') = A$ and $u = 1_A$, we write A'/F instead of u/F.

Let $\mathcal{C}$ and $\mathcal{S}$ be two categories such that $\mathcal{S}$ is small. A functor $D : \mathcal{S} \to \mathcal{C}$ is also called a *diagram*. We usually identify a diagram $D : \mathcal{S} \to \mathcal{C}$ with its image in $\mathcal{C}$, $D(\mathcal{S})$. A *cocone* of D is a natural transformation $\mu : D \Longrightarrow V$ between the functor D and the constant functor mapping all objects to V and all morphisms to 1_V; V is an object in $\mathcal{C}$, the *vertex* of the colimit, and the components of μ are the *structural morphisms* of the colimit. Any partially ordered set $(I, \leq)$ can be regarded as a category in the obvious way, with the arrows being pairs $i \leq j$. A non-empty partially ordered set $(I, \leq)$ is said to be *directed* if for all $i, j \in I$, there exists $k \in I$ such that $i \leq k$ and $j \leq k$, and is called a *chain* if the order $\leq$ is total. A diagram defined on a directed set (on a chain) shall be called *directed diagram (chain diagram)* and a colimit of such a diagram, *directed colimit (chain colimit)*. A *final subset* of a directed partially ordered set $(I, \leq)$ is a subset K of I such that for all $i \in I$, there exists $k \in K$ such that $i \leq k$. For instance, given $i \in I$, $I_i = \{j \in I \mid i \leq j\}$ is a final subset of $(I, \leq)$. A *sub-diagram* of a directed diagram $D : (I, \leq) \to \mathcal{C}$ is the restriction of D to $(K, \leq)$, where K is a subset of $(I, \leq)$; the sub-diagram is said to be *final* if K is final. An object A in a category $\mathcal{C}$ is called *finitely presented* if for each directed diagram $D : (I, \leq) \to \mathcal{C}$ with colimit $\{Di \xrightarrow{\mu i} B\}_{i \in I}$, and for each morphism $A \xrightarrow{h} B$, there exists $j \in I$ and $A \xrightarrow{g} D_j$ such that $g; \mu_j = h$.

Let $\mathcal{C}'$ be a subcategory of $\mathcal{C}$. $\mathcal{C}'$ is called a *broad subcategory* if it contains all the objects of $\mathcal{C}$. $\mathcal{C}'$ is said to be *closed under directed colimits (chain colimits)* if for any directed diagram (chain diagram) $D : (I, \leq) \to \mathcal{C}$ such that $D(i \leq j)$ is in $\mathcal{C}'$ for all $i \leq j$, any colimit $\{Di \xrightarrow{\mu} B\}_{i \in I}$ of D has all the structural morphisms μ_i in $\mathcal{C}'$. $\mathcal{C}'$ is said to be *closed under pushouts* if for each pushout $(A_2 \xleftarrow{h_2} A \xrightarrow{h_1} A_1, A_2 \xrightarrow{h_1'} A' \xleftarrow{h_2'} A_1)$ in $\mathcal{C}$, h_1' is in $\mathcal{C}'$ whenever h_1 is in $\mathcal{C}'$. Note that the notion of 'closed under' that we adopt is stronger for pushouts than for directed or chain colimits. The following lemma is proved in [19].

Lemma 1
Let $(I, \leq)$ be a directed set, $\mathcal{C}$ a category, $D : (I, \leq) \to \mathcal{C}$ a diagram, and $\{Di \xrightarrow{\mu} A\}_{i \in I}$ its colimit. If K is a final subset of $(I, \leq)$, then $\{Di \xrightarrow{\mu} A\}_{i \in K}$ is a colimit of the corresponding final sub-diagram of D.

2.2 Institutions

Institutions were introduced in [14] with the original goal of providing an abstract, logic-independent framework for algebraic specifications of computer science systems. By isolating the essence of a logical system in the abstract *satisfaction relation*, these structures achieve an appropriate level of generality for the development of abstract model theory, as shown by a whole series of (old and new) papers: [34–36, 31, 32, 7, 8, 10, 9, 13, 27]. See also [26] for an up-to-date discussion on institutions as abstract logics.

An institution [14] consists of:

(1) A category *Sign*, whose objects are called *signatures*;
(2) A functor $Sen : Sign \to Set$, providing for each signature a set whose elements are called *(Σ-)sentences;*
(3) A functor $Mod : Sign \to Cat^{op}$, providing for each signature Σ a category whose objects are called *(Σ-)models* and whose arrows are called *(Σ-)morphisms;* and
(4) A relation $\models_\Sigma \subseteq |Mod(\Sigma)| \times Sen(\Sigma)$ for each $\Sigma \in |Sign|$, called *(Σ-)satisfaction*, such that for each morphism $\varphi : \Sigma \to \Sigma'$ in *Sign*, the *satisfaction condition*

$$M' \models_{\Sigma'} Sen(\varphi)(e) \text{ iff } Mod(\varphi)(M') \models \Sigma^{e}$$

holds for all $M' \in |Mod(\Sigma')|$ and $e \in Sen(\Sigma)$. Following the usual notational conventions, we sometimes let $\_\upharpoonright_\varphi$ denote the reduct functor $Mod(\varphi)$ and φ denote the sentence translation $Sen(\varphi)$. When $M = M'\upharpoonright_\varphi$, we say that M' is a *φ-expansion* of M and that M is the *φ-reduct* of M', and similarly for model morphisms.

For all the following concepts related to institutions that we recall in the following text, the reader is referred to [14] unless some other place is explicitly indicated.

Let Σ be a signature. Then,

- For each $E \subseteq Sen(\Sigma)$, let $E^* = \{M \in |Mod(\Sigma)| \mid M \models_\Sigma e \text{ for all } e \in E\}$.
- For each class $\mathcal{M}$ of Σ-models, let $\mathcal{M}^* = \{e \in Sen(\Sigma) \mid M \models e \text{ for all } M \in \mathcal{M}\}$.

With no danger of confusion, we let $\bullet$ denote any of the two compositions $**$ of the two operators $*$. Each of the two bullets is a closure operator. When E and E' are sets of sentences of the same signature Σ, we let $E \models \Sigma E'$ denote the fact that $E^* \subseteq E'^*$. The relation $\models_\Sigma$ between sets of sentences is called the *(Σ-)semantic consequence relation* (notice that it is written just like the satisfaction relation). If $E' = \{e'\}$, we might write $E \models_\Sigma e'$. In order to simplify notation, we usually write $\models$ instead of $\models_\Sigma$ for both the satisfaction relation and the semantic consequence relation. Two sentences e and e' are called *equivalent*, denoted $e \equiv e'$, if $\{e\}^* = \{e'\}^*$. Dually, two models M and M' are called *elementary equivalent*, denoted $M \equiv M'$, if $\{M\}^* = \{M'\}^*$.

An institution is called *semi-exact* [22] if the model functor $Mod: Sign \to Cat^{op}$ preserves pushouts. Semi-exactness implies the following *amalgamation property* for any pushout of signature morphisms $(\Sigma_2 \xleftarrow{\varphi_2} \Sigma \xrightarrow{\varphi_1} \Sigma_1, \Sigma_2 \xrightarrow{\varphi_1'} \Sigma' \xleftarrow{\varphi_2'} \Sigma_1)$: for any $M_1 \in |Mod(\Sigma_1)|$, $M_2 \in |Mod(\Sigma_2)|$ such that $M_1\upharpoonright_{\varphi_1} = M_2\upharpoonright_{\varphi_2}$, there exists a unique model $M' \in |Mod(\Sigma')|$ such that $M'\upharpoonright_{\varphi_2'} = M_1$ and $M'\upharpoonright_{\varphi_1'} = M_2$. An analogous property is implied about model morphisms. An institution is called *liberal on signature morphisms* if the functor $Mod(\varphi)$ has a left adjoint for each signature morphism φ.

A *presentation* is a pair (Σ, E), where $E \subseteq Sen(\Sigma)$. A *theory* is a presentation (Σ, E) with E closed, i.e. with $E^\bullet = E$. One usually calls 'presentation' or 'theory' only the set E, and not the whole pair (Σ, E). A *presentation morphism* $\varphi : (\Sigma, E) \to (\Sigma', E')$ is a signature morphism $\varphi : \Sigma \to \Sigma'$ such that $\varphi(E) \subseteq E'^\bullet$. A presentation morphism between theories is called theory morphism. For a presentation (Σ, E), we let $Mod(\Sigma, E)$ denote the category of all Σ-models A such that $A \models E$.

A sentence $\rho \in Sen(\Sigma)$ is called basic [7] if there exists a Σ-model M_ρ such that, for all Σ-models M, $M \models \rho$ iff there exists a morphism $M_\rho \to M$. If, in addition, M_ρ is a finitely presented object in $Mod(\Sigma)$, ρ is called *finitary basic* [10]. Basic sentences tend to be the starting building blocks for sentences in concrete institutions. For instance, in the institution of first-order predicate logic, FOPL (see the following text the examples of institutions), conjunctions of ground atoms are basic. In this article, we shall be interested in institutions whose sentences are accessible from basic sentences by means of several first-order constructs; hence, it suffices to identify *enough* basic sentences, like the conjunctions of ground atoms above, in order to ensure accessibility. However, the concept of basic sentence turns out to be quite comprehensive in concrete cases; for instance, existentially quantified atoms are also basic in FOPL. The attribute 'finitary' is usually equivalent, in concrete cases, to the property that the sentence has only a finite number of symbols. All basic sentences in FOPL are also finitary basic, because FOPL is a 'finitary' logic; this is not the case of the institution of infinitary first-order predicate logic, IFOPL, where basic sentences with an infinite number of symbols can be constructed by means of infinite conjunctions of atoms.

The sentences of an institution $\mathcal{I}$ can be naturally extended with first-order-like constructions [34]: if $\varphi : \Sigma \to \Sigma'$, $\rho, \delta \in Sen(\Sigma)$, $\rho' \in Sen(\Sigma')$, and $E \subseteq Sen(\Sigma)$, one can build the sentences $\neg\rho$, $\delta \wedge \rho$, $\delta \vee \rho$, $\bigwedge E$, $\bigvee E$, $(\forall\varphi)\rho'$, $(\exists\varphi)\rho'$ by means of *negation, conjunction, disjunction, arbitrary conjunction, arbitrary disjunction, universal and existential quantification (over signature morphisms)*, respectively, with the following semantics for each Σ-model M:

- $M \models \neg\rho$ iff $M \not\models \rho$;
- $M \models \delta \wedge \rho$ iff $M \models \delta$ and $M \models \rho$;
- $M \models \delta \vee \rho$ iff $M \models \delta$ or $M \models \rho$;
- $M \models \bigwedge E$ iff $M \models e$ for each $e \in E$;
- $M \models \bigvee E$ iff $M \models e$ for some $e \in E$;
- $M \models (\forall\varphi)\rho'$ iff $M' \models \rho'$ for each φ-expansion M' of M;
- $M \models (\exists\varphi)\rho'$ iff $M' \models \rho'$ for some φ-expansion M' of M.

It might be the case that the newly constructed sentences are equivalent to some existing sentences in $\mathcal{I}$. The notion of a class of sentences *closed under* either one of the aforementioned constructions (e.g. under conjunction, or under universal quantification over a morphism φ) should be clear. An institution is said to *admit negation* if the class of all its sentences is closed under negation.

2.2.1 Examples of institutions

(1) FOPL — the institution of (many-sorted) first-order predicate logic (with equality). The signatures are triplets (S, F, P), where S is the set of sorts, $F = \{F_{w,s}\}_{w \in S^*, s \in S}$ is the ($S^* \times S$ -indexed) set of operation symbols and $P = \{P_w\}_{w \in S^*}$ is the (S^*-indexed) set of relation symbols. If $w = \lambda$, an element of $F_{w,s}$ is called a constant symbol, or a *constant*. By a slight notational abuse, we let F and P also denote $\bigcup_{(w,s) \in S^* \times S} F_{w,s}$ and $\bigcup_{w \in S^*} P_w$, respectively. A signature morphism between (S, F, P) and (S', F', P') is a triplet $\varphi = (\varphi^{\mathrm{sort}}, \varphi^{\mathrm{op}}, \varphi^{\mathrm{rel}})$, where $\varphi^{\mathrm{sort}} : S \to S'$, $\varphi^{\mathrm{op}} : F \to F'$, $\varphi^{\mathrm{rel}} : P \to P'$ such that $\varphi^{\mathrm{op}}(F_{w,s}) \subseteq F'_{\varphi^{\mathrm{sort}}(w), \varphi^{\mathrm{sort}}(s)}$ and $\varphi^{\mathrm{rel}}(P_w) \subseteq P'_{\varphi^{\mathrm{sort}}(w)}$ for all $(w, s) \in S^* \times S$. When there is no danger of confusion, we may let φ denote each of φ^{sort}, φ^{rel} and φ^{op}. Given a signature $\Sigma = (S, F, P)$, a Σ-model A is a triplet $A = (\{A_s\}_{s \in S}, \{A_{w,s}(\sigma)\}_{(w,s) \in S^* \times S, \sigma \in F_{w,s}}, \{A_w(R)\}_{w \in S^*, R \in P_w})$ interpreting each sort s as a set A_s, each operation symbol $\sigma \in F_{w,s}$ as a function $A_{w,s}(\sigma) : A^w \to A_s$ (where A^w stands for $A_{s_1} \times \ldots \times A_{s_n}$ if $w = s_1 \ldots s_n$), and each relation symbol $R \in P_w$ as a relation $A_w(R) \subseteq A^w$. When there is no danger of confusion we may let A_σ and A_R denote $A_{w,s}(\sigma)$ and $A_w(R)$, respectively. Morphisms between models are the usual Σ-homomorphisms, i.e. S-sorted functions that preserve the structure. The Σ-sentences are obtained from atoms, i.e. equality atoms $t_1 = t_2$, where $t_1, t_2 \in (T_F)_s$,[2] or relational atoms $R(t_1, \ldots, t_n)$, where $R \in P_{s_1 \ldots s_n}$ and $t_i \in (T_F)_{s_i}$ for each $i \in \{1, \ldots, n\}$, by applying for a finite number of times:

- negation, conjunction, and disjunction;
- universal or existential quantification over finite sets of constants (variables).

Satisfaction is the usual first-order satisfaction and is defined using the natural interpretations of ground terms t as elements A_t in models A. The definitions of functors *Sen* and *Mod* on morphisms are the natural ones: for any signature morphism $\varphi : \Sigma \to \Sigma'$, $Sen(\varphi) : Sen(\Sigma) \to Sen(\Sigma')$ translates sentences symbolwise, and $Mod(\varphi) : Mod(\Sigma') \to Mod(\Sigma)$ is the forgetful functor.

(2) $(\Pi \cup \Sigma)_n^0$ — the fragment of FOPL containing only sentences that are equivalent to sentences in prenex normal form that have at most n alternated blocks of quantifiers (universal and existential). Within a given signature, the mentioned set of sentences actually puts together two well-known types of first-order sentences: Π_n^0 and Σ_n^0 [4].

(3) PFOPL [6] — the institution of partial first-order predicate logic, an extension of FOPL whose signatures $\Sigma = (S, F, F', P)$ contain, besides relation and (total) operation symbols (in F and P), partial operation symbols too in F'. Models of course interpret the symbols in F' as partial operations of appropriate ranks. Σ-model morphisms $h : A \to B$ are S-sorted functions that commute with the total operations and relations in the usual way, and with the partial operations $\sigma \in F'_{s_1 \ldots s_n, s}$ in the following way: for each $(a_1, \ldots, a_n) \in A_{s_1 \ldots s_n}$, if $A_\sigma(a_1, \ldots, a_n)$ is defined, then so is $B_\sigma(h_{s_1}(a_1), \ldots, h_{s_n}(a_n))$,

[2] T_F is the ground term algebra over F.

and in this case the latter is equal to $h_s(A_\sigma(a_1, \ldots, a_n))$. A signature morphism between (S, F, F', P) and (S_0, F_0, F'_0, P_0) is a FOPL-signature morphism $\varphi : (S, F \cup F', P') \to (S_0, F_0 \cup F'_0, P'_0)$ such that, for each $\sigma \in F$, it holds that $\varphi^{\mathrm{op}}(\sigma) \in F_0$. Thus, signature morphisms are allowed to map partial operation symbols to total operation symbols, but not vice versa. There exist two kinds of atoms: (*existential*) *equality atoms* $t = t'$ and relational atoms $R(t_1, \ldots, t_n)$, having syntax just like at FOPL. An equality atom $t = t'$ holds in a model A when both terms are defined and equal ($A_t = A_{t'}$). A relational atom $R(t_1, \ldots, t_n)$ holds when all terms t_i are defined and their interpretations A_{ti} stay in relation A_R. The sentences are obtained from atoms just like in the case of FOPL (quantification over variables is allowed in the usual sense, which corresponds to considering the quantified variables as new total constants). Note that other kinds of sentences usually considered in partial algebraic frameworks can be expressed here: *definedness*, $t\downarrow$, as $t = t$; *strong equality*, $t =_s t'$ (either both t and t' are undefined, or both are defined and equal), as $(\neg\, t\downarrow \wedge \neg t'\downarrow) \vee\ t = t'$; *weak equality*, $t =_w t'$ (if both t and t' are defined, then they are equal), as $(\neg\, t\downarrow) \vee (\neg\, t'\downarrow) \vee\ t = t'$. The functor *Mod* is defined similarly to the case of FOPL.

(4) PA — the institution of partial algebra, a fragment of PFOPL having signatures without relation symbols. Partial algebras and their applications were extensively studied in [28] and [3].

(5) IFOPL — the institution of infinitary first-order logic, an infinitary extension of FOPL, which allows conjunction on arbitrary sets of sentences. This logical system is known under the name $L_{\infty,\omega}$ [21, 20] and plays an important role in categorical logic.

(6) IFOPL$_\alpha$ (where α is an infinite cardinal) — a fragment of IFOPL, admitting only conjunction on sets of sentences with cardinal smaller than α. This logical system is usually called $L_{\alpha,\omega}$ [18]. Note that IFOPL$_\omega$ is FOPL.

(7) PosFOPL — the institution of positive first order predicate logic, a fragment of FOPL, with sentences constructed by means of $\wedge$, $\vee$, $\forall$, $\exists$, but not negation $\neg$. Here $\vee$ and $\exists$ are no longer reducible to $\wedge$ and $\forall$ or vice versa. Positive sentences are defined and studied for example in [4, 25].

(8) EQL — the institution of equational logic [14], a fragment of FOPL, with no relation (only operation) symbols, and with sentences constructed from atoms only by means of universal quantification (no logical connectives).

(9) EQLN — a minimal extension of EQL with negation, allowing sentences obtained from atoms and negations of atoms through only one round of quantification, either universal or existential, over a set of variables. More precisely, all sentences have the form $(Q\,X)t_1\, k\, t_2$, where $Q \in \{\forall, \exists\}$ and $k \in \{=, \neq\}$. Note that this institution admits negation.

(10) RWL — the institution of (unconditional) rewriting logic. It has the same signatures as EQL, but models have in addition a preorder relation on each sort carrier, compatible with the operations, and model morphisms have to be increasing with respect to these preorders. The sentences are usual equations as in EQL and transitions $(\forall X)t \to t'$, with $\to$ interpreted as the model preorder. This logic cannot be seen as a fragment of FOPL, due to the built-in nature of the preorder on models. Rewriting logic was introduced in [23] with models having a categorical structure where arrows express different transitions between states; a simplified and more amenable version of this logic, which forces this categorical structure to be a preorder, is used in specification languages such as CafeObj [11] or Maude [5]; this simplified version was considered here.

(11) OSL — the institution of order-sorted (equational) logic [29], an extension of EQL where each signature has a partial order on the set of sorts. Thus a signature is a triplet $(S, \leq, F)$, where $(S, \leq)$ is a partially ordered set and (S, F) is an EQL-signature. A $(S, \leq, F)$-model is an (S, F)-model in EQL subject to two additional requirements:
   - For each $s, s' \in S$ with $s \leq s'$, it holds that $A_s \subseteq A_{s'}$;
   - For each $(w, s), (w', s') \in S^* \times S$ such that $w = s_1 \ldots s_n$, $w' = s'_1 \ldots s'_n$, $s_i \leq s'_i$, and $s \leq s'$, and each $\sigma \in F_{w,s} \cap F_{w',s'}$, it holds that $A_{w,s}(\sigma) : A^w \to A_s$ restricts and corestricts $A_{w',s'}(\sigma) : A^{w'} \to A_{s'}$.

   A $(S, \leq, F)$-morphism between A and B is a (S, F)-morphism in EQL, $h : A \to B$, such that for all $s, s' \in S$ with $s \leq s'$, it holds that h_s restricts and corestricts $h_{s'}$. Given a signature $\Sigma = (S, \leq, F)$, one can construct the ground term Σ-algebra T_Σ similarly to the case of EQL, just that one needs to consider the subsort relationship $\leq$ too. An $(S, \leq, F)$-sentence is an equation $(\forall X)t = t'$, where X is an S-sorted set of variables and $t, t' \in T_{(S, \leq, F \cup X)}$.[3] Satisfaction of a sentence by a model is defined in the obvious way. The functor *Mod* acts just like in the case of EQL.

(12) ML — the institution of (unconditional) membership equational logic, an extension of EQL, which calls the usual sorts 'kinds', and allows on each kind a set of sorts that are to be interpreted, on models, as subsets of the kind carrier. Thus a signature is a triplet $(K, \{S_k\}_{k \in K}, F)$, where (K, F) is an EQL-signature and for each kind $k \in K$, S_k is the set of sorts for this kind. Besides equations, this logic also has membership assertions: $(\forall X)t : s$, where $t \in (T_\Sigma(X))_k$ and $s \in S_k$, meaning that 't is of sort s'. This logic, introduced in [24], can be seen as a fragment of FOPL, which only uses unary relation symbols and has only universally quantified atoms as sentences. As shown in [24], ML naturally embeds (a variation of) OSL.

(13) EHL — the institution of extended hidden logic. The signatures are triplets (H, V, F), where:
   - H is the set of *hidden sorts*;
   - V is the set of *visible* sorts, $V \cap H = \emptyset$;
   - $(H \cup V, F)$ is an EQL-signature (i.e. F is an $(H \cup V)^* \times (H \cup V)$-indexed set of operation symbols).

   The (H, V, F)-models are the usual $(H \cup V, F)$-models from EQL. For a model A, one defines its behavioral equivalence $\equiv_A$ to be the least congruence on A, which is an identity on visible sorts. The (H, V, F)-morphsism are the $(H \cup V, F)$-morphisms from EQL that preserve behavioral equivalence. There are two kinds of atoms: (usual) *equality atoms* $t = t'$ and *behavioral equality atoms* $t \equiv t'$. Satisfaction of equality atoms is the usual first-order satisfaction. For a (H, V, F)-model A, $A \models t \equiv t'$ iff $A_{c[x_1 \leftarrow a_1, \ldots, x_n \leftarrow a_n, z \leftarrow t]} = A_{c[x_1 \leftarrow a_1, \ldots, x_n \leftarrow a_n, z \leftarrow t']}$ for each sort $v \in V$, each sequence $x_1, \ldots, x_n$ of variables of various sorts, each context c in $T_F(\{x_1, \ldots, x_n, z\})$ of sort v, and each sequence $a_1, \ldots, a_n$ of elements in A of appropriate sorts. (Here a context is a term with only one occurrence of the variable z; z is assumed to have the same sort as t and t'. Also, for instance $c[x_1 \leftarrow a_1, \ldots, x_n \leftarrow a_n, z \leftarrow t]$ is a ground $(H \cup V, F)$-term parameterized by elements in A, and $A_{c[x_1 \leftarrow a_1, \ldots, x_n \leftarrow a_n, z \leftarrow t]}$ is its natural interpretation as an element in A.) All sentences are constructed from atoms by means of first-order

[3] The variables in X are interpreted as new constants.

connectives and quantifiers just like in the case of FOPL (quantification is allowed over variables of both hidden and visible sorts). A signature morphism between (H, V, F) and (H', V', F') is an EQL-signature morphism $\varphi : (H \cup V, F) \to (H' \cup V', F')$ such that:

- $\varphi(V) \subseteq V'$, $\varphi(H) \subseteq H'$;
- For each $(w', s') \in (H' \cup V')^* \times (H' \cup V')$ such that w' contains a sort in $\varphi^{\mathrm{sort}}(H)$, and $\sigma' \in F'_{w', s'}$, there exists $\sigma \in F$ such that $\varphi^{\mathrm{op}}(\sigma) = \sigma'$.

On signature morphisms, the functors *Mod* and *Sen* act as in the case of EQL. The preceding description of EHL was adapted from [8]. See [15] for details about hidden logic, and [2] for the description of (a variation of) full first-order hidden logic.

(14) HL — [16, 12] the institution of hidden logic, a fragment of EHL, with sentences constructed from atoms only by means of universal quantification (no logical connectives).

3 Representable and quasi-representable signature morphisms

The institutional notions of *representable* and *quasi-representable* signature morphisms are abstract concepts meant to capture the phenomena of quantification over (sets of) first-order variables. Both notions start from the fact that semantics of quantification in first-order-like logics can be given in terms of signature extensions: $M \models_{(S,F,P)} (\forall X)e$ ($M \models_{(S,F,P)} (\exists X)e$) iff $M' \models_{(S,F \cup X,P)} e$ for each (for some) $(S, F \cup X, P)$-expansion M' of M. Thus, in order to reach first-order quantification institutionally, one needs to define somehow what 'injective signature morphism that only adds constant symbols' (such as $\iota : (S, F, P) \to (S, F \cup X, P)$) means.

Definition 2
A signature morphism $\varphi : \Sigma \to \Sigma'$ is called:

- Representable [7], if there exists a Σ-model M_φ (called the representation of φ) and an isomorphism of categories $I_\varphi : Mod(\Sigma') \to M_\varphi / Mod(\Sigma)$ such that $I_\varphi ; U = Mod(\varphi)$, where $U : M_\varphi / Mod(\Sigma) \to Mod(\Sigma)$ is the usual forgetful functor;
- Finitely representable [7], if it is representable and M_φ is a finitely presented object in $Mod(\Sigma)$;
- Quasi-representable [6], if for each $A' \in |Mod(\Sigma')|$, the canonical functor $A'/Mod(\varphi) : A'/Mod(\Sigma') \to A'\!\upharpoonright_\varphi / Mod(\Sigma)$ is an isomorphism of categories; and
- Finitely quasi-representable [6], if it is quasi-representable and for each colimit $(A_i \xrightarrow{\mu_i} A)_{i \in I}$ of a directed diagram of Σ-models $(A_i \xrightarrow{h_{i,j}} A_j)_{i,j \in I, i \leq j}$ and each φ-expansion A' of A, there exists an index $i \in I$ and a φ-expansion μ_i' of μ_i.

The notion of representability is built on the intuition that, in FOPL, an expansion of a $\Sigma = (S, F, P)$-model A over a signature inclusion $\iota : \Sigma \to \Sigma' = (S, F \cup X, P)$ that only adds constants can be viewed as a pair (M, v), where $v : X \to M$ is a function interpreting the new constants in X, and furthermore as a pair $(M, \overline{v})$, where $\overline{v} : T_\Sigma(X) \to M$ is a model-morphism.[4] Hence ι is represented by $T_\Sigma(X)$. And ι is finitely representable, i.e. $T_\Sigma(X)$ is finitely presented in $Mod(\Sigma)$, if X is finite.

[4] $T_\Sigma(X)$ is the term algebra over variables X and operations in F, with all relations in P empty.

On the other hand, quasi-representability follows the intuition that the aforementioned signature inclusion ι does not allow multiple expansions of Σ-morphisms $A \xrightarrow{h} B$ having a fixed source A' (where A' is a ι-expansion of A). This is because A' already 'indicates', via h, how the constants in X should be interpreted in the target model B' of a presumptive ι-expansion h' of h; and of course h' has to be identical, as a function, to h. Thus, ι is also quasi-representable. And again, ι is finitely quasi-represented if X is finite. Intuitively, if we regard directed colimits of Σ-models as 'unions', all the interpretations in the 'union model' A of the finite number of constants in X will eventually be reached by one of the members A_i of the union; hence the 'inclusion' of A_i into A has a ι-expansion.

For most concrete institutions (at least for those admitting initial objects in the categories of models, like our examples 1–14), the notions of representability and quasi-representability coincide, as shown by the following lemma.

Lemma 3 [6]
Let $\varphi\colon\Sigma\rightarrow\Sigma'$ be a signature morphism. Then [φ is (finitely) quasi-representable and $Mod(\Sigma)$ has an initial object] iff φ is (finitely) representable.

It is shown in [6] that quasi-representable signature morphisms create directed colimits. Throughout the article, we are going to use intensively a similar property:

Lemma 4
Let $\varphi : \Sigma \rightarrow \Sigma'$ be a quasi-representable signature morphism and let $(A_i \xrightarrow{f_{i,j}} A_j)_{i,j\in I, i\leq j}$ be a directed diagram in $Mod(\Sigma)$ with colimit $(A_i \xrightarrow{\mu_i} A)_{i\in I}$. Also, let $k \in I$, $I_k = \{i \in I \mid k \leq i\}$, and B a φ-expansion of A_k. Then there exists a directed diagram $(A'_i \xrightarrow{f'_{i,j}} A'_j)_{i,j\in I_k, i\leq j}$ in $Mod(\Sigma')$, with colimit $(A'_i \xrightarrow{\mu'_i} A')_{i\in I_k}$, such that:

- $A'_k = B$;
- For each $i,j \in I_k$ with $i \leq j$, $f'_{i,j}$ is a φ-expansion of $f_{i,j}$;
- For each $i \in I_k$, μ'_i is a φ-expansion of μ_{i}.

Proof. For each $i \in I_k$, we define $(f'_{k,i} : B \rightarrow A'_i, A'_i)$ to be $(A_k/Mod(\varphi))^{-1}(f_{k,i} : A_k \rightarrow A_i, A_i)$. In particular, for $i = k$, we have $A'_k = B$ and $f'_{k,k} = 1_B$. Also, for each $i,j \in I_k$ with $k < i \leq j$, we define $f'_{i,j}$ to be $(A_k/Mod(\varphi))^{-1}(f_{i,j} : (f_{k,i}, A_i) \rightarrow (f_{k,j}, A_j))$. Then $Dg' = (A'_i \xrightarrow{f'_{i,j}} A'_j)_{i,j\in I_k, i\leq j}$ is a directed diagram in $Mod(\Sigma')$ and its φ-reduct is $Dg = (A_i \xrightarrow{f_{i,j}} A_j)_{i,j\in I_k, i\leq j}$. Now we define $(\mu'_k : B \rightarrow A', A')$ to be $(A_k/Mod(\varphi))^{-1}(\mu_k : A_k \rightarrow A, A)$ and for each $i \in I_k$, μ'_i to be $(A_k/Mod(\varphi))^{-1}(\mu_i : (f_{k,i}, A_i) \rightarrow (\mu_k, A))$. Then $CC' = (A'_i \xrightarrow{\mu'_i} A')_{i\in I_k}$ is a cocone of Dg', and the φ-reduct of CC' is $CC = (A_i \xrightarrow{\mu_i} A)_{i\in I_k}$. The fact that CC' is an actual colimit for Dg' follows at once by the quasi-representability of φ: for any cocone $(A'_i \xrightarrow{\nu'_i} A'')_{i\in I_k}$ of Dg', we get that its reduct $(A'_i\upharpoonright_\varphi \nu'_i\upharpoonright_\varphi \rightarrow A''\upharpoonright_\varphi)_{i\in I_k}$ is a cocone of Dg; thus if one takes u to be the universal arrow from CC to $(A'_i\upharpoonright_\varphi \nu'_i\upharpoonright_\varphi \rightarrow A''\upharpoonright_\varphi)_{i\in I_k}$ (according to Lemma 2), then $(A_k/Mod(\varphi))^{-1}(u : (\mu_k, A) \rightarrow (\nu'_k\upharpoonright_\varphi, A''\upharpoonright_\varphi))$ is the desired universal arrow from CC' to $(A'_i\nu'_i \rightarrow A'')_{i\in I_k}$. ∎

The next lemma shows that (quasi-)representable signature morphisms behave well under composition and pushouts.

Lemma 5 [6]

(1) (Finitely) quasi-representable signature morphisms form a subcategory of *Sign*.
(2) (Finitely) representable signature morphisms form a subcategory of *Sign*.

(3) If the institution is semi-exact, then the class of (finitely) quasi-representable signature morphisms is closed under pushouts.
(4) If the institution is semi-exact and liberal on signature morphisms, then the class of (finitely) representable signature morphisms is closed under pushouts.

Of course, representability and quasi-representability are only abstract approximations for 'injective morphisms that only add constants'. What will be relevant for the results of this paper is that in all our examples 1–14 of institutions, (quasi-)representable signature morphism include the desired types of morphisms. Formally, let $\varphi : \Sigma \to \Sigma'$ be a signature morphism in any of the examples 1–14 of institutions. We say that φ is an injective signature morphism that only adds constants if the following conditions hold: φ^{sort} is bijective, φ^{rel} is bijective, φ^{op} is injective and $F' - \varphi^{\mathrm{op}}(F)$ contains only (total) operation symbols. (Here, F and F' stand for the sets of all (partial and total) operation symbols of Σ and Σ', respectively.) If in addition $F' - \varphi^{\mathrm{op}}(F)$ is finite, we say that φ is an *injective signature morphism that only adds finitely many constants*.

Proposition 6
In any of the examples 1–14 of institutions, all injective signature morphisms that only add (finitely many) constants are (finitely) representable, hence also (finitely) quasi-representable. Moreover, in each case, the (broad) subcategory of *Sign* of such morphisms is closed under pushouts.

Proof. The fact that such morphisms are (finitely) representable can be shown using arguments very similar to the ones for FOPL. The only slightly more exotic cases are PFOPL, PA, OSL, RL, HL, and EHL; however, in each case the algebra freely generated by a set of (total) variables exists in any signature. The cases of HL and EHL actually require a small separate discussion. Let (H, V, F) be a signature in either of the two institutions. Since each free algebra $T_F(X)$ has its behavioral equivalence equal to the identity, every EQL-morphism with source $T_F(X)$ is also an EHL- and HL-morphism, thus an EHL- or HL-signature morphism that only adds (finitely many) constants X is indeed (finitely) represented by $T_F(X)$.

As for closure under pushouts, this follows easily from the fact that, in the category of sets and functions, the subcategory of injective functions is closed under pushouts. ∎

Although not strictly needed in this article, but helpful for getting an idea on how close the aforementioned approximation is, we recall a concrete characterization of representable (and quasi-representable) signature morphisms in FOPL.

Proposition 7
[33] Let $\varphi : \Sigma = (S, F, P) \to \Sigma' = (S', F', P')$ be a signature morphism in FOPL. Then the following are equivalent:

(1) φ is representable;
(2) φ is quasi-representable and
(3) φ^{sort} and φ^{rel} are bijective, and for all $(w', s') \in S'^{*} \times S'$ with $w' \neq \lambda$, for all $\sigma' \in F'_{w',s'}$, there exists a unique $\sigma \in F$ such that $\varphi^{\mathrm{op}}(\sigma) = \sigma'$ (in other words, φ is bijective with respect to all items except constant symbols).

Moreover, the preceding three conditions stay equivalent if we add the word 'finitely' to the first two and add the requirement that $F' - \varphi^{\mathrm{op}}(F)$ be finite to the third.

4 Elementary morphisms

In classical first-order logic [4], an injective model morphism $A \xrightarrow{h} B$ is called an *elementary embedding* if one of the following equivalent conditions holds:

(1) For each formula $e(x_1, \ldots, x_n)$ and each sequence $a_1, \ldots, a_n \in A$, $A \models e(a_1, \ldots, a_n)$ iff $B \models e(h(a_1), \ldots, h(a_n))$;
(2) For each formula $e(x_1, \ldots, x_n)$ and each sequence $a_1, \ldots, a_n \in A$, $A \models e(a_1, \ldots, a_n)$ implies $B \models e(h(a_1), \ldots, h(a_n))$.

The notion of elementary embedding has an immediate generalization to the many-sorted case *FOPL*; and it was extended to cope with infinitary first-order logics too [18, 17]. Our next institutional generalization reads the concept of elementary embedding in the following way: the morphism h preserves sentences in any language extending with constants the original language, regardless of the actual interpretation of these constants. Notice that the two alternative definitions of elementary embeddings listed here are equivalent thanks to the existence of negations in full first-order logic; however, this is not the case in less expressive logics, such as *PosFOPL* or EQL. We prefer to consider the second variant and interpret elementarity as a *sentence preservation* property rather than a refinement of elementary equivalence. This subjective choice is motivated by our belief that taking into consideration the direction of the arrow h in the definition is a more fruitful approach. In an institution, the 'languages extended with constants' are captured by quasi-representable signature morphisms φ having as source the given language/signature, and satisfaction inside such an extended language is captured by usual satisfaction by φ-expansions. To keep the discussion general and to avoid certain intricacies in the particular cases resulting from considering all quasi-representable, or representable, signature morphisms, we parameterize our definition by a class $\mathcal{Q}$ of quasi-representable signature morphisms. Thus let $A \xrightarrow{h} B$ be a Σ-morphism. Formulae $e(x_1, \ldots, x_n)$ are expressed by usual sentences e' in signatures Σ', where $\varphi : \Sigma \to \Sigma'$ is a signature morphism in $\mathcal{Q}$. Satisfaction of such sentences $e' \in Sen(\Sigma')$ makes sense in φ-expansions of A and B, expansions which are to be seen as models together with some designated constants.[5] However, asking that '$A' \models e'$ implies $B' \models e'$' for all φ-expansions A' and B' of A and B and for all $e' \in Sen(\Sigma')$ is not appropriate, since the quoted implication should be required only about constants in A and B connected through h; the connection is realised by first considering φ-expansions h' of h.

The injectivity assumption in the definition of elementary embedding for classical first-order logic is superfluous. We did not consider it in the preceding discussion; this could be seen as yet another subjective choice, meant to emphasize once more the idea of sentence preservation, this time to the prejudice of the algebraic property of model embedding. This choice has an important terminological consequence: we define and study 'elementary morphisms', and not 'elementary embeddings', although the elementary morphisms yield in particular the FOPL-elementary embeddings.

For the whole section, we fix an institution $\mathcal{I}$ and a broad subcategory $\mathcal{Q}$ of *Sign* (i.e. a class of signature morphisms containing all identity morphisms and closed under composition), consisting of quasi-representable signature morphisms. In particular, by taking further

[5] Recall from Section 3 the connection between quasi-representable signature morphisms and first-order variables/constants.

mild assumptions on $\mathcal{I}$ such as semi-exactness or liberality on signature morphisms, according to Lemma 5, possible choices for $\mathcal{Q}$ are given by either of the following four types of signature morphisms: quasi-representable, finitely quasi-representable, representable, finitely representable.

Definition 8
Let Σ be a signature. A Σ-morphism $A \xrightarrow{h} B$ is called $\mathcal{Q}$-elementary if for all signature morphisms $\varphi : \Sigma \to \Sigma'$ in $\mathcal{Q}$, φ-expansions $A' \xrightarrow{h} B'$ of h, and sentences $e' \in \mathrm{Sen}(\Sigma')$, it holds that $A' \models e'$ implies $B' \models e'$.

Remark 9

(1) Because each morphism in $\mathcal{Q}$ is quasi-representable, in Definition 8 the φ-expansion h' of h is uniquely determined by the φ-expansion A' of A.
(2) If the institution admits negations, then the condition in Definition 8 can be equivalently stated by replacing '$A' \models e'$ implies $B' \models e'$' with '$A' \models e'$ iff $B' \models e'$'.

Let us see what our general concept of $\mathcal{Q}$-elementary morphism becomes for our examples 1–14 of institutions. In what follows, for all these institutions, we shall simply call *elementary morphisms* the $\mathcal{Q}$-elementary morphisms with $\mathcal{Q}$ being the category of injective signature morphisms that only add constants. (Note that in logics with finite sentences, such as FOPL and all its fragments, elementarity with respect to arbitrary signature morphisms that only add constants is equivalent to elementarity with respect to morphisms that only add *finitely many* constants; this is because just a finite set of the newly added constants are contained in a given sentence.) Known and relatively well-studied cases are the following:

- For FOPL, the elementary embeddings from (the many-sorted version of) classical model theory [4];
- For PA, the elementary embeddings of partial algebras [3];
- For IFOPL and IFOPL_α, the L_{∞,w^-} and L_{α,w^-} elementary embeddings from infinitary model theory [18, 21, 17];
- For $(\Pi \cup \Sigma)^0_1$, the existentially closed embeddings [17] and
- For $(\Pi \cup \Sigma)^0_n$, the Σ^0_n-extensions [4].

Up to our knowledge, elementary embeddings for the other examples of institutions were not considered so far in the literature. However, such notions are meaningful instances of the logic-independent concept of elementary morphism that we propose here. In each case, an elementary morphism is one that preserves satisfaction of all sentences with elements of the source model as parameters. The next proposition gives some expected properties of $\mathcal{Q}$-elementary morphisms.

Proposition 10
Let $\chi : \Sigma \to \Sigma'$ be a signature morphism. Then the following hold;

- The $\mathcal{Q}$-elementary morphisms in $Mod(\Sigma)$ form a subcategory of $Mod(\Sigma)$;
- Assume that the institution has pushouts of signatures and is semi-exact and that $\mathcal{Q}$ is closed under pushouts. If $A' \xrightarrow{h'} B'$ is a $\mathcal{Q}$-elementary morphism in $Mod(\Sigma')$, then $h' \upharpoonright_\chi$ is also $\mathcal{Q}$-elementary;
- If χ is in $\mathcal{Q}$, $A \xrightarrow{h} B$ is a $\mathcal{Q}$-elementary morphism in $Mod(\Sigma)$, and $A' \xrightarrow{h'} B'$ is a χ-expansion of h, then h' is also $\mathcal{Q}$-elementary.

Proof.

(1) If $A \in |Mod(\Sigma)|$, then 1_A is $\mathcal{Q}$-elementary because any expansion of 1_A along a quasi-representable signature morphism is an identity model-morphism itself. Let now $A \xrightarrow{h} B$ and BC be two $\mathcal{Q}$-elementary morphisms in $Mod(\Sigma)$ and let $f = h; g$. In order to show that f is $\mathcal{Q}$-elementary, let $\varphi : \Sigma \to \Sigma_0$ be a signature morphism in $\mathcal{Q}$ and $A_0 f_0 \to C_0$ a φ-expansion of f. Let $h_0 : A_0 \to B_0$ such that $(A_0, h_0) = (A_0/Mod(\varphi))^{-1}(A, h)$. Let $g_0 = (A_0/Mod(\varphi))^{-1}(g : (A, h) \to (A, f))$. Then $h_0; g_0 = f_0$. Since h and g are $\mathcal{Q}$-elementary, it follows that $\{A_0\}^* \subseteq \{B_0\}^* \subseteq \{C_0\}^*$.

(2) Let $A \xrightarrow{h} B$ denote the χ-reduct of h'. In order to prove h elementary, let $\varphi : \Sigma \to \Sigma_0$ be a signature morphism in $\mathcal{Q}$ and $A_0 \xrightarrow{h_0} B_0$ a φ-expansion of h. Consider the pushout $\Sigma' \varphi' \to \Sigma_1 \chi' \Sigma_0$ of the signature morphism span $(\Sigma' \chi \Sigma \varphi \to \Sigma_0)$. Then φ' is also in $\mathcal{Q}$. By semi-exactness, since h' and h_0 have a common reduct (that is, h), they also have a common expansion $A_1 \xrightarrow{h_1} B_1$ in $Mod(\Sigma_1)$. Because h' is elementary, $\{A_1\}^* \subseteq \{B_1\}^*$. Finally, using the satisfaction condition, we get $\{A_0\}^* \subseteq \{B_0\}^*$.

(3) Immediate from the definition of $\mathcal{Q}$-elementary morphisms and the fact that $\mathcal{Q}$ is closed under composition. ■

5 Elementary chain property

Throughout this section, we again fix an institution $\mathcal{I}$ and a broad subcategory $\mathcal{Q}$ of *Sign* consisting of quasi-representable signature morphisms.

A $\mathcal{Q}$-*elementary chain* is a chain diagram in $Mod(\Sigma)$ for some signature Σ, such that all its morphisms are $\mathcal{Q}$-elementary. The *elementary chain property* (parameterized by $\mathcal{Q}$ and abbreviated $\mathcal{Q}$-ECP) asks that, for each colimit of each $\mathcal{Q}$-elementary chain, all the structural morphisms be $\mathcal{Q}$-elementary. In other words, it asks that for each signature Σ, the subcategory of $Mod(\Sigma)$ of $\mathcal{Q}$-elementary morphisms be closed under chain colimits. We are going to prove that, under appropriate accessibility assumptions on sentences, $\mathcal{Q}$-ECP holds in an arbitrary institution. But first we need to consider some technical concepts and results.

We say that a sentence $e \in Sen(\Sigma)$ for some signature Σ is *preserved (reflected)* by directed colimits of $\mathcal{Q}$-*elementary morphisms, abbreviated* $\mathcal{Q}$-*preserved* (*reflected*), if for each directed diagram of $\mathcal{Q}$-elementary Σ-morphisms $(A_i \xrightarrow{f_{i,j}} A_j)_{i,j \in I, i \leq j}$ with colimit $(A_i \xrightarrow{\mu_i} A)_{i \in I}$ and each $k \in I$, $A_k \models e$ implies $A \models e$ ($A \models e$ implies $A_k \models e$ respectively).

Proposition 11

The class of sentences preserved by directed colimits of $\mathcal{Q}$-elementary morphisms

(1) Contains all basic sentences,
(2) Is closed under arbitrary conjunction and disjunction,
(3) Is closed under existential quantification over morphisms in $\mathcal{Q}$ and
(4) Is closed under universal quantification over finitely quasi-representable morphisms in $\mathcal{Q}$.

Proof. Let $e \in Sen(\Sigma)$, $(A_i \xrightarrow{f_{i,j}} A_j)_{i,j \in I, i \leq j}$ a directed diagram of $\mathcal{Q}$-elementary Σ-morphisms, with colimit $(A_i \xrightarrow{\mu_i} A)_{i \in I}$, and let $k \in I$. Assume that $A_k \models e$. We need to prove that $A \models e$.

(1) Assume e is a basic sentence. Since $A_k \models e$, there exists a Σ-morphism $M_e \to A_k$, hence, by composition with μ_k, we find a morphism $M_e \to A$, implying $A \models e$.

(2) Assume e is equivalent to $\bigwedge E$, where $E \subseteq Sen(\Sigma)$ such that for all $e' \in E$, e' is $\mathcal{Q}$-preserved. Since $A_k \models e$, it holds that $A_k \models e'$ for all $e' \in E$, hence $A \models e'$ for all $e' \in E$, hence $A \models e$. The proof for disjunction goes similarly.

(3) Assume e is equivalent to $(\exists\varphi)e'$, where $\varphi : \Sigma \to \Sigma'$ is in $\mathcal{Q}$ and e' is $\mathcal{Q}$-preserved. Then there exists a φ-expansion A'_k of A_k such that $A'_k \models e'$. By Lemma 4, there exists a directed diagram $(A'_i \xrightarrow{f'_{i,j}} A'_j)_{i,j\in I_k, i\leq j}$ in $Mod(\Sigma')$, with colimit $(A'_i \xrightarrow{\mu'_i} A')_{i\in I_k}$ such that, for each $i,j \in I_k$ with $i \leq j$, $f'_{i,j}$ is a φ-expansion of $f_{i,j}$ and μ'_i is a φ-expansion of μ_i. In particular, $A'\upharpoonright_\varphi = A$. According to Proposition 10.(3), each $f'_{i,j}$ is $\mathcal{Q}$-elementary. Applying the fact that e' is $\mathcal{Q}$-preserved, we obtain that $A' \models e'$, hence $A \models (\exists\varphi)e'$, i.e. $A \models e$.

(4) Assume e is equivalent to $(\forall\varphi)e'$, where $\varphi : \Sigma \to \Sigma'$ is a finitely quasi-representable signature morphism in $\mathcal{Q}$ and e' is $\mathcal{Q}$-preserved. Let A' be a φ-expansion of A. We need to show that $A' \models e'$. Because φ is finitely quasi-representable, there exists $q \in I$ and a φ-expansion $\xi' : A'_q \to A'$ of μ_q. Since $(I, \leq)$ is directed, there exists $p \in I$ such that $q \leq p$ and $k \leq p$. Thus, because $f_{k,p}$ is $\mathcal{Q}$-elementary, we get $A_p \models e$. Define $(A'_q \xrightarrow{f'} A'_p, A'_p)$ to be $(A_q/Mod(\varphi))^{-1}(A_q \xrightarrow{f_{q,p}} A_p, A_p)$ and ν' to be $(A_q/Mod(\varphi))^{-1}(\mu_p : (A_q, f_{q,p}) \to (A_q, \mu_q))$. Note that $\nu' : A'_p \to A'$. By Lemma 4 applied to the index p, there exists a directed diagram $(A'_i \xrightarrow{f'_{i,j}} A'_j)_{i,j\in I_p, i\leq j}$ in $Mod(\Sigma')$, with colimit $(A'_i \xrightarrow{\mu'_i} B')_{i\in I_p}$, such that, for each $i,j \in I_p$ with $i \leq j$, $f'_{i,j}$ is a φ-expansion of $f_{i,j}$ and μ'_i is a φ-expansion of μ_i. Again, Proposition 10.(3) assures us that each $f'_{i,j}$ is $\mathcal{Q}$-elementary. Since both (ν', A') and (μ'_p, B') are equal to $(A_p/Mod(\varphi))^{-1}(\mu_p, A)$, it follows that $A' = B'$ and $\nu' = \mu'_p$. Finally, since e' is $\mathcal{Q}$-preserved and $A'_p \models e'$ (A'_p being a φ-expansion of A_p), we obtain that $A' \models e'$.

Proposition 12

Assume that the institution admits negation. Then the class of sentences preserved and reflected by directed colimits of $\mathcal{Q}$-elementary morphisms

(1) Contains all finitary basic sentences,
(2) Is closed under arbitrary conjunction and disjunction, and under negation and
(3) Is closed under universal and existential quantification over finitely quasi-representable morphisms in $\mathcal{Q}$.

Proof. Let $e \in Sen(\Sigma)$, $(A_i \xrightarrow{f_{i,j}} A_j)_{i,j\in I, i\leq j}$ a directed diagram of $\mathcal{Q}$-elementary Σ-morphisms, with colimit $(A_i \xrightarrow{\mu_i} A)_{i\in I}$, and let $k \in I$. We need to prove $[A_k \models e$ iff $A \models e]$.

- Assume e is a finitary basic sentence. That $A_k \models e$ implies $A \models e$ follows from Proposition 11(1). Therefore let us suppose $A \models e$, i.e. there exists a Σ-morphism $g : M_e \to A$, and let us show that $A_k \models e$. Because M_e is finitely presentable, there exists $j \in I$ and Σ-morphism $h : M_e \to A_j$ such that $h; \mu_j = g$. Because $(I, \leq)$ is directed, there exists $i \in I$ such that $k \leq i$ and $j \leq i$. Then, by the existence of the morphism $M_e \xrightarrow{h;f_{j,i}} A_i$, it follows that $A_i \models e$. Moreover, since $f_{k,i} : A_k \to A_i$ is $\mathcal{Q}$-elementary and the institution admits negation, we obtain $A_k \models e$.
- Similar to the proof of Proposition 11(2) for conjunction and disjunction. For negation, the property is obvious thanks to its symmetry.
- Because the institution admits negation, universal and existential quantifications are mutually definable. Therefore, let us focus on existential quantification. Assume e is equivalent to $(\forall\varphi)e'$, where $\varphi : \Sigma \to \Sigma'$ is a finitely quasi-representable morphism in $\mathcal{Q}$ and e' is a Σ'-sentence $\mathcal{Q}$-[preserved and reflected]. That $A_k \models e$ implies $A \models e$ follows from Proposition 11(3). Let us now suppose $A \models (\exists\varphi)e'$ and let us show that $A_k \models (\exists\varphi)e'$. We have that $A \models \neg(\forall\varphi)\neg e'$, which means $A \not\models (\forall\varphi)\neg e'$. By point (2), $\neg e'$ is $\mathcal{Q}$-[preserved and reflected]. Thus, by Proposition 11.(4), $(\forall\varphi)\neg e'$ is $\mathcal{Q}$-preserved, hence $A_k \models (\forall\varphi)\neg e'$ would imply $A \models (\forall\varphi)\neg e'$, which is a contradiction. Thus $A_k \models \neg(\forall\varphi)\neg e'$, i.e. $A_k \models (\exists\varphi)e'$. ■

Proposition 13

Assume that all sentences of the institution are preserved by directed colimits of $\mathcal{Q}$-elementary morphisms. Then for each signature Σ, the subcategory of $Mod(\Sigma)$ of $\mathcal{Q}$-elementary morphisms is closed under directed colimits.

Proof. Let $(A_i \xrightarrow{f_{i,j}} A_j)_{i,j\in I, i\leq j}$ be a directed diagram in $Mod(\Sigma)$ such that each $f_{i,j}$ is $\mathcal{Q}$-elementary, and let $(A_i \xrightarrow{\mu_i} A)_{i\in I}$ be its colimit. Let $k \in I$. In order to prove that μ_k is $\mathcal{Q}$-elementary, let $\nu' : A'_k \to A'$ be a φ-expansion of μ_k and let $e' \in Sen(\Sigma')$ such that $A'_k \models e'$. By Lemma 4, there exists a directed diagram $(A'_i \xrightarrow{f'_{i,j}} A'_j)_{i,j\in I_k, i\leq j}$ in $Mod(\Sigma')$, with colimit $(A'_i \xrightarrow{\mu'_i} B')_{i\in I_k}$, such that, for each $i,j \in I_k$ with $i \leq j$, $f'_{i,j}$ is a φ-expansion of $f_{i,j}$ and μ'_i is a φ-expansion of μ_i. Just like in the proof of Proposition 11(4), one obtains that each $f'_{i,j}$ is $\mathcal{Q}$-elementary and that $A' = B'$ and $\nu' = \mu'_k$. Thus, according to our hypothesis, $A' \models e'$.

Definition 14

An institution $\mathcal{I}$ is called $\mathcal{Q}$-first-order-accessible if one of the two following properties holds:

(1) All sentences of $\mathcal{I}$ are (equivalent to ones) obtained from basic sentences by applying a finite number of times the following rules:
- Arbitrary conjunction;
- Arbitrary disjunction;
- Existential quantification over morphisms in $\mathcal{Q}$ and
- Universal quantification over finitely quasi-representable morphisms in $\mathcal{Q}$.

(2) I admits negation and all sentences of $\mathcal{I}$ are (equivalent to ones) obtained from finitary basic sentences by applying a finite number of times the following rules:
- Arbitrary conjunction;
- Arbitrary disjunction;
- Negation;
- Existential quantification over finitely quasi-representable morphisms in $\mathcal{Q}$ and
- Universal quantification over finitely quasi-representable morphisms in $\mathcal{Q}$.

Proposition 15

All the examples 1–14 of institutions are $\mathcal{Q}$-first-order-accessible, where $\mathcal{Q}$ is each time the category of injective signature morphisms that only add constants.

Proof. Let us first see that IFOPL and all its fragments are $\mathcal{Q}$-first-order accessible. Indeed, for each signature $\Sigma = (S, F, P)$, an equality atom $t = t'$ is finitary basic thanks to the model $T_\Sigma/t = t'$, that is, the (S, F)-algebra $T_F/t = t'$ (a quotient of the ground term algebra over F) with all the relations in P empty, while a relational atom $R(t_1, \ldots, t_n)$ is basic, thanks to the model consisting of T_F together with all relations in P empty, except R, which is the singleton $\{(t_1, \ldots, t_n)\}$. Moreover, quantification over finite or infinite sets of variables are particular cases of quantification over signature morphisms in $\mathcal{Q}$. Thus, PosFOPL, EQL and ML fall into case 1 of Definition 14, and FOPL, $(\Pi \cup \Sigma)^0_n$, IFOPL, IFOPLα and EQLN into case 2.

A similar argument as the preceding holds for PFOPL and PA too, since, for instance given a PA-signature and a set of equality atoms, there exists the initial algebra in the category of algebras satisfying these atoms [3]. And similarly for RWL and OSL.

As for EHL and HL, one has to notice mainly two things. First, all usual equality atoms are basic; indeed, the algebra $T_F/t = t'$ has the property that, for each (H, V, F)-model A satisfying the usual equality atom $t = t'$, the unique EQL-morphism between $T_F/t = t'$ and A preserves behavioral equivalence; hence, it is also an EHL- and HL-morphism.

Second, the behavioural equality atoms are equivalent to (infinite) conjunctions of universally quantified usual equality atoms; indeed, it holds that $A \models t \equiv t'$ iff $A \models \bigwedge\{(\forall\{x_1, \ldots, x_n\})\; c[z \leftarrow t] = c[z \leftarrow t'] \mid v \in V,\, x_1, \ldots, x_n$ variables, $c \in T_F(\{x_1, \ldots, x_n, z\})_v$context$\}$. Thus HL falls into case 1 of Definition 14. Moreover, EHL falls into case 2, since although behavioural equality atoms are not finitary, they are nevertheless obtainable from finitary basic sentences by means of the rules of universal quantification over finitely quasi-representable morphisms in $\mathcal{Q}$ (i.e. over finite sets of variables) and arbitrary conjunction.

Theorem 16

(Elementary Chain Theorem) Assume that the institution is $\mathcal{Q}$-first-order-accessible. Then for each signature Σ, the subcategory of *Mod(*Σ*)* of $\mathcal{Q}$-elementary morphisms is closed under directed colimits. In particular, the institution enjoys the $\mathcal{Q}$-*ECP*.

Proof. Follows immediately: for case 1 of Definition 14 from Propositions 11 and 13, and for case 2 from Propositions 12 and 13.

The separation on two cases in Definition 14 covers mainly the following situations: the institution $\mathcal{I}$ either admits negation, or has no negation—intermediate cases are not covered. Some important examples of institutions to which our Theorem 16 does not apply are all variations of Horn logic—in fact, for those institutions, we conjecture that the elementary chain property does not hold.

Corollary 17

All the examples 1–14 of institutions enjoy the $\mathcal{Q}$-ECP.

Note that Theorem 16 is applicable to a whole variety of other logics resulted from other different combinations of connectives and quantifiers. An interesting example which takes full advantage of Proposition 11 is a version of positive infinitary first-order logic admitting arbitrary conjunction and disjunction, existential quantification over arbitrary sets of variables, and universal quantification over finite sets of variables. Moreover, the case of *fragments* of languages (over transitive sets) in infinitary first-order logic [18] also seems to fall into our framework, provided that one takes the trouble of formalizing this as an institution.

6 Elementary morphisms by diagrams

An alternative definition of elementary embeddings in classical model theory is given in terms of elementary diagrams [4, 25]. There, the *elementary diagram* $EDg(A)$ of a model A is the set of all sentences in $\Sigma(A)$ (the language Σ of A extended with all elements of A as constants) that are true in A. Then, an embedding $A \xrightarrow{h} B$ is elementary iff $h(EDg(A)) \subseteq EDg(B)$, where $h(EDg(A))$ is the obvious translation through h of the sentences in $EDg(A)$. The main difference to the original definition (discussed at the beginning of Section 4) is that a language which includes parameter symbols for the source model, $\Sigma(A)$, is *a priori* given and the desired property is stated locally, in that fixed language. By adapting an existing institutional concept of diagram, we can discuss this alternative definition in a logic-independent framework.

6.1 Institutional diagrams

Diagrams are a basic concept in classical model theory [4]. They were first generalized to the institutional framework in [35, 36]; there, it is defined the concept of *abstract* *algebraic institution*, which is an institution subject to some additional natural requirements (like finite-exactness, existence of direct products of models, etc.) and enriched with a system of diagrams. The reason for introducing diagrams there was *making all algebras accessible*, for specification purposes. In this article, we need a more elaborated notion of institutional diagram, defined in [8], which takes into consideration not only models, but also model morphisms. The concept was introduced under the name 'elementary diagram'. For reasons that will be pointed out soon, here we prefer to use, like in [35], the name 'positive diagram' instead.

An institution $\mathcal{I} = (Sign, Sen, Mod, \models)$ is said to have positive diagrams [8] if

(1) For each signature Σ and Σ-model A there exists a signature morphism $\iota_\Sigma(A) : \Sigma \to \Sigma_A$ and a set E_A of Σ_A-sentences (called the *positive diagram* of A) such that $Mod(\Sigma_A, E_A)$ and $A/Mod(\Sigma)$ are isomorphic by an isomorphism $i_{\Sigma,A}$ making the following diagram commutative:

$$\begin{tikzcd} Mod(\Sigma_A, E_A) \arrow[rr, "i_{\Sigma,A}"] \arrow[dr, "-\upharpoonright_{\iota_\Sigma(A)}"'] & & A/Mod(\Sigma) \arrow[dl, "\text{forgetful}"] \\ & Mod(\Sigma) & \end{tikzcd}$$

(2) ι is 'functorial', i.e. for each signature morphism $\varphi : \Sigma \to \Sigma'$, each $A \in |Mod(\Sigma)|$, $A' \in |Mod(\Sigma')|$ and $h : A \to A'\upharpoonright_\varphi$ in $Mod(\Sigma)$, there exists a presentation morphism $\iota_\varphi(h) : (\Sigma_A, E_A) \to (\Sigma'_{A'}, E_{A'})$ making the following diagram commutative:

$$\begin{tikzcd} \Sigma \arrow[r, "\iota_\Sigma(A)"] \arrow[d, "\varphi"'] & \Sigma_A \arrow[d, "\iota_\varphi(h)"] \\ \Sigma' \arrow[r, "\iota_{\Sigma'}(A')"'] & \Sigma'_{A'} \end{tikzcd}$$

(3) i is natural, i.e. for each signature morphism $\varphi : \Sigma \to \Sigma'$, each $A \in |Mod(\Sigma)|$, $A' \in |Mod(\Sigma')|$ and $h : A \to A'\upharpoonright_\varphi$ in $Mod(\Sigma)$, the following diagram is commutative:

$$\begin{tikzcd} Mod(\Sigma_A, E_A) \arrow[r, "i_{\Sigma,A}"] & A/Mod(\Sigma) \\ Mod(\Sigma'_{A'}, E_{A'}) \arrow[r, "i_{\Sigma',A'}"] \arrow[u, "-\upharpoonright_{\iota_\varphi(h)}"] & A'/Mod(\Sigma') \arrow[u, "h/Mod(\varphi)"'] \end{tikzcd}$$

Here are some notational conventions that we hope will make the reader's life easier. Let $\varphi : \Sigma \to \Sigma'$ be a signature morphism, $A' \in |Mod(\Sigma')|$ and $h : A \to B$ in $Mod(\Sigma)$. We write $\iota_\Sigma(h)$ instead of $\iota_{1_\Sigma}(h)$ and $\iota_\varphi(A'\upharpoonright_\varphi)$ instead of $\iota_\varphi(1_{(A'\upharpoonright_\varphi)})$. Let A be a fixed object in $Mod(\Sigma)$ and let $B, C \in |Mod(\Sigma)|$ and $f : A \to B$, $g : A \to C$, $u : B \to C$ morphisms in $Mod(\Sigma)$ such that $f; u = g$. Then (f, B) and (g, C) are objects in $A/Mod(\Sigma)$ and u is also a morphism in $A/Mod(\Sigma)$ between (f, B) and (g, C). We establish the following notations: $B_f = i^{-1}_{\Sigma,A}(f, B)$ (and, similarly, $C_g = i^{-1}_{\Sigma,A}(g, C)$), $u_{f,g} = i^{-1}_{\Sigma,A}((f, B) \xrightarrow{u} (g, C))$. Thus, for instance, let $f : A \to B$ be a Σ-model morphism. Then $f_{1_A,f}$ is the image through $i^{-1}_{\Sigma,A}$ of the morphism $f : (1_A, A) \to (f, B)$ in $A/Mod(\Sigma)$, and has source $A_{(1_A)}$ and target B_f. We shall write A_A instead of $A_{(1_A)}$ and $f_{A,f}$ instead of $f_{1_A,f}$.

In classical model theory, Σ_A is the signature Σ enriched with all the elements of A as constants, $\iota_\Sigma(A) : \Sigma \to \Sigma_A$ is the inclusion of signatures, and E_A is a set of parameterized sentences which hold in A, depending on the considered type of morphism between models. If arbitrary model homomorphisms are allowed as morphisms, like in FOPL, one gets the 'positive diagram'; if just model embeddings are considered, one gets the 'diagram'; and only if just elementary embeddings are allowed, one gets what is classically called 'elementary diagram' (see [4] for the pointed standard terminology in classical model theory). Thus, the preceding institutional definition of diagrams particularises to elementary diagrams for classical first-order logic only if a notion of elementary morphism is assumed as *previously* defined. However, it is precisely the latter notion that we want to capture using diagrams. Therefore, we prefer to use the term 'positive diagram', in accordance to the particularization of the concept to the concrete institution FOPL, widely accepted as the institution of first-order logic. Thus we view the set of sentences E_A as the positive, rather than elementary, diagram of A, but of course keeping for it the same understanding as in [8]: that E_A axiomatizes the class of Σ-morphisms with source A. And we use the term elementary diagram of A for the set $\{A_A\}^*$, of all sentences satisfied by the self-parameterized extension A_A of A.

In [8], there are presented positive diagrams for FOPL, RWL, PA, and HL. Most institutions that were built starting from 'working' logical systems tend to have elementary diagrams. We next recall the system of positive diagrams for FOPL. Let $\Sigma = (S, F, P)$ be a FOPL-signature and $A \in |Mod(\Sigma)|$. Define $\Sigma_A = (S, F_A, P)$, where F_A extends F by adding, for each $s \in S$, all elements of A_s as constants of sort s. Further, we define:

(1) $A_A \in |Mod(\Sigma_A)|$, as the Σ_A-expansion of A which interprets each constant $a \in A$ by a;
(2) E_A, as the set of all atoms in $Sen(\Sigma_A)$ satisfied by A_A;
(3) $\iota_\Sigma(A)$, as the signature inclusion of Σ into Σ_A;
(4) The functor $i_{\Sigma,A} : Mod(\Sigma_A, E_A) \to A/Mod(\Sigma)$, as:

- $i_{\Sigma,A}(B') = (A \xrightarrow{h} B, B)$, where $B = B'\upharpoonright_{\iota_\Sigma(A)}$ and, for each $s \in S$ and $a \in A_s$, $h_s(a) = B'_a$.
- $i_{\Sigma,A}(f) = f$.

Let $\varphi : \Sigma = (S, F, P) \to \Sigma' = (S', F', P')$ be a signature morphism, $A \in |Mod(\Sigma)|$, $B \in |Mod(\Sigma')|$ and $h : A \to B\upharpoonright_\varphi$ in $Mod(\Sigma)$. Then the natural presentation morphism $\iota_\varphi(h) : (\Sigma_A, E_A) \to (\Sigma'_B, E_B)$ from the definition of positive diagrams is the following: if $e \in Sen(\Sigma_A)$, then $\iota_\varphi(h)(e)$ is obtained from e by symbolwise translation, mapping:

- Each $\sigma \in F$ into $\varphi^{\mathrm{op}}(\sigma)$,
- Each $R \in P$ into $\varphi^{\mathrm{rel}}(R)$,
- For all $s \in S$, each $a \in A_s$ into $h_s(a)$,
- For all $s \in S$, each variable $x{:}s$ of sort s into a variable $x{:}\varphi^{\mathrm{sort}}(s)$ and
- Each other symbol u that appears in e (e.g. logical connectives and quantifiers) into u.

As a general rule easily seen to hold about positive diagrams, one has that:

- If an institution $\mathcal{I}'$ extends an institution $\mathcal{I}$ and has the same category *Sign* and functor *Mod* (thus only adds new sentences), then positive diagrams are inherited by $\mathcal{I}'$ from $\mathcal{I}$;
- If an institution $\mathcal{I}'$ restricts an institution $\mathcal{I}$, has the same category *Sign* and functor *Mod* (thus only restricts the sets of sentences), but $\mathcal{I}'$ still has all the sentences in the positive diagrams E_A of $\mathcal{I}$, then positive diagrams are inherited by $\mathcal{I}'$ from $\mathcal{I}$.

Thus, the earlier described positive diagrams of FOPL are inherited by IFOPL, IFOLPLα, PosFOPL, EQL, EQLN, ML. The positive diagrams for the other mentioned institutions can be constructed with a similar pattern as those of FOPL; as remarked in [8], the sentences E_A are always the basic sentences satisfied by the model A expanded to Σ_A with constants in A pointing to themselves.

For what follows, we fix an institution with positive diagrams, $\mathcal{I}$.

DEFINITION 18
Let Σ be a signature.

- Given a Σ-model A, the elementary diagram of A is the set $\{A_A\}^*$ (of Σ_A-sentences satisfied by A_A).[6]
- A Σ-morphism $A \xrightarrow{h} B$ is called elementary by diagrams (d-elementary) if one of the following two equivalent conditions holds:
  - $A_A^* \subseteq B_h^*$;
  - $\iota_\Sigma(h)(A_A^*) \subseteq B_B^*$.

That the two conditions in point (2) of the preceding definition are equivalent follows from the satisfaction condition together with the fact that, by the naturality of i, $B_B{\upharpoonright}_{\iota_\Sigma(h)} = B_h$.

Thus, we defined elementary morphisms by means of elementary diagrams. We can spell out this definition as follows: $A \xrightarrow{h} B$ is d-elementary if the elementary diagram of A is embedded, via h, into the elementary diagram of B.

6.2 The relationship between $\mathcal{Q}$-elementary and d-elementary

The notion of d-elementary morphism is more compact than that of $\mathcal{Q}$-elementary morphism, but the former needs a lot of further structure on top of the plain institutional structure. We next provide conditions under which the two concepts are equivalent. For all this section, we fix a broad subcategory $\mathcal{Q}$ of *Sign* consisting of representable signature morphisms.[7]

PROPOSITON 19
The positive diagrams of $\mathcal{I}$ are said to be $\mathcal{Q}$-normal if for each representable signature morphism $\varphi : \Sigma \to \Sigma'$ (represented by M_φ) there exists a signature morphism $\chi : \Sigma' \to \Sigma_{M_\varphi}$ such that $\varphi; \chi = \iota_\Sigma(M_\varphi)$ and $Mod(\chi); I_\varphi = i_{\Sigma, M_\varphi}$.[8]

In examples 1–14 of institutions, for the usual choice of $\mathcal{Q}$, i.e. to consist of all injective signature morphisms that only add constants, the signature morphisms $\iota_\Sigma(A)$ of the positive diagrams are all in $\mathcal{Q}$. Moreover, in each case, the positive diagrams are also $\mathcal{Q}$-normal. Indeed, for example, in FOPL, given an injective signature morphism φ that only adds constants, which we can assume without loss of generality to be an inclusion $\Sigma = (S, F, P) \to \Sigma' = (S, F \cup X, P)$, represented by the Σ-model $T_\Sigma(X)$, the desired morphism χ such that $\varphi; \chi = \iota_\Sigma(T_\Sigma(X))$ is the signature inclusion $(S, F \cup X, P) \to (S, F \cup T_\Sigma(X), P)$ given by the set inclusion $X \to T_\Sigma(X)$. In order to see that the corresponding condition on model

[6] Since $A_A \models E_A$, the positive diagram of A is included in the elementary diagram of A.

[7] Note that we require more than usual for the subcategory $\mathcal{Q}$, namely representability instead of quasi-representability.

[8] Here, we made the slight notational abuse of letting $Mod(\chi)$ denote the restriction of $Mod(\chi) : Mod(\Sigma_{M_\phi}) \to Mod(\Sigma')$ to $Mod(\Sigma_{M_\varphi}, E_{M\varphi})$.

categories holds, let N be a $\Sigma_{T_\Sigma(X)}$-model that satisfies $E_{T_\Sigma(X)}$. Then $i_{\Sigma, T_\Sigma(X)}(N) = (h : T_\Sigma(X) \to N\upharpoonright_{\iota_\Sigma(T_\Sigma(X))}, N\upharpoonright_{\iota_\Sigma(T_\Sigma(X))})$, where $h(t) = N_t$ for all $t \in T_\Sigma(X)$. On the other hand, $I_\varphi(N\upharpoonright_\chi) = (g : T_\Sigma(X) \to N\upharpoonright_{\varphi;\chi}, N\upharpoonright_{\varphi;\chi}) = (g : T_\Sigma(X) \to N\upharpoonright_{\iota_\Sigma(T_\Sigma(X))}, N\upharpoonright_{\iota_\Sigma(T_\Sigma(X))})$, where g is the unique Σ-morphism extending the mapping $v : X \to N$ defined by $v(x) = N_x$ for all $x \in X$. Thus, by the freeness of $T_\Sigma(X)$, $g = h$. Hence, the functors $Mod(\chi); I_\varphi$ and i_{Σ, M_φ} coincide on objects. That they coincide on morphisms too follows at once from $\varphi; \chi = \iota_\Sigma(T_\Sigma(X))$. Normality of the positive diagrams for the other examples of institutions can be shown similarly to the case of FOPL.

Proposition 20
If the positive diagrams are normal and have each signature morphism $\iota_\Sigma(A)$ in $\mathcal{Q}$, then any model morphism is $\mathcal{Q}$-elementary iff it is d-elementary.

Proof. Let Σ be a signature and $A \xrightarrow{h} B$ a Σ-morphism.

Assume first that h is $\mathcal{Q}$-elementary. Then, since $\iota_\Sigma(A)$ is in $\mathcal{Q}$ and $A_A \xrightarrow{h_A} B_h$ is a $\iota_\Sigma(A)$-expansion of h, we get $\{A_A\}^* \subseteq \{B_h\}^*$. Thus, h is d-elementary.

Conversely, assume that h is d-elementary. Let $\varphi : \Sigma \to \Sigma'$ be a signature morphism in $\mathcal{Q}$ and $A' \xrightarrow{h} B'$ a φ-expansion of h. Let $(M_\varphi \xrightarrow{a} A, A) = I_\varphi(A')$ and $(M_\varphi \xrightarrow{b} B, B) = I_\varphi(B')$. By the naturality of i we have $h_{A,h}\upharpoonright_{\iota_\Sigma(a)} = h_{a,b}$.[9] Since $\{A_A\}^* \subseteq \{B_h\}^*$, by the satisfaction condition, it follows that $\{A_a\}^* \subseteq \{B_b\}^*$. Now, by the normality of diagrams, there exists $\chi : \Sigma' \to \Sigma_{M_\varphi}$ such that $\varphi; \chi = \iota_\Sigma(M_\varphi)$ and $Mod(\chi); I_\varphi = i_{\Sigma, M\varphi}$. Then $Mod(\chi) = i_{\Sigma, M\varphi}; I_\varphi^{-1}$, thus $h_{a,b}\upharpoonright_\chi = I_\varphi^{-1}(i_{\Sigma, M_\varphi}(h_{a,b})) = I_\varphi^{-1}(h) = h'$. Hence, $A_a\upharpoonright_\chi = A'$ and $B_b\upharpoonright_\chi = B'$. Finally, by the satisfaction condition, we get $\{A'\}^* \subseteq \{B'\}^*$.

Corollary 21
In all the examples 1–14 of institutions (with their mentioned diagrams), a model morphism is elementary iff it is d-elementary.

7 Concluding remarks

We outline the contributions of the present article:

- Introduced an abstract notion of elementary morphism, parameterized by a class of signature morphisms;
- Studied the connection between elementary morphisms and positive diagrams in an arbitrary institution, by giving an alternative diagrammatic definition of elementarity and
- Showed how the general results particularize to many concrete cases of logical systems, yielding different known results in a unitary fashion but also some new results; in particular, the less conventional cases of partial algebra, hidden logic and rewriting logic fall into our framework.

An open problem that we consider worthwhile is the institutional relationship between elementary morphisms and model embeddings. Given the fact that classically elementary morphisms are also embeddings, a result stating that, under certain assumptions on the expressive power of sentences, all elementary morphisms are embeddings (where 'embeddings'

[9] $h_{A,h} : A_A \to B_h$ is a morphism in Σ_A and $h_{a,b} : A_a \leftarrow B_b$ is a morphism in Σ_{M_φ}—recall the notational conventions about elementary diagrams.

can be defined either strictly categorically, as subobjects, or by means of inclusion or factorization systems), would be very desirable.

A more in depth study of elementary morphisms in some particular cases might also prove to be interesting. Take for instance the equational framework. In EQL and EQLN, the elementary morphisms do not seem very amenable. It is not clear to us how they look like. Note that a surjective morphism is always elementary in EQL, and an elementary morphism has to be injective in EQLN. The case of HL is even more intricate, and the notion of 'elementary behavioral morphism', complementing that of bisimulation, is potentially fruitful in the algebraic study of systems and behaviour.

Acknowledgements

Răzvan Diaconescu taught us institution-independent model theory during some lectures that he held at Şcoala Normală Superioară, Romania. He also read several preliminary versions of this article, and his comments improved the article substantially. Our colleagues, especially Traian Şerbănuţă, Marius Petria, and Cristian Cucu, pointed out lots of bugs in our proofs and minds with respect to elementary chains that we hope we managed to correct. Last but not least, our two referees, who criticized the article constructively.

References

[1] J. Barwise and J. Feferman. *Model-Theoretic Logics*. Springer, New York, 1985.
[2] M. Bidoit and R. Hennicker. Proving behavioural theorems with standard firstorder logic. *In Algebraic and Logic Programming (ALP'94)*, vol. 850 of LNCS, pp. 41–58, 1994.
[3] P. Burmeister. *A Model Theoretic Oriented Appraoch to Partial Algebras*. Akademie-Verlag Berlin, 1986.
[4] C. C. Chang and H. J. Keisler. *Model Theory*. North Holland, Amsterdam, 1973.
[5] M. Clavel, S. Eker, P. Lincoln and J. Meseguer. Principles of Maude. In *Proceedings of the First International Workshop on Rewriting Logic,* J. Meseguer, ed, vol. 4 *of Electronic Notes in Theoretical Computer Science*, pp. 65–89. Elsevier, 1996.
[6] R. Diaconescu. *Institution-independent Model Theory*. Forthcoming, Springer.
[7] R. Diaconescu. Institution-independent ultraproducts. *Fundamenta Informaticae*, **55**, 321–348, 2003.
[8] R. Diaconescu. Elementary diagrams in institutions. *Journal of Logic and Computation*, **14**, 651–674, 2004.
[9] R. Diaconescu. Herbrand theorems in arbitrary institutions. *Information Processing Letters*, **90**, 29–37, 2004.
[10] R. Diaconescu. An institution-independent proof of Craig interpolation theorem. *Studia Logica*, **77**, 59–79, 2004.
[11] R. Diaconescu and K. Futatsugi. *CafeOBJ Report: The Language, Proof Techniques, and Methodologies for Object-Oriented Algebraic Specification*. World Scientific, 1998. AMAST Series in Computing, vol. 6.
[12] R. Diaconescu and K. Futatsugi. Behavioural coherence in object-oriented algebraic specification. *Universal Computer Science*, **6**, 74–96, 2000.
[13] D. Găină and A. Popescu. An institution-independent proof of Robinson consistency theorem. *Studia Logica*, in press.

[14] J. Goguen and R. Burstall. Institutions: Abstract model theory for specification and programming. *Journal of the Association for Computing Machinery*, **39**, 95–146, 1992.

[15] J. Goguen and G. Malcolm. A hidden agenda. *Theoretical Computer Science*, **245**, 55–101, 2000.

[16] J. Goguen and G. Roşu. Hiding more of hidden algebra. In *Formal Methods 1999 (FM'99)*, vol. 1709 of *Lecture Notes in Computer Sciences*, pp. 1704–1719. Springer-Verlag, 1999.

[17] J. Hirschfeld and W. Wheeler. Forcing, arithmetic, division rings. *Lecture Notes in Mathematics*, (454), 1975.

[18] H. J. Keisler. *Model Theory for Infinitary Logic*. North-Holland, 1971.

[19] S. Mac Lane. *Categories for the Working Mathematician*. Springer, New York, 1971.

[20] M. Makkai and R. Paré. *Accessible Categories: The Foundations of Categorical Model Theory*. Providence, 1989.

[21] M. Makkai and G. Reyes. First order categorical logic. *Lecture Notes in Mathematics*, (611), 1977.

[22] J. Meseguer. General logics. In H.-D. Ebbinghaus *et al.*, ed., *Proceedings, Logic Colloquium* 1987, pp. 275–329. North-Holland, 1989.

[23] J. Meseguer. Conditional Rewriting Logic as a Unitied Model of Concurrency. *Theoretical Computer Science*, pp. 73–155, 1992.

[24] J. Meseguer. Membership algebra as a logical framework for equational specification. In *Proceedings, WADT'97*, vol. 1376 of *LNCS*, pp. 18–61. Tarquinia, Italy, 1998.

[25] J. D. Monk. *Mathematical Logic*. Springer-Verlag, New York, 1976.

[26] T. Mossakowski, J. Goguen, R. Diaconescu, and A. Tarlecki. What is a logic? In J.-Y. Beziau, ed., *Logica Universalis*, pp. 113–133. Birkhauser, 2005.

[27] M. Petria and R. Diaconescu. Abstract Beth definability institutionally. *Journal of Symbolic Logic*, in press.

[28] H. Reichel. *Structural Induction on Partial Algebras*. Akademie-Verlag Berlin, 1984.

[29] G. Roşu. The institution of order-sorted equational logic. *Bulletin of EATCS*, **53**, 250–255, 1994.

[30] G. E. Sacks. *Saturated Model Theory*. W. A. Bejamin, Inc., Reading, 1972.

[31] A. Salibra and G. Scollo. A soft stairway to institutions. In *Recent Trends in Data Type Specification*, M. Bidoit and C. Choppy, eds, vol. 655 of *Lecture Notes in Computer Science*, pp. 310–329. Springer-Verlag, Dourdan, France, 1992.

[32] A. Salibra and G. Scollo. Interpolation and compactness in categories of preinstitutions. *Mathematical Structures in Computer Science*, **6**, 261–286, 1996.

[33] T. Şerbănuţă. Institutional concepts in first-order logic, parameterized specifications, and logic programming. Master Disertation. University of Bucharest, 2004.

[34] A. Tarlecki. Bits and pieces of the theory of institutions. In D. Pitt, S. Abramsky, A. Poigné, and D. Rydeheard, eds, *Proceedings, Summer Workshop on Category Theory and Computer Programming, Lecture Notes in Computer Science*, vol. 240, pp. 334–360. Springer, GuildFord, 1986.

[35] A. Tarlecki. On the existence of free models in abstract algebraic institutions. *Theoretical Computer Science*, **37**, 269–304, 1986.

[36] A. Tarlecki. Quasi-varieties in abstract algebraic institutions. *Journal of Computer and System Sciences*, **33**, 333–360, 1986.