The institution-theoretic scope of logic theorems

Răzvan Diaconescu

Simion Stoilow Institute of Mathematics of the Romanian Academy

Till Mossakowski

DFKI GmbH Bremen, Germany, and University of Bremen, Germany

Andrzej Tarlecki

Institute of Informatics, University of Warsaw, Poland

Abstract

In this essay we analyse and clarify the method to establish and clarify the scope of logic theorems offered within the theory of *institutions*. The method presented pervades a lot of abstract model theoretic developments carried out within institution theory. The power of the proposed general method is illustrated with the examples of (Craig) interpolation and (Beth) definability, as they appear in the literature of institutional model theory. Both case studies illustrate a considerable extension of the original scopes of the two classical theorems. Our presentation is rather narrative with the relevant logic and institution theory concepts introduced and explained gradually to the non-expert reader.

1. Institution theory – a very brief introduction

Institution theory is a categorical abstract model theory that arose about three decades ago [20] as a response to the explosion of the population of logical systems used for formal software specification. Its original aim was to develop as much computing science as possible in a general, uniform way, independently of particular logical systems. This has been achieved to an extent even greater than originally envisaged. The theory of institutions became the most fundamental mathematical tool underlying algebraic specification theory (in its wider meaning) [29], also being increasingly used in other areas of computer science. Moreover, institution theory is a major trend in the so-called 'universal logic' (in the sense envisaged by Jean-Yves Béziau [4, 5]) which is considered by many a true renaissance of mathematical logic. A lot of model theory has been gradually developed at the level of abstract institutions (see [14]). A relatively recent survey of the vast area of institution theory is [16].

The starting concept of institution theory is the formal definition of a logical system; this includes the syntax, the semantics and the satisfaction relation between them. It plays the same role as for example the definition of group plays for group theory. Although the definition of group is very simple, group theory is a vast sophisticated mathematical area. The same with the definition of institution and institution theory.

An *institution* is a tuple $(\mathrm{Sign}, \mathrm{Sen}, \mathrm{Mod}, (\models_\Sigma)_{\Sigma \in |\mathrm{Sign}|})$ that consists of

Email addresses: `Razvan.Diaconescu@imar.ro` (Răzvan Diaconescu), `Till.Mossakowski@dfki.de` (Till Mossakowski), `tarlecki@mimuw.edu.pl` (Andrzej Tarlecki)

Preprint submitted to UNILOG 2013 CONTEST *February 18, 2013*

- a category Sign whose objects are called *signatures*,
- a functor Sen: Sign → **Set** (to the category of sets) giving for each signature a set whose elements are called *sentences* over that signature,
- a (contravariant) functor Mod: $(\mathrm{Sign})^{op} \to$ **CAT** (to the 'category' of categories), giving for each signature Σ a category whose objects are called Σ-*models*, and whose arrows are called Σ-*(model) homomorphisms*, and
- a relation $\models_\Sigma \subseteq |\mathrm{Mod}(\Sigma)| \times \mathrm{Sen}(\Sigma)$ for each $\Sigma \in |\mathrm{Sign}|$, called the *satisfaction relation*,

such that for each morphism $\varphi\colon \Sigma \to \Sigma' \in \mathrm{Sign}$, the *Satisfaction Condition*

$$M' \models_{\Sigma'} \mathrm{Sen}(\varphi)(\rho) \text{ if and only if } \mathrm{Mod}(\varphi)(M') \models_\Sigma \rho \tag{1}$$

holds for each $M' \in |\mathrm{Mod}(\Sigma')|$ and $\rho \in \mathrm{Sen}(\Sigma)$.

The literature (e.g. [14, 29]) shows myriads of logical systems from computing or from mathematical logic captured as institutions. In fact, an informal thesis underlying institution theory is that any 'logic' may be captured by the above definition. While this should be taken with a grain of salt, it certainly applies to any logical system based on satisfaction between sentences and models of any kind.

Given a signature Σ in an institution, for any sets E and E' of Σ-sentences by $E \models E'$ we denote the situation that for each Σ-model M, $M \models E$ implies $M \models E'$. The institution is *compact* when given $E \models E'$ if E' is finite then there exists finite $E_0 \subseteq E$ such that $E_0 \models E'$.

2. The method to clarify the scope of logic theorems

Among mathematicians it is common to think that there cannot be an in-depth understanding of a result in the absence of the understanding of the proof of this result. The understanding of the scope of logic theorems may be therefore considered at two different levels. A coarse level refers to the actual results (statements of logic theorems), and a subtle level refers to the causalities leading to the results (i.e. dependencies among logic theorems and methods to prove these theorems). Clearly, answers at the subtle level determine answers at the coarse level.

Institution theory and its abstract approach to model theory develop a distinctive and clear way to determine the scope of logic theorems, with emphasis on revealing bare causalities that are stripped off irrelevant details. In many situations this led not only to quite unexpected extensions of the scopes of logic theorems, but even to a reformed and a more realistic understanding of fundamental logic concepts (these points will be illustrated also by the case studies discussed in this essay). The following scheme captures a method to determine the scope of logic theorems that proved effective in numerous developments within institution theory:

1. Choose a proof of a logic theorem.
2. Extract its essence by leaving out the irrelevant details and by identifying the conceptual structure and the causalities underlying the result.
3. Formulate the conceptual structure at the level of an abstract institution.
4. Lift the proof considered to the level of an abstract institution, shaping an abstract, generic scope of the result.
5. Determine the actual scope by analysing the abstract conditions used in the proof.

In the following we illustrate this scheme with a few representative cases from institutional model theory.

3. First case study: interpolation by axiomatisability

Because of its many applications in logic and computer science, interpolation is one of the most desired and studied properties of logical systems. Although it has a strikingly simple and elementary formulation, in general it is very difficult to establish. The famous result of Craig [9] marks perhaps the birth of the study of interpolation, proving it for first-order logic. The actual scope of Craig's result has been gradually extended to many other logical systems (for example in the world of modal logics, see [19]).

It has been widely believed that equational logic lacks interpolation; likewise for Horn-clause logic and other such fragments of first-order logic. As far as we know, Piet Rodenburg was the first to point out that this is a misconception due to a basic misunderstanding of interpolation, rooted in the heavy dependency of logic culture on classical first-order logic with all its distinctive features taken for granted. Then it follows the grave fault of exporting a coarse understanding of concepts dependent on details of a particular logical system to other logical systems of a possibly very different nature, where some detailed features may not be available. In the case of interpolation, the gross confusion has to do with looking for an interpolant as a single sentence. In first-order logic, which has conjunction, looking for interpolants as finite sets of sentences ($\{\rho_1, \ldots, \rho_n\}$) is just the same as looking for interpolants as single sentences ($\rho_1 \wedge \cdots \wedge \rho_n$). Hence, the common formulation of interpolation requires a single-sentence interpolant. However, this is not an adequate formulation for equational logic which lacks conjunction, i.e., conjunction $\rho_1 \wedge \rho_2$ of universally quantified equations ρ_1 and ρ_2 cannot be captured as a universally quantified equation in general. Rodenburg [27, 28] proved that equational logic has interpolation with the interpolant being a finite set of sentences, and this apparently weaker interpolation property is quite sufficient in both computer science and logic applications. While the proof in [27] is syntactic and consists of rather lengthy combinatorial arguments, the second proof reported in [28] is purely algebraic and displays a very elegant reliance upon Birkhoff's variety theorem (any class of algebras over a common signature is axiomatisable by a set of equations if and only if it is closed under products, sub-algebras, and homomorphic images). We take the proof of [28] as item 1. of the scheme sketched in Sect. 2 to determine the scope of interpolation by axiomatisability.

At item 2. of the scheme, we go through a process of understanding that the essence of the proof in [28] is in fact independent of equational logic and of Birkhoff's variety theorem. The key is a causal relationship between interpolation and a 'Birkhoff-style' axiomatisability property of the logic, both of them considered in a rather general sense. In other words, the proof in [28] carries over in essence to various logics that support some kind of axiomatisability property resembling Birkhoff's variety theorem.

At the next step (item 3.) we work towards formalising the above understanding of the proof in [28]

- by defining the concept of interpolation at the level of abstract institutions, and
- by formulating a general axiomatisability concept at the level of abstract institutions that captures the essence of the use of Birkhoff's variety theorem in the proof of [28] as abstractly as possible.

Interpolation in abstract institutions.

The standard formulation of (Craig) interpolation property for first-order logic is as follows. Given signatures Σ_1, Σ_2, Σ_1-sentence ρ_1 and Σ_2-sentence ρ_2, if ρ_2 is a consequence of ρ_1 (written $\rho_1 \vdash \rho_2$) then there exists an 'interpolant' $(\Sigma_1 \cap \Sigma_2)$-sentence ρ such that $\rho_1 \vdash \rho$ and $\rho \vdash \rho_2$.

It is by far not straightforward how to express this property at the level of abstract institutions. First, we have to interpret the consequence relation between (sets of) sentences $\vdash$ as the semantic consequence $\models$, which is naturally defined in any institution. Then, in order to free our discussion from the existence of conjunctions, we replace single sentences by finite sets of sentences. Finally, we have to capture the relationship between signatures Σ_1, Σ_2 and their union $\Sigma_1 \cup \Sigma_2$ (where the consequence $\rho_1 \vdash \rho_2$ happens) and intersection $\Sigma_1 \cap \Sigma_2$ (the signature of the interpolant), depicted by the following diagram where arrows indicate the obvious inclusions:

$$\begin{array}{ccc} \Sigma_1 \cap \Sigma_2 & \longrightarrow & \Sigma_1 \\ \downarrow & & \downarrow \\ \Sigma_2 & \longrightarrow & \Sigma_1 \cup \Sigma_2 \end{array}$$

While intersections $\cap$ and unions $\cup$ are more of less obvious for signatures as used in first-order logic and in many other standard logics, they are not so in some other logical systems, and certainly not at the level of abstract institutions where signatures are just objects of an arbitrary category. One immediate response to this problem would be to add an infrastructure to the abstract category of signatures that would support concepts of $\cap$ and $\cup$; in fact this is already available in the institution theoretic literature and is called *inclusion system* [17, 14]. However this may not be the best solution. In many computer science applications it is very meaningful to consider non-inclusive signature morphisms in the role of inclusions in the square above. The abstraction exercise concerning interpolation we undertake is an ideal occasion to follow [31] and incorporate the corresponding generalisation as well. The category-theoretic property of the above intersection-union square that makes things work is that it is a *pushout*. These considerations lead to the following abstract formulation of the interpolation property.[1]

An institution has *Craig interpolation (Ci)* when for each pushout square of signatures

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

and any finite sets of sentences $E_1 \subseteq \mathrm{Sen}(\Sigma_1)$ and $E_2 \subseteq \mathrm{Sen}(\Sigma_2)$, if $\theta_1(E_1) \models \theta_2(E_2)$ then there exists a finite set E of Σ-sentences such that $E_1 \models \varphi_1(E)$ and $\varphi_2(E) \models E_2$.

However, this concept proves a bit too strong (for example many-sorted first-order logic does not support this [21, 7, 14]). It is useful to restrict (of course abstractly) the signature morphisms involved to pre-defined classes of signature morphisms, $\mathcal{L}$ for φ_1 and $\mathcal{R}$ for φ_2. We thus arrive at a realistic concept of *Craig* $(\mathcal{L}, \mathcal{R})$-*interpolation* which these days prevails institution-theory literature (e.g. [14]).

The use of pushout squares above meets so-called *model amalgamation* property, which is a crucial technical property pervading most of the developments in institutional model theory. This requires that for any pushout of signatures as above, Σ_1-model M_1 and Σ_2-model M_2 with common reduct to Σ, i.e., $\mathrm{Mod}(\varphi_1)(M_1) = \mathrm{Mod}(\varphi_2)(M_2)$, admit a unique common expansion M' to Σ', i.e., $\mathrm{Mod}(\theta_1)(M') = M_1$ and $\mathrm{Mod}(\theta_2)(M') = M_2$. This property is evident in most institutions of interest, and is tacitly assumed in many model-theoretic developments. Often its weaker variant, that does not require the uniqueness of M', suffices; this is called *weak model amalgamation*.

Abstract Birkhoff institutions.

The other concept that plays a crucial role in our analysis of the scope of interpolation by axiomatisability is that of Birkhoff(-style) axiomatisability. Let us start from the classical result of Birkhoff [6] about the equationally defined classes of algebras, which gives the following algebraic characterisation of the

---

[1]Given a signature morphism $\varphi\colon \Sigma \to \Sigma'$, we abbreviate $\mathrm{Sen}(\varphi)$ as φ, and so for a set of sentences $E \subseteq \mathrm{Sen}(\Sigma)$, $\varphi(E)$ is the image of E under $\mathrm{Sen}(\varphi)$.

so-called *equationally axiomatisable hulls*: for any class $\mathcal{A}$ of algebras (of a common signature) we have that

$$\mathcal{A}^{**} = HSP(\mathcal{A})$$

where $\mathcal{A}^*$ is the set of the (universally quantified) equations satisfied by all algebras in $\mathcal{A}$, $\mathcal{A}^{**}$ is the class of algebras that satisfy all sentences of $\mathcal{A}^*$, and $HSP(\mathcal{A})$ is the closure of $\mathcal{A}$ first under direct products (P), then under sub-algebras (S), and finally under homomorphic images (H). The Birkhoff variety theorem is just one of a large spectrum of algebraic characterisations of axiomatisable hulls $\mathcal{M}^{**}$ as closures of $\mathcal{M}$ first under a certain class of filtered products, and then under some relation between models defined in terms of certain classes of model homomorphisms. The class of the filtered products and the relation considered are interdependent with the kind of sentences considered. For example, if in the Birkhoff variety theorem discussed above we generalise the sentences to conditional equations (equations with finite sets of equational premises) then we have to consider closure under all filtered products instead of P, and closure under isomorphisms instead of H — this is another famous axiomatisability result, the so-called quasi-variety theorem of Mal'cev [24].

The concept of *Birkhoff institution* [12, 14] below captures abstractly this kind of algebraic characterisations of axiomatisable hulls. A tuple $(\mathrm{Sign}, \mathrm{Sen}, \mathrm{Mod}, \models, \mathcal{F}, \mathcal{B})$ is a *Birkhoff institution* when

- $(\mathrm{Sign}, \mathrm{Sen}, \mathrm{Mod}, \models)$ is an institution such that for each signature $\Sigma \in |\mathrm{Sign}|$, the category $\mathrm{Mod}(\Sigma)$ of Σ-models has $\mathcal{F}$-filtered products,
- $\mathcal{F}$ is a class of filters with $\{\{*\}\} \in \mathcal{F}$, and
- $\mathcal{B}_\Sigma \subseteq |\mathrm{Mod}(\Sigma)| \times |\mathrm{Mod}(\Sigma)|$ is a binary relation for each signature $\Sigma \in |\mathrm{Sign}|$, which is closed under isomorphisms, i.e., $(\mathcal{B}_\Sigma \circ \cong_\Sigma) = \mathcal{B}_\Sigma = (\cong_\Sigma \circ \mathcal{B}_\Sigma)$ (where $\cong_\Sigma$ denotes the isomorphism relation between Σ-models),

such that

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

for each signature Σ and each class of Σ-models $\mathcal{M} \subseteq |\mathrm{Mod}(\Sigma)|$, and where $\mathcal{F}\mathcal{M}$ is the class of all F-filtered products of models from $\mathcal{M}$ for all filters $F \in \mathcal{F}$. (Given a filter F over a set I, the *F-filtered product* of a family $\{M_i \mid i \in I\}$ of Σ-models is defined as the co-limit of the F-shaped diagram of projections between corresponding direct products of models from the family; this is the common concept of filtered product from categorical model theory [25, 23, 1, 11, 14].) $\mathcal{B}_\Sigma^{-1}(\mathcal{F}\mathcal{M})$ stands for the class $\{M' \mid$ there exists $M \in \mathcal{F}\mathcal{M}$ such that $(M', M) \in \mathcal{B}_\Sigma\}$.

From the perspective of the problem of the scope of logic theorems, the concept of Birkhoff institution may be regarded as a very abstract definition of the (coarse-level) scope of Birkhoff's variety theorem. The literature abounds with examples of Birkhoff institutions that correspond to various axiomatisability results (a general result with hundreds of concrete instances has been the subject of [1]).

Interpolation theorem by axiomatisability.

Let us now get back to our general scheme of Sect. 2, to item 4. The result below (taken from [14] and extending the original version in [12]) gives a *generic scope* for interpolation by axiomatisability:

Theorem 3.1. *Any Birkhoff institution* $(\mathrm{Sign}, \mathrm{Sen}, \mathrm{Mod}, \models, \mathcal{F}, \mathcal{B})$ *with the weak model amalgamation property, has Craig* $(\mathcal{L}, \mathcal{R})$*-interpolation when*

1. *for each* $\varphi \in \mathcal{L}$*,* $\mathrm{Mod}(\varphi)$ *preserves* $\mathcal{F}$*-filtered products, and*
2. – *for each* $\varphi \in \mathcal{R}$*,* $\mathrm{Mod}(\varphi)$ *lifts* $\mathcal{B}$*, or*

– *for each* $\varphi \in \mathcal{L}$, Mod(φ) *lifts* $\mathcal{B}^{-1}$ *and model isomorphisms.*

The requirements in this theorem (which include the conditions underlying the definition of Birkhoff institutions) represent a clarification of the system of detailed technical causalities in the proof of interpolation by axiomatisability. Their analysis corresponds to item 5. of our scheme, that determines the *actual scope*. The conditions of amalgamation and preservation of filtered products are rather obvious in the concrete situations (at least in the conventional institutions). They are thus easy to overlook; however, at the abstract level their role becomes evident. The lifting conditions are less obvious, even if in the applications they may be established rather smoothly (see [12, 14]). Their definition is as follows. For $\varphi\colon \Sigma \to \Sigma'$, Mod($\varphi$) *lifts* $\mathcal{B}$ when for each Σ'-model M' and each Σ-model N such that $(\mathrm{Mod}(\varphi)(M'), N) \in \mathcal{B}_\Sigma$ there exists a φ-expansion N' of N such that $(M', N') \in \mathcal{B}_{\Sigma'}$. Semantic in nature, the lifting conditions represent the technical link between the closure by $\mathcal{B}$ in the algebraic characterisation of the axiomatisable hulls (the formula $\mathcal{B}_\Sigma^{-1}(\mathcal{F}\mathcal{M})$) and the syntactic restriction on the signature morphisms involved in the interpolation property.

In the rather classical context of fragments of (many-sorted) first-order logic, let us give two perhaps surprising concrete instances of the general interpolation theorem by axiomatisability.

Corollary 3.1. *[14] Many-sorted Horn-clause logic (with equality) has* $(\mathcal{L},\mathcal{R})$*-interpolation for any classes* $\mathcal{L}$ *and* $\mathcal{R}$ *of signature morphisms such that*

– $\mathcal{R}$ *consists only of injective signature morphisms, or*
– $\mathcal{L}$ *consists only of signature morphisms that are injective on the sorts and that 'encapsulate' the operations in the sense that there is no 'new' operation with an 'old' result sort.*

4. Second case study: definability by interpolation

In first-order logic, an operation or relation symbol σ is *defined implicitly* by a theory E' in a signature Σ' when the interpretation of σ in any model M' of E' is uniquely defined by the interpretations of all the other symbols of Σ' in M'. An example is the inverse operation in group theory: it is an easy exercise to see that any monoid may be expanded by at most one inverse operation so that the expanded model is a group. This notion of definability is clearly semantic. There is also a syntactic counterpart: σ is *defined explicitly* by E' when for each finite set of variables X and each $(\Sigma' + X)$-sentence ρ there exists a $((\Sigma' \setminus \{\sigma\}) + X)$-sentence E_ρ such that $E' \vdash (\forall X)(\rho \Leftrightarrow E_\rho)$. The fact that the inverse operation is defined explicitly by the theory of groups is much less obvious than in the case of implicit definability. Beth's definability theorem [3] establishes the equivalence between these two notions of definability in first-order logic. Since (in first-order logic) it is trivial to see that explicit definability implies implicit definability, the real substance of Beth's definability result is the opposite implication. In fact, in the literature the term "Beth definability result" often refers to this difficult derivation of explicit from implicit definability. The standard proof of this result in classical first-order model theory [8, 22] relies on Craig interpolation and uses classical implication. However, in [26] it is shown that Beth definability may be derived more directly from axiomatisability properties. In the following we study the scope of Beth's definability theorem *by interpolation*. In the process we will see how the classical first-order logic framework may lead to a misunderstanding of *how* interpolation causes definability.

Proceeding to item 3. of our general scheme in Sect. 2, we establish the concepts of definability, interpolation, and implication at the level of abstract institutions. Interpolation has already been discussed above, so we are left with the other two concepts.

Definability in abstract institutions.

We present here an abstract institutional version of definability following [26], see also [14].

As formulated above, the classical concept of definability concerns a signature extension by an individual new symbol (such as the inverse operation of groups); the natural generalisation is to consider an arbitrary signature morphism. This may appear as a big generalisation step, and one may think to recover some of the details of the classical situation by considering only inclusions here, or imposing some finiteness properties on the signature morphism. However, it turns out that for definability these are irrelevant details, and moreover it is definitely simpler to do without them. Hence we say that a signature morphism $\varphi\colon \Sigma \to \Sigma'$ is *defined implicitly by* a Σ'-theory E' when the corresponding reduct $\mathrm{Mod}(\varphi)\colon \mathrm{Mod}(\Sigma', E') \subseteq \mathrm{Mod}(\Sigma') \to \mathrm{Mod}(\Sigma)$ is injective. This is quite a straightforward lifting of implicit definability to the abstract institutional setting we consider. A signature morphism $\varphi\colon \Sigma \to \Sigma'$ is *defined explicitly by* E' when for each pushout square

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

and each sentence $\rho \in \mathrm{Sen}(\Sigma'_1)$, there exists a finite set of sentences $E_\rho \subseteq \mathrm{Sen}(\Sigma_1)$ such that

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

The above formulation does not assume that the set of sentences of our institution is closed under the equivalence connective and universal quantification. It should simply be read as follows: each θ'-expansion of each model of E' satisfies ρ if and only if it satisfies $\varphi_1(E_\rho)$. Referring to the above classical definition of explicit definability in first-order logic, θ captures the set X of variables; thus θ (and consequently θ') extends a signature by a finite number of variables. This kind of additional restriction on θ may be handled also at the abstract level with some technical effort. However it turns out that this is another unnecessary detail which may be skipped at the abstract level — and so we may consider arbitrary signature morphisms here. The pushout square is used to capture the translation of quantified formulae that is standard in institution theory; in the classical framework of first-order logic such a translation is taken for granted, but at the abstract level here it becomes visible and its expected properties have to be ensured. A more substantial difference with respect to the definition of definability in first-order logic is that we permit E_ρ to be a finite *set of sentences* rather than a single sentence; this is quite similar to what have been discussed above for the concept of interpolant, and plays the same role.

A signature morphism φ *has the definability property* if and only if a theory defines φ explicitly whenever it defines φ implicitly. We will not discuss the opposite implication (deriving implicit from explicit definability); we just note that while this is immediate within first-order logic (and therefore gets almost no attention in the literature), as shown in [26] it is highly non-trivial at the general level of abstract institutions.

Implication; but is it really needed?

Besides interpolation, the first-order logic proof of Beth's definability theorem uses implication — this feature of first-order logic so obvious that it is hardly ever mentioned explicitly in this context. Its definition at the level of abstract institutions is straightforward [31]: an institution *has implication* when for every signature Σ and Σ-sentences ρ_1, ρ_2, there exists a Σ-sentence ρ such that for each Σ-model M,

$$M \models \rho \text{ if and only if } M \models \rho_2 \text{ whenever } M \models \rho_1.$$

However, in the context of definability results, we may render implication unnecessary by reformulating the interpolation property. The trick is to 'parametrise' each instance of interpolation by a set of 'secondary' premises. In [18, 30, 33] this is called *Craig-Robinson interpolation (*CRi*)*; it also plays an important role in specification theory, e.g. [2, 17, 18, 14]. Let us recall here explicitly its institution-theoretic formulation. An institution has *Craig-Robinson* $(\mathcal{L},\mathcal{R})$*-interpolation* when for each pushout square of signatures with $\varphi_1 \in \mathcal{L}$ and $\varphi_2 \in \mathcal{R}$

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

and finite sets of sentences $E_1 \subseteq \mathrm{Sen}(\Sigma_1)$ and $E_2, \Gamma_2 \subseteq \mathrm{Sen}(\Sigma_2)$, if $\theta_1(E_1) \cup \theta_2(\Gamma_2) \models \theta_2(E_2)$ then there exists a finite set E of Σ-sentences such that $E_1 \models \varphi_1(E)$ and $\varphi_2(E) \cup \Gamma_2 \models E_2$.

Clearly, *CRi* implies *Ci*. In any compact institution with implication, *CRi* and *Ci* are equivalent [15, 14] (so for instance within first-order logic, the two properties coincide). This means that *CRi* alone in principle is weaker than *Ci* and implication. But is it properly so? Is there a significant example of an institution lacking implication but having *CRi*? Through a rather sophisticated technique of so-called Grothendieck institutions [10, 13], a result in [14] gives a general method to lift *Ci* to *CRi* in institutions that may not have implication but are embedded in a certain way into institutions having implication. The following corollary is an example of a consequence of this result based on the *Ci* property stated in Cor. 3.1.

Corollary 4.1. *Many-sorted Horn-clause logic (with equality) has* $(\mathcal{L},\mathcal{R})$*-CRi when* $\mathcal{L}$ *consists only of signature morphisms that are injective on sorts and 'encapsulate' the operations.*

As hinted above, in the context of definability results, *CRi* may replace *Ci* and implication: according to [26], the proof of Beth's definability theorem can be carried out using *CRi* alone instead of *Ci* and implication. While such a change is not of interest for the classical Beth definability theorem for first-order logic (which has implication), this general result becomes very interesting in other logical contexts, where implication may be lacking.

Definability theorem by interpolation.

The generic scope of Beth's definability theorem is given by the following result from [26], which corresponds to item 4. of the scheme given in Sect. 2.

Theorem 4.1. *In any compact institution that has the model amalgamation property and* $(\mathcal{L},\mathcal{R})$*-CRi for classes of signature morphisms* $\mathcal{L}$*,* $\mathcal{R}$ *that are stable under pushouts, any signature morphism in* $\mathcal{L} \cap \mathcal{R}$ *has the definability property.*

The actual scope of definability by interpolation is determined by the analysis of the conditions of Thm. 4.1. Compactness is a common property of institutions (or logics they capture) that may be established by various means, such as completeness of a finitary proof calculus, or preservation by ultraproducts (for the latter method at the level of abstract institutions see [11, 14]). Model amalgamation is another common condition that has already been been discussed. The above discussion of *CRi* also makes demands behind these requirements sufficiently clear. Stability of $\mathcal{L}$ under pushouts means that in any pushout square of signature morphisms as below

$$\begin{array}{ccc} \Sigma & \xrightarrow{\varphi} & \Sigma' \\ {\scriptstyle\theta}\downarrow & & \downarrow{\scriptstyle\theta'} \\ \Sigma_1 & \xrightarrow[\varphi']{} & \Sigma_1' \end{array}$$

if $\varphi \in \mathcal{L}$ then $\varphi' \in \mathcal{L}$ too (and similarly for $\mathcal{R}$). This is a purely technical condition, which is rather mild for typical choices of $\mathcal{L}$ and $\mathcal{R}$ in applications. Summing up, this analysis shows that the only substantial condition for definability by interpolation is the *CRi* property.

Let us present two sample applications of Thm. 4.1. The following corollary represents a twofold extension of the scope of Beth's definability theorem in first-order logic, to the many-sorted case and to arbitrary signature morphisms. It is based on the interpolation result for many-sorted first-order logic [7, 21].

Corollary 4.2. *In many-sorted first-order logic any signature morphism that is injective on sorts has the definability property.*

Another corollary extends the scope of Beth's definability theorem to a logic without implication. This is a consequence of Thm. 4.1 and Cor. 4.1.

Corollary 4.3. *In many-sorted Horn-clause logic any signature morphism that is injective on sorts and 'encapsulates' the operations has the definability property.*

5. Conclusions

We have put forward a general institution-theoretic scheme for clarifying the scope of logic theorems and illustrated it with the cases of Craig interpolation and Beth definability results. Our approach distinguishes two different levels: that of the scope of results themselves and another, more subtle, of the scope of methods to prove the results. We claim that the abstraction process involved in developments of logic theorems within the theory of institutions almost always implies a clarification and a significant expansion of the scope of the most important logic concepts involved, often correcting some common conceptual misunderstandings of rather subtle nature. We have noticed this also with our two case studies.

Lack of space prevented us from discussing other important achievements that clarify the scope of logic theorems in institution theory, following the general pattern proposed by this essay. They include Birkhoff and Gödel completeness theorems, compactness theorems (by ultraproducts as well as by completeness), Craig interpolation (by Robinson consistency), Beth definability (by axiomatisability), the Keisler-Shelah isomorphism theorem (by saturated ultraproducts), axiomatisability theorems, etc.

For the same reason we restricted our presentation of the concrete scope of Craig interpolation and Beth definability to the rather conventional contexts of simple fragments of many-sorted first-order logic, where we presented general results that significantly expand the scope of both theorems. We think that Craig interpolation and Beth definability in the many-sorted context considerably strengthen arguments [32] that dramatically invalidate a rather common view expressed for instance in [30]: namely, that many-sorted logics are 'inessential variations' on their single-sorted versions. Institution theory literature contains many examples of less conventional logics, which for instance may involve partial functions, ordered models, modalities and Kripke semantics, many-valued truth, etc. Although we have not discussed this here explicitly, it should be clear that the general results outlined in this essay easily instantiate to such contexts as well, and determine the scopes of so generalised logic theorems in a variety of new logical situations at hand.

Acknowledgements

The work has been partially supported by a grant of the Romanian National Authority for Scientific Research, CNCS-UEFISCDI, project number PN-II-ID-PCE-2011-3-0439 (RD), DFG CRC/TR 8 "Spatial Cognition", project I1-[OntoSpace] (TM), and by the Polish Ministry of Science and Higher Education grant N206 493138 (AT).

We thank Oliver Kutz for comments on a draft of this paper.

References

[1] Hajnal Andréka and István Németi. A general axiomatizability theorem formulated in terms of cone-injective subcategories. In B. Csakany, E. Fried, and E.T. Schmidt, editors, *Universal Algebra*, pages 13–35. North-Holland, 1981. Colloquia Mathematics Societas János Bolyai, 29.

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

[3] Evert Willem Beth. On Padoa's method in the theory of definition. *Indagationes Mathematicæ*, 15:330–339, 1953.

[4] Jean-Yves Béziau. 13 questions about universal logic. *Bulletin of the Section of Logic*, 35(2/3):133–150, 2006.

[5] Jean-Yves Béziau, editor. *Universal Logic: an Anthology*. Studies in Universal Logic. Springer Basel, 2012.

[6] Garrett Birkhoff. On the structure of abstract algebras. *Proceedings of the Cambridge Philosophical Society*, 31:433–454, 1935.

[7] Tomasz Borzyszkowski. Generalized interpolation in first-order logic. *Fundamenta Informaticæ*, 66(3):199–219, 2005.

[8] Chen-Chung Chang and H. Jerome Keisler. *Model Theory*. North Holland, Amsterdam, 1990.

[9] William Craig. Three uses of the Herbrand-Gentzen theorem in relating model theory and proof theory. *Journal of Symbolic Logic*, 22:269–285, 1957.

[10] Răzvan Diaconescu. Grothendieck institutions. *Applied Categorical Structures*, 10(4):383–402, 2002. Preliminary version appeared as IMAR Preprint 2-2000, ISSN 250-3638, February 2000.

[11] Răzvan Diaconescu. Institution-independent ultraproducts. *Fundamenta Informaticæ*, 55(3-4):321–348, 2003.

[12] Răzvan Diaconescu. An institution-independent proof of Craig Interpolation Theorem. *Studia Logica*, 77(1):59–79, 2004.

[13] Răzvan Diaconescu. Interpolation in Grothendieck institutions. *Theoretical Computer Science*, 311:439–461, 2004.

[14] Răzvan Diaconescu. *Institution-independent Model Theory*. Birkhäuser, 2008.

[15] Răzvan Diaconescu. Borrowing interpolation. *Journal of Logic and Computation*, 22(3):561–586, 2012.

[16] Răzvan Diaconescu. Three decades of institution theory. In Jean-Yves Béziau, editor, *Universal Logic: an Anthology*, pages 309–322. Springer Basel, 2012.

[17] Răzvan Diaconescu, Joseph Goguen, and Petros Stefaneas. Logical support for modularisation. In Gerard Huet and Gordon Plotkin, editors, *Logical Environments*, pages 83–130. Cambridge, 1993. Proceedings of a Workshop held in Edinburgh, Scotland, May 1991.

[18] Theodosis Dimitrakos and Tom Maibaum. On a generalized modularization theorem. *Information Processing Letters*, 74:65–71, 2000.

[19] Dov M. Gabbay and Larisa Maksimova. *Interpolation and Definability: modal and intuitionistic logics*. Oxford University Press, 2005.

[20] Joseph Goguen and Rod Burstall. Institutions: Abstract model theory for specification and programming. *Journal of the Association for Computing Machinery*, 39(1):95–146, 1992.

[21] Daniel Găină and Andrei Popescu. An institution-independent proof of Robinson consistency theorem. *Studia Logica*, 85(1):41–73, 2007.

[22] Wilfrid Hodges. *Model Theory*. Cambridge University Press, 1993.

[23] Michael Makkai. Ultraproducts and categorical logic. In C.A. DiPrisco, editor, *Methods in Mathematical Logic*, volume 1130 of *Lecture Notes in Mathematics*, pages 222–309. Springer Verlag, 1985.

[24] Anatoly Malcev. *The Metamathematics of Algebraic Systems*. North-Holland, 1971.

[25] Günter Matthiessen. Regular and strongly finitary structures over strongly algebroidal categories. *Canadian Journal of Mathematics*, 30:250–261, 1978.

[26] Marius Petria and Răzvan Diaconescu. Abstract Beth definability in institutions. *Journal of Symbolic Logic*, 71(3):1002–1028, 2006.

[27] Pieter-Hendrik Rodenburg. Interpolation in conditional equational logic, 1989. Preprint from Programming Research Group at the University of Amsterdam.

[28] Pieter-Hendrik Rodenburg. A simple algebraic proof of the equational interpolation theorem. *Algebra Universalis*, 28:48–51, 1991.

[29] Donald Sannella and Andrzej Tarlecki. *Foundations of Algebraic Specifications and Formal Software Development*. Springer, 2012.

[30] Joseph Shoenfield. *Mathematical Logic*. Addison-Wesley, 1967.

[31] Andrzej Tarlecki. Bits and pieces of the theory of institutions. In David Pitt, Samson Abramsky, Axel Poigné, and David Rydeheard, editors, *Proceedings, Summer Workshop on Category Theory and Computer Programming*, volume 240 of *Lecture Notes in Computer Science*, pages 334–360. Springer, 1986.

[32] Andrzej Tarlecki. Some nuances of many-sorted universal algebra: A review. *Bulletin of the EATCS*, 104:89–111, 2011.

[33] Paulo Veloso. On pushout consistency, modularity and interpolation for logical specifications. *Information Processing Letters*, 60(2):59–66, 1996.