Reducing first-order logic to Df_3, free algebras.

Hajnal Andréka and István Németi

Alfred Tarski in 1953 formalized set theory in the equational theory of relation algebras [37, 38]. Why did he do so? Because the equational theory of relation algebras (RA) corresponds to a logic without individual variables, in other words, to a propositional logic. This is why the title of the book [39] is "Formalizing set theory without variables". Tarski got the surprising result that a propositional logic can be strong enough to "express all of mathematics", to be the arena for mathematics. The classical view before this result was that propositional logics in general were weak in expressive power, decidable, uninteresting in a sense. By using the fact that set theory can be built up in it, Tarski proved that the equational theory of RA is undecidable. This was the first propositional logic shown to be undecidable.

From the above it is clear that replacing RA in Tarski's result with a "weaker" class of algebras is an improvement of the result and it is worth doing. For more on this see the open problem formulated in Tarski-Givant [39, p.89, line 2 bottom up – p.90, line 4 and footnote 17 on p.90].

A result of J. D. Monk says that for every finite n there is a 3-variable first-order logic (FOL) formula which is valid but which can be proved (in FOL) with more than n variables only (cf. [14, 3.2.85]). Intuitively this means that during any proof of this formula there are steps when we have to use n independent data (stored in the n variables as in n machine registers). For example, the associativity of relation composition of binary relations can be expressed with 3 variables but 4 variables are needed for any of its proofs.

Tarski's main idea in [39] is to use pairing functions to form ordered pairs, and so to store two pieces of data in one register. He used this technique to translate usual infinite-variable first-order logic into the three-variable fragment of it. From then on, he used the fact that any three-variable -formula about binary relations can be expressed by an RA-equation, [14, 5.3.12]. He used two registers for storing the data belonging to a binary relation and he had one more register available for making computations belonging to a proof.

The finite-variable fragment hierarchy of FOL corresponds to the appropriate hierarchy of cylindric algebras (CA_n's). The n-variable fragment $\mathcal{L}_n$ of FOL consists of all FOL-formulas which use only the first n variables. By Monk's result, $\mathcal{L}_n$ is essentially incomplete for all $n \geq 3$, it cannot have a finite Hilbert-style complete and strongly sound inference system. We get a finite Hilbert-style inference system $\vdash_{\overline{n}}$ for $\mathcal{L}_n$ by restricting a usual complete one for infinite-variable FOL to the first n variables (see [14, sec. 4.3]). This inference system

$\vdash_n$ belonging to $\mathcal{L}_n$ is a translation of an equational axiom system for CA_n, it is strongly sound but not complete: $\vdash_n$ is much weaker than validity $\models_n$ (which is the restriction of $\models$ to the formulas in $\mathcal{L}_n$).

Relation algebras are halfway between CA_3 and CA_4, the classes of 3-dimensional and 4-dimensional cylindric algebras, respectively. We sometimes jokingly say that RA is $\mathsf{CA}_{3.5}$. Why is RA stronger than CA_3? Because, the so-called relation-algebra-type reduct of a CA_3 is not necessarily an RA, e.g., associativity of relation composition can fail in the reduct. See [14, sec 5.3], and for more in this line see Németi-Simon [32]. Why is CA_4 stronger than RA? Because not every RA can be obtained, up to isomorphism, as the relation-algebra-type reduct of a CA_4, and consequently not every 4-variable sentence can be expressed as an RA-term. However, the same equations are true in RA and in the class of all relation-algebra-type reducts of CA_4's (Maddux's result, see [14, sec 5.3]). Thus Tarski formulated set theory, roughly, in CA_4, i.e., in $\mathcal{L}_4$ with $\vdash_4$, or in $\mathcal{L}_3$ with validity $\models_3$.

Németi [26], [27] improved this result by formalizing set theory in CA_3, i.e., in $\mathcal{L}_3$ with $\vdash_3$ in place of validity $\models_3$. The main idea for this improvement was using the pairing functions to store all data always, during every step of a proof, in one register only, so as to get two registers to work with in the proofs. In this approach one represents binary relations as unary ones (of ordered pairs).

First-order logic has equality as a built-in relation. One of the uses of equality in FOL is that it can be used to express (simulate) substitutions of variables, thus to "transfer" content of one variable to the other. The reduct SCA_3 of CA_3 "forgets" equality d_{ij} but retains substitution in the form of the term-definable operations s^i_j. The logic belonging to SCA_3 is weaker than 3-variable fragment of FOL. Zalán Gyenis [13] improved parts of Németi's result by extending them from CA_3 to SCA_3.

We get a much weaker logic by forgetting substitutions, too, this is the logic corresponding to Df_3 in which FOL and set theory were formalized in Andréka-Németi [4].

Three-dimensional diagonal-free cylindric algebras, Df_3's, are Boolean algebras with 3 commuting complemented closure operators, see [14, 1.1.2] or [17]. The logic $\mathcal{L}df_3$ corresponding to Df_3 has several intuitive forms, one is 3-variable equality- and substitution-free fragment of first-order logic with a rather weak proof system $\vdash^{df}$, another form of this same logic is modal logic $[S5, S5, S5]$, see [17] and [12]. Not only set theory but the whole of FOL is recaptured in $\mathcal{L}df_3$. This is a novelty w.r.t. previous results in this line. All the formalizability theorems mentioned above follow from this last result.

In section 1 we define our weak "target logic" $\mathcal{L}df_3$ and we state the existence of a structural translation mapping of FOL with countably many relation symbols of arbitrary ranks, $\mathcal{L}_\omega$, into $\mathcal{L}df_3$ with a single ternary relation symbol, see Theorem 1.6. If equality is available in our target logic, then we can do with one binary relation symbol, we do not need a ternary one, see Theorem 1.7. For theories in which a conjugated pair of quasi-pairing functions can be defined, such as most set theories, we can define a similar translation function which preserves meaning of formulas a bit more closely, see Theorem 1.6(ii), Theorem 1.7(ii). Theorem 1.6(ii) is a very strong version of Tarski's main result in [39, Theorem (xxxiv),p.122], which states roughly the same for the logic corresponding to RA in place of Df_3. After Theorem 1.7 we discuss the conditions in both Theorem 1.6 and Theorem 1.7, and we obtain that almost all of them are needed and that they cannot be substantially weakened.

In sections 2 and 3 we concentrate on the applications of the theorems stated in section 1. In section 2 we show that our translation functions are useful in proving properties for n-variable logics as well as for other "weak" logics. In particular, we prove a partial completeness theorem for the n-variable fragment of FOL ($n \geq 3$) and we prove that Gödel's incompleteness property holds for it.

In section 3 we review some results and problems on free cylindric-like algebras from the literature since 1985. As an application of the theorems in section 1, we show that the free cylindric algebras are not atomic (solution for [14, Problem 4.14]) and that these free algebras are not "wide", i.e., the $k+1$-generated free cylindric algebra cannot be embedded into the k-generated one, but these free algebras have many k-element irredundant non-free generator sets (solution for [14, Problem 2.7]).

1 Interpreting FOL in its small fragments

Instead of Df_3 and CA_3, we will work with fragments of FOL which are equivalent to them because this will be convenient when stating our theorems. We treat FOL as [14] does, i.e., with equality and with no operation symbols. We deviate from [14] in that our connectives are $\vee, \neg, \exists v_i, v_i = v_j,\ i, j \in \omega$ and we treat the rest as derived ones, by defining $\varphi \wedge \psi \stackrel{d}{=} \neg(\neg\varphi \vee \neg\psi)$, $\bot \stackrel{d}{=} (v_0 = v_0 \wedge \neg v_0 = v_0)$, $\top \stackrel{d}{=} \neg\bot$. We will use the derived connectives $\forall, \rightarrow, \leftrightarrow$, too, as abbreviations: $\forall v\varphi \stackrel{d}{=} \neg\exists v\neg\varphi$, $\varphi \rightarrow \psi \stackrel{d}{=} \neg\varphi \vee \psi$, $\varphi \leftrightarrow \psi \stackrel{d}{=} (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$. We will use x, y, z to denote the first three variables v_0, v_1, v_2. Sometimes we will write, e.g., $\exists xy$ or $\forall xyz$ in place of $\exists x\exists y$ or $\forall x\forall y\forall z$, respectively.

We begin with defining the fragment $\mathcal{L}df_3(P, 3)$ of FOL. It contains three variables and one ternary relation symbol P. It is a fragment of FOL in which we omit the equality, quantifiers $\exists v$ for v distinct from x, y, z, and atomic formulas $P(u, v, w)$ for $uvw \neq xyz$; and we omit all relation symbols distinct from P.

Definition 1.1. (3-variable restricted FOL without equality $\mathcal{L}df_3(P, 3)$)

(i) The *language* of our system contains one atomic formula, namely $P(x, y, z)$. (E.g., the formula $P(y, x, z)$ is not available in this language, this feature is what the adjective "*restricted*" refers to.) The logical connectives are $\vee, \neg, \exists x, \exists y, \exists z$. Thus, the set Fdf_3 of formulas of $\mathcal{L}df_3(P, 3)$ is the smallest set F containing $P(x, y, z)$ and such that $\varphi \vee \psi, \neg\varphi, \exists x\varphi, \exists y\varphi, \exists z\varphi \in F$ whenever $\varphi, \psi \in F$.

(ii) The *proof system* $\vdash^{df}$ which we will use is a Hilbert-style one with the following logical axiom schemes and rules.

The logical axiom schemes are the following. Let $\varphi, \psi \in Fdf_3$ and $v, w \in \{x, y, z\}$.
((1)) φ, if φ is a propositional tautology.
((2)) $\forall v(\varphi \to \psi) \to (\exists v\varphi \to \exists v\psi)$.
((3)) $\varphi \to \exists v\varphi$.
((4)) $\exists v\exists v\varphi \to \exists v\varphi$.
((5)) $\exists v(\varphi \vee \psi) \leftrightarrow (\exists v\varphi \vee \exists v\psi)$.
((6)) $\exists v\neg\exists v\varphi \to \neg\exists v\varphi$.
((7)) $\exists v\exists w\varphi \to \exists w\exists v\varphi$.
The inference rules are Modus Ponens ((MP), or detachment), and Generalization ((G)).

(iii) We define $\mathcal{L}df_3(P, 3)$ as the logic with formulas Fdf_3 and with proof system $\vdash^{df}$.

(iv) We define $\mathcal{L}df_n(\mathcal{R}, \rho)$ where n is an ordinal, $\mathcal{R}$ is a sequence of relation symbols and ρ is the sequence of their ranks (i.e., numbers of arguments), all $\leq n$, analogously to $\mathcal{L}df_3(P, 3)$. When we do not indicate $\mathcal{R}, \rho$ in $\mathcal{L}df_n(\mathcal{R}, \rho)$, we mean to have infinitely many n-place relation symbols. $\square$

The fragment $\mathcal{L}ca_3$ is similar to the above fragment $\mathcal{L}df_3$, except that we do not omit equality from the language, hence we will have $u = v$ as formulas for $u, v \in \{x, y, z\}$, and we will have two more axiom schemes concerning equality in the proof system. Since we have equality, our "smallest interesting" language will be when we have one binary relation symbol E.

Definition 1.2. (3-variable restricted FOL with equality $\mathcal{L}ca_3(E, 2)$)

(i) The *language* of our system contains one atomic formula, namely $E(x, y)$. The logical connectives are $\vee, \neg, \exists x, \exists y, \exists z$ together with $u = v$ for $u, v \in \{x, y, z\}$ as zero-place connectives. Thus, $x = x, x = y$, etc are formulas of $\mathcal{L}ca_3$. We denote the set of formulas (of $\mathcal{L}ca_3$) by Fca_3.

(ii) The *proof system* $\vdash^{ca}$ which we will use is a Hilbert-style one with the logical axiom schemes and rules of $\mathcal{L}df_3$ (understood as schemes for $\mathcal{L}ca_3$) extended with the following two axiom schemes:

Let $\varphi \in Fca_3$ and $u, v, w \in \{x, y, z\}$.
((8)) $(u = v \to v = u) \wedge (u = v \wedge v = w \to u = w) \wedge \exists v \;\; u = v$.
((9)) $u = v \wedge \exists v(u = v \wedge \varphi) \to \varphi$, where u, v are distinct.

(iii) We define $\mathcal{L}ca_3(E, 2)$ as the logic with formulas Fca_3 and with proof system $\vdash^{ca}$.

(iv) We define $\mathcal{L}ca_n(\mathcal{R}, \rho)$ where n is an ordinal, $\mathcal{R}$ is a sequence of relation symbols and ρ is the sequence of their ranks, all $\leq n$, analogously to $\mathcal{L}ca_3(E, 2)$. When we do not indicate $\mathcal{R}, \rho$ in $\mathcal{L}ca_n(\mathcal{R}, \rho)$, we mean to have infinitely many n-place relation symbols. $\square$

Remark 1.3. (On the fragment $\mathcal{L}df_3$ of FOL)

(i) The proof system $\vdash^{df}$ is a direct translation of the equational axiom system of Df_3. Axiom ((2)) is needed for ensuring that the equivalence relation defined on the formula algebra by $\varphi \equiv \psi \Leftrightarrow \vdash^{df} \varphi \leftrightarrow \psi$ be a congruence with respect to (w.r.t.) the operation $\exists v$. It is congruence w.r.t. the Boolean connectives $\vee, \neg$ by axiom ((1)). Axiom ((1)) expresses that the formula algebra factorized with $\equiv$ is a Boolean algebra, axiom ((5)) expresses that the quantifiers $\exists v$ are operators on this Boolean algebra (i.e., they distribute over $\vee$), axioms ((3)),((4)) express that these quantifiers are closure operations, axiom ((6)) expresses that they are complemented closure operators (i.e., the negation of a closed element is closed again). Together with ((5)) they imply that the closed elements form a Boolean subalgebra, and hence the quantifiers are normal operators (i.e., the Boolean zero is a closed element). Finally, axiom ((7)) expresses that the quantifiers commute with each other. We note that ((1)) is not an axiom scheme in the sense of [6] since it is not a formula scheme, but it can be replaced with three formula schemes, see [14, Problem 1.1] (solved in [23]).

(ii) The logic $\mathcal{L}df_3$ corresponds to Df_3 in the sense of [14, sec.4.3], as follows. What is said in (i) above immediately implies that the proof-theoretic (Lindenbaum-Tarski) formula algebra of $\mathcal{L}df_3$ (which is just the natural formula-algebra factorized by the equivalence relation $\equiv$ defined in (i) above) is the infinitely generated Df_3-free algebra, and that of $\mathcal{L}df_3(P,3)$ is the one-generated Df_3-free algebra. Moreover, valid formulas of $\mathcal{L}df_3$ correspond to equations valid in Df_3, namely we claim that $\mathcal{L}df_3 \vdash^{df} \varphi \Longleftrightarrow \mathsf{Df}_3 \models \tau\mu(\varphi) = 1$ for all $\varphi \in Fdf_3$, where $\tau\mu(\varphi)$ is as defined in [14, 4.3.55].

(iii) The logic $\mathcal{L}df_3$ inherits a natural semantics from first-order logic (namely Mod, the class of models of FOL, and $\models_3$, the validity relation restricted to 3-variable formulas). The proof system $\vdash^{df}$ is strongly sound with respect to this semantics, but it is not complete, for more on this see section 2. We note that just as Df_3 corresponds to the logic $\mathcal{L}df_3 = \langle Fdf_3, \vdash^{df} \rangle$, the class of algebras corresponding to $\langle Fdf_3, \models_3\rangle$ is the class RDf_3 of representable diagonal-free cylindric algebras ([14, 5.1.33(v)]). For more on connections between logics and classes of algebras, besides [14, sec.4.3], see [6], or [33].

The expressive power of $\mathcal{L}df_3$ is seemingly very small. It's not only that "we cannot count" due to lack of the equality, we cannot transfer any information from one variable to the other by the use of the equality, so all such transfer must go through an atomic formula. Hence if we have only binary relation symbols, in the restricted language there is just no way of meaningfully using the third variable z, and we basically have two-variable logic which is decidable. However, Theorem 1.6 below says that if we have at least one ternary relation symbol and we are willing to express formulas in a more complicated way (than the most natural one would be), then we can express any sentence that we can in FOL.

(iv) In the present paper we will use $\mathcal{L}df_3$ as introduced above because it will be convenient to consider it a fragment of FOL. However, $\mathcal{L}df_3$ has several different but equivalent forms, each of which has advantages and disadvantages. Some of the different forms are reviewed in [4, sec.2]. We mention two of the equivalent forms. One is modal logic $[S5, S5, S5]$, this is equivalent to $\mathcal{L}df_3$ (while the modal logic $S5 \times S5 \times S5$ introduced in [17] is equivalent with $\langle Fdf_3, \models_3\rangle$), see [12, p.379]. The other equivalent form is just equational logic with the defining equations of Df_3 as extra axioms. $\square$

Remark 1.4. (On the fragment $\mathcal{L}ca_3$ of FOL)

(i) The logic $\mathcal{L}ca_3$ corresponds to CA_3 just the way $\mathcal{L}df_3$ corresponds to Df_3. The proof system $\stackrel{ca}{\vdash}$ is a direct translation of the equational axiom system of CA_3. ((8)) expresses that $=$ is an equivalence relation and ((9)) expresses that formulas do not distinguish equivalent (equal) elements. Take the Hilbert-style proof system with axiom schemes ((1))-((9)) and rules as (MP) and (G). Add the axioms

((0)) $R(v_{i1}, ..., v_{in}) \leftrightarrow \exists v_j R(v_{i1}, ..., v_{in})$ for R an n-place relation symbol and $j \notin \{i1, ..., in\}$.

Then the so obtained proof system is complete for FOL (with usual semantics Mod, $\models$). Hence, $\mathcal{L}df_3$ and $\mathcal{L}ca_3$ are "proof-theoretic" fragments of FOL when taking this complete proof system for FOL.

(ii) The expressive power of $\mathcal{L}ca_3$ is much greater than that of $\mathcal{L}df_3$, due to the presence of equality. E.g., one can express that a binary relation is actually a function, one can express composition of binary relations, one can express (simulate) substitution of variables. However, the proof system $\stackrel{ca}{\vdash}$ is still very weak, e.g., one can express but cannot prove the following: the composition of two functions is a function again, composition of binary relations is associative, converse of the converse of a binary relation is the original one, interchanging the variables x, y in two different ways by using z as "auxiliary register" results in an equivalent formula (this is the famous Merry Go Round equation [14, 3.2.88], see also [11], [36]). More precisely, one cannot prove these statements if one expresses them the most natural ways. Our theorems below say that if we express the same statements in more involved ways, they become $\stackrel{ca}{\vdash}$-provable. $\square$

Let $\mathcal{L}_\omega$ denote usual FOL with countably many variables and with countably many relation symbols for each rank, i.e., we have countably many n-place relation symbols for all positive n. Let L_ω denote the set of formulas of $\mathcal{L}_\omega$. Thus $\mathcal{L}_\omega = \langle L_\omega, \vdash\rangle$ where $\vdash$ is either the proof system outlined in Remark 1.4(i) above, or just the usual semantic consequence relation $\models$. We assume that E is a binary and P is a ternary relation symbol in $\mathcal{L}_\omega$. Then $L_\omega(E, 2)$ denotes the set of formulas in $\mathcal{L}_\omega$ in which only E occurs from the relation symbols. Zermelo-Fraenkel set theory written up in $L_\omega(E, 2)$ is denoted by ZF. A formula of $\mathcal{L}_\omega$ is called a *sentence* if it does not contain free variables.

Definition 1.5. (Structural translations) Let $\mathcal{L} = \langle F, \vdash\rangle$ be a logic (in the sense of Remark 1.8(i) below). Assume that $\vee, \neg$ are connectives in $\mathcal{L}$, and let $\rightarrow$ denote the corresponding derived connective in $\mathcal{L}$, too. Let $f : L_\omega \rightarrow F$ be an arbitrary function. We say that f is *structural* iff the following (i)-(ii) hold for all sentences $\varphi, \psi \in L_\omega$.

(i) $\vdash f(\varphi \vee \psi) \leftrightarrow [f(\varphi) \vee f(\psi)]$,

(ii) $\vdash f(\varphi \rightarrow \psi) \rightarrow [f(\varphi) \rightarrow f(\psi)]$. $\square$

The following is proved in Andréka-Németi [4] together with [5], by defining concrete translations tr. For the role of $\neg\mathsf{tr}(\bot)$ in (ii) below see Remark 1.8(iii),(iv).

Theorem 1.6. (Formalizability of FOL in $\mathcal{L}df_3$)

(i) There is a structural computable translation function $\mathsf{tr} : \mathcal{L}_\omega \longrightarrow \mathcal{L}df_3(P,3)$ *such that* tr *has a decidable range and the following (a),(b) are true for all sets of sentences* $Th \cup \{\varphi\}$ *in* $\mathcal{L}_\omega$*:*

(a) $Th \models \varphi$ *iff* $\mathsf{tr}(Th) \mid\!\frac{df}{}\ \mathsf{tr}(\varphi)$.

(b) $Th \models \varphi$ *iff* $\mathsf{tr}(Th) \models \mathsf{tr}(\varphi)$.

(ii) There is a structural computable translation function $\mathsf{tr} : \mathcal{L}_\omega(E,2) \longrightarrow \mathcal{L}df_3(P,3)$ *such that* tr *has a decidable range and the following (c),(d) are true, where* Δ *denotes the set of the following two formulas:*

$E(x,y) \longleftrightarrow \forall z P(x,y,z) \wedge \exists xyz[P(x,y,z) \wedge \neg\forall z P(x,y,z)]$,

$x = y = z \longleftrightarrow P(x,y,z) \wedge \neg E(x,y)$.

(c) Statements (a) and (b) in (i) above hold for all sets of sentences $Th \cup \{\varphi\}$ *in* $\mathcal{L}_\omega(E,2)$ *such that* $Th \cup \Delta \models \neg\mathsf{tr}(\bot)$*. Further,* $ZF \cup \Delta \models \neg\mathsf{tr}(\bot)$.

(d) $\Delta \cup \{\neg\mathsf{tr}(\bot)\} \models \varphi \leftrightarrow \mathsf{tr}(\varphi)$ *for all sentences* $\varphi \in L_\omega(E,2)$. $\square$

The following is proved in Németi [26], [27] (taken together with [5]), by constructing concrete tr's. It says that we can replace the ternary relation symbol P with a binary one in Theorem 1.6 if we have equality.

Theorem 1.7. (Formalizability of FOL in $\mathcal{L}ca_3$)

(i) There is a computable, structural translation function $\mathsf{tr} : \mathcal{L}_\omega \longrightarrow \mathcal{L}ca_3(E,2)$ *such that* tr *has a decidable range and the following (a),(b) are true for all sets of sentences* $Th \cup \{\varphi\}$ *in* $\mathcal{L}_\omega$*:*

(a) $Th \models \varphi$ *iff* $\mathsf{tr}(Th) \mid\!\frac{ca}{}\ \mathsf{tr}(\varphi)$.

(b) $Th \models \varphi$ *iff* $\mathsf{tr}(Th) \models \mathsf{tr}(\varphi)$.

(ii) There is a computable, structural translation function $\mathsf{tr} : \mathcal{L}_\omega(E,2) \longrightarrow \mathcal{L}ca_3(E,2)$ *such that* tr *has a decidable range and the following (c),(d) are true:*

(c) Statements (a) and (b) in (i) above hold for all sets of sentences $Th \cup \{\varphi\}$ *in* $\mathcal{L}_\omega(E,2)$ *such that* $Th \models \neg\mathsf{tr}(\bot)$*. Further,* $ZF \models \neg\mathsf{tr}(\bot)$.

(d) $\neg\mathsf{tr}(\bot) \models \varphi \leftrightarrow \mathsf{tr}(\varphi)$*, for all sentences* $\varphi \in L_\omega(E,2)$. $\square$

On the proof of Theorem 1.7: The most difficult part of this theorem is proving $\models \varphi \Rightarrow \;\vdash^{ca} \mathsf{tr}(\varphi)$, for all $\varphi \in L_\omega$, therefore we will outline the ideas for proving this part. So, we want to prove a kind of completeness theorem for $\vdash^{ca}$.

Formulas $\varphi(x, y)$ with two free variables x, y represent binary relations and then the natural way of expressing relation composition of binary relations is the following:

$$(\varphi \circ \psi)(x, y) \;\stackrel{d}{=}\; \exists z(\varphi(x, z) \wedge \psi(z, y)), \quad \text{where}$$

$$\varphi(x, z) \;\stackrel{d}{=}\; \exists y(y = z \wedge \varphi(x, y)) \quad \text{and} \quad \psi(z, y) \;\stackrel{d}{=}\; \exists x(x = z \wedge \psi(x, y)).$$

Now, assume that we have two unary *partial* functions, p, q which form pairing functions, i.e. for which the following formula π holds:

$$\pi \;\stackrel{d}{=}\; \forall xy\exists z(\mathsf{p}(z) = x \wedge \mathsf{q}(z) = y).$$

For supporting intuition, let us write $z_0 = x$ and $z_1 = y$ in place of $\mathsf{p}(z) = x$ and $\mathsf{q}(z) = y$, and let $\langle x, y\rangle$ denote an arbitrary z for which $z_0 = x$ and $z_1 = y$. Now, we can "code" binary relations as unary ones, i.e., if $\varphi(x)$ is a formula with one free variable x, then we can think of it as representing the binary relation $\{\langle x_0, x_1\rangle : \varphi(x)\}$. With this in mind then a natural way of representing relation composition is the following

$$(\varphi \odot \psi)(x) \;\stackrel{d}{=}\; \exists y(\varphi(y_0) \wedge \psi(y_1) \wedge x_0 = y_{00} \wedge y_{01} = y_{10} \wedge y_{11} = x_1), \quad \text{see Figure 1.}$$

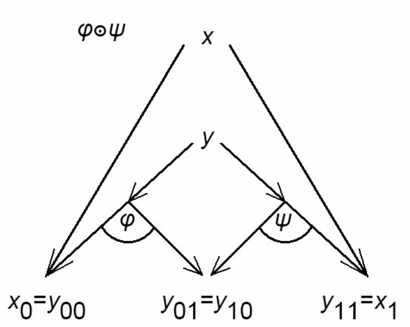

Figure 1: Illustration of $\varphi \odot \psi$

As we said in Remark 1.4(ii), associativity of $\circ$ cannot be proved by $\vdash^{ca}$, i.e., there are formulas φ, ψ, η in Fca_3 such that

$$\not\vdash^{ca} ((\varphi \circ \psi) \circ \eta)(x, y) \leftrightarrow (\varphi \circ (\psi \circ \eta))(x, y).$$

However, associativity of relation composition expressed in the unary form can be proved, by assuming a formula $\mathsf{Ax} \in Fca_3$ which is semantically equivalent with π but proof-theoretically stronger:

$\mathsf{Ax} \models^{ca} ((\varphi \odot \psi) \odot \eta)(x) \leftrightarrow (\varphi \odot (\psi \odot \eta))(x)$

for all formulas φ, ψ, η with one free variable x such that $\models^{ca} \varphi(x) \to \mathsf{pair}(x)$, etc, where $\mathsf{pair}(x) \stackrel{d}{=} \exists y x_0 = y \wedge \exists y x_1 = y$. (I.e., $\mathsf{pair}(x)$ holds for x if both p and q are defined on it.) We note that π is not strong enough for proving associativity of $\odot$, and even Ax is not strong enough for proving associativity of $\circ$, see [27, 15T(ii),(iv)]. We mentioned already that $\models^{ca}$ cannot prove that composition of functions is a function again. Roughly, Ax is π together with stating that composition of at most three "copies" of p, q is a function again (i.e., $\mathsf{p} \circ \mathsf{p} \circ \mathsf{q}$, $\mathsf{p} \circ \mathsf{q}$ etc are all functions). Similarly to the above, we can express converse of binary relations and the identity relation (coded for their unary form) and prove for these by $\models^{ca}$ all the relation algebraic equations, from Ax of course. Thus we defined relation-algebra-type operations on the set of formulas of form $\varphi(x) \wedge \mathsf{pair}(x)$, and we can prove from Ax that these operations form an RA. If p, q can be expressed as above, then we have a so-called quasi-projective RA, a QRA, which are representable by [39, 8.4(iii)], and we know that representation theorems help us to get provability from validity (i.e., the hard direction of completeness theorems). It remains to get suitable pairing formulas p, q (see (1) below) and to translate all FOL-formulas, in a meaning-preserving way, to the above QRA-fragment of $\mathcal{L}ca_3$ (see (2) below).

(1) We can get p, q by "brute force": we add a new binary relation symbol E to our language, intuitively we will think of it as the element-of relation $\in$. Then we express ordered pairs the way usually done in set theory (i.e., $\langle x, y \rangle \stackrel{d}{=} \{\{x\}, \{x, y\}\}$), and realize that we can write up the two projection functions belonging to these using only three variables. By using these projection functions we can convert every FOL-formula to one in Fca_3 so that we preserve validity (we can use the pairing-technique to code all the relations into one binary one, and then we can code up all the variables into the first three ones). This part is not so difficult because we may think "semantically".

(2) It remains now to translate all 3-variable formulas $\varphi \in Fca_3$ into the QRA-fragment of $\mathcal{L}ca_3$ we obtained above. The paper Simon [34] comes to our aid. In [34], to every QRA a CA_3-type subreduct is defined which is representable, i.e., which is in RCA_3. Let $\mathfrak{C}$ be this subreduct of our above-defined QRA, then the universe of $\mathfrak{C}$ is a subset of Fca_3 and the operations of $\mathfrak{C}$ are defined in terms of formulas of Fca_3, too. Let $f : Fca_3 \to \mathfrak{C}$ be a homomorphism (where the CA_3-type operations of Fca_3 are the natural ones), and then we define $\mathsf{tr} : Fca_3 \to Fca_3$ by $\mathsf{tr}(\varphi) \stackrel{d}{=} \mathsf{Ax} \to f(\varphi)$. Now, one can check that $\models \varphi \Rightarrow \models^{ca} \mathsf{tr}(\varphi)$. For the details of the proof outlined above and for the proof of Theorem 1.6 we refer the reader to [4, 5, 26, 27].

Remark 1.8. (Discussion of the conditions in Theorems 1.6,1.7)

(i) In Abstract Algebraic Logic, AAL, and/or in Universal Logic the key concept is a logical system (logic in short) $\langle F, \vdash \rangle$ where F is a set (thought of as the set of formulas) and $\vdash \subseteq \mathsf{Sb}(F) \times F$ (thought of as a consequence relation), where $\mathsf{Sb}(F)$ denotes the powerset of F. If $f : F \to F'$ is a function between two logics $\mathcal{L} = \langle F, \vdash\rangle$ and $\mathcal{L}' = \langle F', \vdash'\rangle$, then f is called a *translation* iff it preserves $\vdash$, i.e., iff $Th \vdash \varphi \Rightarrow f(Th) \vdash' f(\varphi)$ holds for all $Th \cup \{\varphi\} \subseteq F$, and f is called a *conservative translation* if $\Rightarrow$ can be replaced by $\Leftrightarrow$ in the above. Jeřábek [15] proved that FOL can be conservatively translated even to classical propositional calculus (CPC); and moreover, every countable logic can be conservatively translated to CPC. In this sense, the existence of conservative translations does not mean much in itself. However, if we require the translation to be computable in addition, then undecidability is preserved along the translation, and so FOL can be translated to undecidable logics only (i.e., where $\{\varphi \in F : \emptyset \vdash \varphi\}$ is undecidable), and so it cannot be translated to CPC. For this reason, the conditions that we have at least one at least ternary relation symbol, we have at least 3 quantifiers (closure operators), and that they commute are all necessary conditions for our target logic in Theorem 1.6 since without these conditions we get decidable logics. (We have seen that $\mathcal{L}df_3(P,2)$ is basically 2-variable logic which is decidable [14, 4.2.7], and the logic we get from $\mathcal{L}df_n$ by omitting the axiom scheme ((7)) requiring that the quantifiers commute is proved to be decidable in [27, Chap.III], [29, Theorem 1.1]). If we require more properties for the translation function to hold, then more properties are preserved along them. E.g., structural computable conservative translations preserve Gödel's incompleteness property from one logic to the other, see Theorem 2.4.

(ii) The achievement (of Theorems 1.6,1.7) that the range of the translation is decidable can be omitted, since if we have a translation function then by using the trick in [10] we can modify this function so that its range becomes decidable and keep all the other good properties, at least in our case when our logics are extensions of CPC.

(iii) There are sentences in $L_\omega(E,2)$ which are not equivalent semantically to any formula in $L_3(E,2)$, hence there is no function $f : L_\omega \to L_3(E,2)$ for which $\models \varphi \leftrightarrow f(\varphi)$ would hold for all sentences $\varphi \in L_\omega$. For this reason, $\neg\mathsf{tr}(\bot)$ cannot be omitted in Theorems 1.6(ii)(d),1.7(ii)(d). For example, such a 4-variable sentence is exhibited in [27, p.39]. We note that Δ in Theorem 1.6(ii) is an explicit "definition" of E and $=$ from P.

(iv) Our translation functions are not Boolean homomorphisms in general, e.g., the translations tr we define in the proofs of Theorems 1.6,1.7 do not preserve negation in the way they preserve disjunction. Consequently, $\neg\mathsf{tr}(\bot)$ is not the same as $\mathsf{tr}(\top)$, and more importantly, $\neg\mathsf{tr}(\bot)$ is not a valid formula. From the fact that tr is structural, it can be proved that $\vdash \mathsf{tr}(\varphi) \leftrightarrow [\neg\mathsf{tr}(\bot) \to \mathsf{tr}(\varphi)]$. Hence, $\neg\mathsf{tr}(\bot)$ seems to be the weakest assumption under which one can expect semantical equivalence of φ with $\mathsf{tr}(\varphi)$. Intuitively, $\neg\mathsf{tr}(\bot)$ is the "background knowledge" we assume for the translation function tr to preserve meaning. This is the role of $\neg\mathsf{tr}(\bot)$ in Theorems 1.6(ii),1.7(ii).

(v) A logic $\langle F, \vdash\rangle$ is defined to be a *propositional logic* (or sentential logic) in AAL if F is built up from some set, called propositional variables, by using connectives and $\vdash$ is *substitutional*, i.e., $\vdash$ is preserved by substitution of arbitrary formulas for propositional variables. In this sense, $\mathcal{L}df_3$ and $\mathcal{L}ca_3$ are propositional logics, but $\mathcal{L}_\omega$ is not, see e.g., [6] or [33].

(vi) Any logic $\mathcal{L} = \langle F, \vdash\rangle$ which is between $\mathcal{L}df_3(P,3)$ and $\mathcal{L}_\omega$ can be taken in

Theorem 1.6 in place of $\mathcal{L}df_3(P, 3)$. (We say that a logic $\mathcal{L} = \langle F, \vdash \rangle$ is contained in another one, $\mathcal{L}' = \langle F', \vdash' \rangle$, if $F \subseteq F'$ and $\vdash \subseteq \vdash'$.) This is easy to check. $\square$

Problem 1.9. (Interpreting FOL in weaker fragments) *Can the requirement of the closure operators being complemented be omitted in our theorems? I.e., is there a computable (structural) conservative translation from $\mathcal{L}_\omega$ to the equational theory EqBf_3 of Bf_3 where Bf_3 denotes the class of all Boolean algebras with three commuting (not necessarily complemented) closure operators? Is EqBf_3 undecidable?*

2 Applications of the interpretation

There are many applications, of different flavors, of the interpretability theorems of which Theorem 1.6 is presently the strongest one. In this and the next sections we state some of these applications. We will concentrate on consequences for cylindric algebras, $\mathsf{CA}_n, \mathsf{RCA}_n$ and their logical counterparts, but analogous results hold for all their variants, e.g., for diagonal-free cylindric algebras $\mathsf{Df}_n, \mathsf{RDf}_n$, substitution-cylindrification algebras SCA_n, RSCA_n, polyadic equality algebras PEA_n, RPEA_n, polyadic algebras PA_n, RPA_n, for relation algebras SA, RA, RRA and their logical counterparts. For the definition of these classes of algebras see, e.g., [14], [19], [22].

The first applications we talk about here concern completeness theorems. Tarski used his translation in [39] to transfer the completeness theorem for $\mathcal{L}_\omega$ into a kind of completeness theorem for his target logic, which in algebraic form is stated as a representation theorem, namely that every quasi-projective relation algebra is representable. (Later Maddux [20] gave a purely algebraic proof for this.) It is shown in [26, 3.7-3.10], [27, 17T(viii)] that RA cannot be replaced with CA_3 in this consequence, namely, quasi-projective CA_3's are rather far from being representable (and the same is true for the class SA of semi-associative relation algebras, in place of CA_3). So, in this respect, Tarski's result cannot be improved.

Yet, we can use our translations in Theorems 1.6,1.7 to prove completeness results for our target logics, but in a different way. We begin with recalling some definitions from [6, D.33, D.48]. A proof system is called *Hilbert-style* if it is given by finitely many axiom schemes and rules where the rules are of form $\varphi_1, ..., \varphi_k \vdash \varphi_0$ for some formula schemes $\varphi_0, ..., \varphi_k$. A proof system $\vdash$ is called *sound* w.r.t. the semantics $\models$ iff $\vdash \varphi$ implies $\models \varphi$, *strongly sound* if $Th \vdash \varphi$ implies $Th \models \varphi$, *complete*, *strongly complete* when "implies" is replaced with "implied by" in the above, for all sets $Th \cup \{\varphi\}$ of formulas. Finally, we define $\mathcal{L}_n = \langle L_n, \vdash_n \rangle$, the usual n-variable fragment of $\mathcal{L}_\omega$, as restricting $\mathcal{L}_\omega$ to those formulas of L_ω which contain only the first n variables. (E.g., $R(y, x, x, z) \in L_3$ when R is a 4-place relation symbol in $\mathcal{L}_\omega$.) More precisely, $\vdash_n$ is the provability relation we get from the axiom schemes ((0))-((9)) understood as schemes for L_n and rules (MP),(G), cf. Remark 1.4(i). Throughout this section, we assume that

E is a binary relation symbol in $\mathcal{L}_\omega$ and $n \geq 3$ is finite. Hence, in Theorem 1.7 we can replace $\mathcal{L}ca_3(E, 2)$ with $\mathcal{L}_n$ (see Remark 1.8(vi)).

We know that $\mathcal{L}_n$ is inherently incomplete, i.e., there is no complete and strongly sound Hilbert-style proof system for the "standard" validity $\models$ restricted to L_n. In the literature, there are approaches aimed at getting around this inherent incompleteness of $\mathcal{L}_n$. One goes by replacing "standard" models and validity with "nonstandard" models and validity which one can obtain from CA_n. This approach originates with Leon Henkin. The other approach is keeping the standard semantics and using new complete inference systems which are sound but not strongly sound. Such inference systems are introduced, e.g., in Venema [40], [41] and in Simon [35]. Problem 7.2 in [33], as well as [14, Problem 4.16], and [3, Problem 1(a) (p.730), Problems 49,50 (p.740)] are strongly related to this direction.

Let $\vdash_{\overline{nt}}$ denote the proof system we obtain from $\vdash_{\overline{n}}$ by adding the rule which infers φ from $\mathsf{tr}(\varphi)$ when φ is a sentence in $\mathcal{L}_n$, where tr is the translation in Theorem 1.7(i). This last rule is sound but not strongly sound, i.e., $\vdash_{\overline{nt}} \varphi$ implies $\models \varphi$, but it is not true that $Th \vdash_{\overline{nt}} \varphi$ implies $Th \models \varphi$ (namely, $\mathsf{tr}(\varphi) \vdash_{\overline{nt}} \varphi$ for all φ, but $\mathsf{tr}(\varphi) \models \varphi$ is not true for all φ).

Our first theorem in this section is an immediate corollary of Theorem 1.7(i). It says that the "standard" Hilbert-style proof system $\vdash_{\overline{n}}$ is strongly complete and strongly sound within a large enough subset of L_n; and the "nonstandard" proof system $\vdash_{\overline{nt}}$ is complete and sound for the whole of $\mathcal{L}_n$.

In more detail, the first part of Corollary 2.1 below says that we can select a subset G of formulas, call it the set of "formulas of good shape", such that the natural Hilbert-style proof system $\vdash_{\overline{n}}$ is strongly complete within this subset; moreover we can decide whether a formula is in good shape, and every formula φ can be algorithmically converted to one in a good-shape such that meaning is preserved in the sense described in Theorem 1.7(i).

Corollary 2.1. *Let G denote the range of* tr *in Theorem 1.7(i) and let $n > 2$ be finite. Then (i)-(iii) below hold:*

(i) $\vdash_{\overline{n}}$ *is strongly complete within $G \subseteq L_n$, i.e., for all $Th \cup \{\varphi\} \subseteq G$ we have that $Th \models \varphi \iff Th \vdash_{\overline{n}} \varphi$.*

(ii) *G is large enough in the sense that $\mathsf{tr}(\varphi) \in G$ for all $\varphi \in L_n$ and $\models \varphi \iff \models \mathsf{tr}(\varphi)$.*

(iii) $\vdash_{\overline{nt}}$ *is complete and sound in the whole of L_n, i.e., for all formulas $\varphi \in L_n$ we have $\models \varphi \iff \vdash_{\overline{nt}} \varphi$.* $\square$

The next corollary concerns connections between RCA_n and its finitary approximation, CA_n. Let $\otimes$ denote the complement of the symmetric difference, i.e., $x \otimes y \stackrel{d}{=} (x \cdot y) + (-x \cdot -y)$. We note that $\otimes$ is the algebraic counterpart of $\leftrightarrow$.

Corollary 2.2. *There is a computable function f mapping* CA_n*-terms to* CA_n*-terms such that for all* CA_n*-terms* τ, σ *we have*

(i) $\mathsf{RCA}_n \models \tau = \sigma$ *iff* $\mathsf{CA}_n \models f(\tau \otimes \sigma) = 1$, *or, in an equivalent form*

(ii) $\mathsf{RCA}_n \models \tau = 1$ *iff* $\mathsf{CA}_n \models f(\tau) = 1$. □

The above corollary of Theorem 1.7 justifies, in a way, the introduction of CA_n. Namely, CA_n was devised in order to "control", have a firm grasp on equations true in RCA_n. Nonfinite axiomatizability of RCA_n implies that this firm grasp cannot be attained in the form of $\mathsf{EqCA}_n = \mathsf{EqRCA}_n$ where EqK denotes the equational theory of the class K of algebras. By contrast, the above theorem says that a firm grasp can be obtained by using the computable function f; the axioms of CA_n together with the definition of f provide a finitary tool that captures (reconstructs completely) EqRCA_n.

A corollary of Theorem 1.6 says that the computational complexity of FOL is the same as that of the equational theory of Df_3. We recall from [9], informally, that the Turing-degree of $S \subseteq \omega$ is less than or equal to that of $Z \subseteq \omega$, in symbols $S \leq_T Z$, if by using a decision procedure for Z we can decide S. The Turing-degrees of S and Z are the same, in symbols $S \equiv_T Z$, if $S \leq_T Z$ and $Z \leq_T S$. The same notion can be applied to the equational theories of various classes of algebras, and to various FOL-theories. Let $Th(\emptyset)$ denote the set of valid formulas of $\mathcal{L}_\omega$.

The following corollary says that if we have a decision procedure for any one of $Th(\emptyset)$, EqK with K one of Df_n, RDf_n, CA_n, RCA_n, SCA_n, RSCA_n, ..., RA, RRA, $3 \leq n < \omega$ then we can e any other of the same list. In short, the Turing-degrees of all these classes are the same. This corollary follows from Theorems 1.6,1.7.

Corollary 2.3. *Let $3 \leq n < \omega$ and let* K *be any one of* Df_n, RDf_n, CA_n, RCA_n, SCA_n, RSCA_n, PA_n, RPA_n, PEA_n, RPEA_n, SA, RA, RRA*. Then (i) and (ii) below hold.*

(i) $\mathsf{EqDf}_3 \equiv_T \mathsf{EqK}$.

(ii) $\mathsf{EqDf}_3 \equiv_T Th(\emptyset)$. □

The above applications are all relevant to Problem 4.1 of [14]. Indeed, from Corollary 2.2 we can get a decidable equational base for EqRCA_n similar to that in [14, 4.1.9], and $\vdash_{\overline{nt}}$ gives a kind of solution for [14, Problem 4.1] similar to [14, 4.1.20].

Now we turn to other kinds of applications. Tarski introduced and used translation functions from a logic $\mathcal{L}$ into a logic $\mathcal{L}'$ in order to transfer some properties of $\mathcal{L}$ to $\mathcal{L}'$. For example, if the translation function is computable, then undecidability of the valid formulas of $\mathcal{L}$ implies the same for $\mathcal{L}'$. This is how Tarski proved that EqRA was undecidable. The same way, Theorem 1.6 immediately implies that the sets of validities of $\mathcal{L}df_n, \mathcal{L}ca_n$ as well as the equational theories of $\mathsf{Df}_n, \mathsf{CA}_n$ for $n \geq 3$ are undecidable. These have been known and have been proved by using Tarski's translation of set theory into RA for $n \geq 4$, and for $n = 3$ it is a result of Maddux [21], proved by an algebraic method.

The next theorem says that structural computable translations are capable of transferring Gödel's incompleteness property. For $\mathcal{L}ca_3$ this is proved in [26, Theorem 1.6], and for $\mathcal{L}df_3$ it is proved in [4, Theorem 2.3].

Theorem 2.4. (Gödel-style incompleteness theorem for $\mathcal{L}df_3$) *There is a formula $\varphi \in Fdf_3$ such that no consistent decidable extension T of φ is complete, and moreover, no decidable extension of φ separates the $\models^{df}$-consequences of φ from the φ-refutable sentences (where ψ is φ-refutable iff $\varphi \models^{df} \neg\psi$). The same is true for $\mathcal{L}ca_3$ and $\mathcal{L}_n$ in place of $\mathcal{L}df_3$.*

The proof of Theorem 2.4 goes by showing that the translation of an inseparable formula which is consistent with $\neg\mathsf{tr}(\bot)$ by a structural computable translation function tr is inseparable again.

In algebraic logic, the algebraic property corresponding to the logical property of Gödel's incompleteness is atomicity of free algebras (see [14, 4.3.32] and [26, Proposition 1.8]). Indeed, Theorem 2.4 above implies non-atomicity of free cylindric algebras, this way providing a solution for [14, Problem 4.14]. We devote the next section entirely to free cylindric algebras, because of their importance.

3 Structure of free cylindric algebras

In general, the free algebras of a variety are important because they show, in a sense, the structure of the different "concepts" (represented by terms) of the variety. In algebraic logic, the free algebras of a variety corresponding to a logic $\mathcal{L}$ are even more important, because they correspond to the so-called Lindenbaum-Tarski algebras of $\mathcal{L}$. This implies that the structures of free cylindric algebras are quite rich, since these reflect the whole of FOL, in a sense. Thus proving properties about free cylindric algebras is not easy in general. Often, one proves properties of free algebras by applying logical results to algebras, and it is a task then to find purely algebraic proofs, too. Chapter 2.5 of [14] is devoted to free cylindric algebras. Most of what we say here about free cylindric algebras generalizes to its variants such as Df_n, RDf_n, PA_n,... by using Theorem 1.6 in place of Theorem 1.7 and keeping Remark 1.8(vi) in mind.

Atoms in the Lindenbaum-Tarski algebras of sentences correspond to finitely axiomatizable complete theories, while atoms in the Lindenbaum-Tarski algebras of formulas with n free variables of FOL are related to the Omitting Type theorems and prime models, see [8, sec.2.3].

$\mathfrak{Fr}_k\mathsf{K}$ denotes the k-generated K-free algebra, see [14, 0.4.19]. We usually assume $k \neq 0$, just for simplicity. $\mathfrak{Fr}_k\mathsf{CA}_n$ is atomless if k is infinite (Pigozzi, [14, 2.5.13]). Assume k is finite, nonzero. If $n < 2$ then $\mathfrak{Fr}_k\mathsf{CA}_n$ is finite ([14, 2.5.3(i)]), hence atomic. $\mathfrak{Fr}_k\mathsf{CA}_2$ is infinite but still atomic (Henkin, [14, 2.5.3(ii), 2.5.7(ii)]). If $2 \leq n < \omega$ then $\mathfrak{Fr}_k\mathsf{CA}_n$ has infinitely many atoms

(Tarski, [14, 2.5.9]), and it was asked in [14] as Problem 4.14 whether it is atomic or not. The following solution is proved in [26], [27].

Theorem 3.1. *Let $0 < k < \omega$ and $n \geq 3$. Then $\mathfrak{F}r_k\mathsf{CA}_n$ is not atomic.* $\square$

A proof of the above theorem is based on Theorem 2.4. This is a metalogical proof, "transferring" Gödel's incompleteness theorem for FOL to three-variable logic. [14, Problem 4.14] also raised the problem of finding purely algebraic proofs for these properties of free algebras. Németi [25] contains direct, purely algebraic proofs showing that $\mathfrak{F}r_k\mathsf{CA}_n$ is not atomic, for $n \geq 4$. However, those proofs do not work for $n = 3$ (counterexamples show that the crucial lemmas fail for $n = 3$), and they are longer than the present metalogical proof.

So, in particular, it remains open to find a direct, algebraic proof for non-atomicity of $\mathfrak{F}r_k\mathsf{Df}_3$, $k > 0$.

In the proof of Theorem 3.1 we had to show that there is an element in the free algebra below which there is no atom. Problem 2.5 in [14], still open, asks if the sum of all atoms in $\mathfrak{F}r_k\mathsf{CA}_n$ exists for finite k and $3 \leq n < \omega$. This problem is equivalent to asking if there is a biggest element in the free algebra below which there is no atom.

Remark 3.2. (On zero-dimensional atoms)

$\mathfrak{F}r_k\mathsf{CA}_n$ has exactly 2^k zero-dimensional atoms (Pigozzi, [14, 2.5.11]). It was conjectured that these are all the atoms if $n \geq \omega$ (see [14, 2.5.12, Problem 2.6]). We note that there may be many more atoms in $\mathfrak{Z}d\mathfrak{F}r_k\mathsf{CA}_n$, the zero-dimensional part of $\mathfrak{F}r_k\mathsf{CA}_n$, than the zero-dimensional atoms of $\mathfrak{F}r_k\mathsf{CA}_n$. I.e., the atoms of $\mathfrak{Z}d\mathfrak{F}r_k\mathsf{CA}_n$ usually are not atoms in $\mathfrak{F}r_k\mathsf{CA}_n$.

The metalogical proof of Theorem 3.1 automatically proves that $\mathfrak{Z}d\mathfrak{F}r_k\mathsf{CA}_n$ is not atomic either, if $2 < n < \omega$. In [18], the locally finite part of $\mathfrak{F}r_k\mathsf{CA}_\alpha$ for infinite α is characterized, and this solves [14, Problem 2.10]. This implies that $\mathfrak{Z}d\mathfrak{F}r_k\mathsf{CA}_\alpha$ is atomic if $\alpha \geq \omega > n$.

On the other hand, the algebraic proofs in [25] show that there is an atom of $\mathfrak{Z}d\mathfrak{F}r_k\mathsf{CA}_n$ (for $0 < k < \omega$ and $4 \leq n < \omega$) below which there is no atom of $\mathfrak{F}r_k\mathsf{CA}_n$. We do not know whether this holds for $n = 3$ or not. As for the conjecture in [14] about the nonzero-dimensional atoms in the case $\alpha \geq \omega$, in [25] we prove that it is true for the free representable CA_α ($\alpha \geq \omega$), and we have some partial results that might point into the opposite direction for the free CA_α. Namely, in [25] we show that there is a nonzero element in $\mathfrak{F}r_k\mathsf{CA}_\alpha$ which is below d_{ij} for all $i, j \in \alpha,\ i, j \notin 2$. This cannot happen in the representable case. $\square$

The proof that $\mathfrak{F}r_k\mathsf{CA}_2$ for finite k is atomic relies on the fact that CA_2 is a discriminator variety and the equational theory of CA_2 is the same as the equational theory of finite CA_2's. See [14, 2.5.7] and [2, Theorem 4.1]. Let n be finite. Then Crs_n, the class of cylindric-relativized set algebras of dimension n [14, 3.1.1(iv)], satisfies the second condition, i.e., it is generated by its finite members as a variety (see Andréka-Hodkinson-Németi [1]) but it is not a discriminator variety. The same holds for the variety NCA_n of non-commutative cylindric algebras (see [27, 5T, p.112]) and for the varieties WA and NA of weakly associative and non-associative relation algebras (by [1], [28])); for the definitions of NA, WA, SA see [22] or, e.g., [2]. It is proved in [26] that neither one of $\mathfrak{Fr}_k\mathsf{RA}$ and $\mathfrak{Fr}_k\mathsf{SA}$ is atomic.

Problem 3.3. *Let k, n be finite. Are the k-generated free Crs_n's and NCA_n's atomic? Is $\mathfrak{Fr}_k\mathsf{WA}$ atomic? Is $\mathfrak{Fr}_k\mathsf{NA}$ atomic?* □

Our next subject is generating and free subsets of free algebras. $\mathfrak{Fr}_k\mathsf{CA}_n$ cannot be generated by fewer than k elements by [14, 2.5.20] and all free generator sets have cardinality k. By [14, 2.5.23], every k-element generator set of $\mathfrak{Fr}_k\mathsf{CA}_n$ is a free generator set, if $n \leq 2$ and k is finite. Problem 2.7 of [14] asks if this continues to hold for $3 \leq n$ and finite k. The following theorem, proved as [27, Theorem 19, p.100], gives a negative answer. Its proof essentially uses the translation mapping in the present Theorem 1.7.

Theorem 3.4. *There is a k-element irredundant non-free generator set in $\mathfrak{Fr}_k\mathsf{CA}_n$, for every $0 < k$ and $3 \leq n$.* □

The proof of the above theorem goes by finding such generator sets in $\mathfrak{Fr}_k\mathsf{RCA}_\omega$, which allows us to think in a model theoretical way, and then using the translation function (in a non-trivial way) of Theorem 1.7 to translate the idea from RCA_ω to CA_n.

Andréka-Jónsson-Németi [2, Theorem 9.1] generalizes the existence of non-free generator sets in Theorem 3.4 from CA_n to many subvarieties of SA. We note that Jónsson-Tarski [16] proves that if a variety is generated by finite algebras then any k-element generator set of the k-generated free algebra generates it freely. Thus CA_n in Theorem 3.4 cannot be replaced with Crs_n or NCA_n or WA or NA.

Are there big, not necessarily generating but free subsets in $\mathfrak{Fr}_k\mathsf{CA}_n$? A way of formalizing this question is whether $\mathfrak{Fr}_{k+1}\mathsf{CA}_n$ can be embedded in $\mathfrak{Fr}_k\mathsf{CA}_n$. This question is investigated thoroughly in [2] and the following negative answer is proved as part of [2, Theorem 10.3].

Theorem 3.5. *$\mathfrak{Fr}_{k+1}\mathsf{CA}_n$ is not embeddable into $\mathfrak{Fr}_k\mathsf{CA}_n$, for $0 < k < \omega$.* □

Many properties of free cylindric algebras are proved in [2]. E.g., [2, Theorem 10.1] gives a complete structural description of the free CA_1's, i.e., of the free monadic algebras. The number of elements of this free algebra is given in [14, 2.5.62]. The cardinality of finitely generated free monadic Tarski algebras (which is a reduct of CA_1) is given in [24].

We had many illuminating discussions in the subject of this paper, during writing it, with Barry Cooper, Steven Givant, Ági Kurucz, András Simon and with the members of the Logic and Philosophy of Mathematics student and faculty seminar of the Eötvös University Budapest. This research was supported by the Hungarian Scientific Research Fund grant No T81188.

References

[1] Andréka, H., Hodkinson, I., and Németi, I., Finite algebras of relations are representable on finite sets. The Journal of Symbolic Logic 64,1 (1999), 243-267.

[2] Andréka, H., Jónsson, B., and Németi, I., Free algebras in discriminator varieties. Algebra Universalis 28 (1991), 401-447.

[3] Andréka, H., Monk, J. D., and Németi, I., eds., Algebraic Logic. Colloq. Math. Soc. J. Bolyai Vol. 54, North-Holland 1991, vi + 746pp.

[4] Andréka, H., and Németi, I., Formalizing set theory in weak logics, searching for the weakest logic with Gödel's incompleteness property. Rényi Institute of Mathematics, Budapest, Preprint, 2011. 26pp. arXiv:1111.0995v1

[5] Andréka, H., and Németi, I., Reducing predicate logic to propositional logic, Part II. Notes for a talk in Department of Logic, ELTE Budapest, December 9, 2011. `http://www.math-inst.hu/pub/algebraic-logic/LapomDf2011b.pdf`

[6] Andréka, H., Németi, I., and Sain, I., Algebraic Logic. In: Handbook of Philosophical Logic Vol. 2, Second Edition, Kluwer, 2001. pp.133-296. `http://www.math-inst.hu/pub/algebraic-logic/handbook.pdf`

[7] Andréka, H., and Sain, I., Connections between algebraic logic and initial algebra semantics of CF languages. In: Mathematical Logic in Computer Science (Proc. Coll. Salgótarján 1978) Eds.: Dömölki, B., and Gergely, T., Colloq. Math. Soc. J. Bolyai Vol. 26, North-Holland, Amsterdam, 1981. pp.25-83.

[8] Chang, C., and Keisler, H. J., Model Theory. North-Holland, Amsterdam, 1990.

[9] Cooper, S. B., Computability theory. Chapman Hall/CRC Mathematics Series, 2003.

[10] Craig, W., and Vaught, R. L., Finite axiomatizability using additional predicates. The Journal of Symbolic Logic 23,3 (1958), 289-308.

[11] Ferenczi, M., A new representation theory: representation of cylindric-like algebras by relativized set algebras. Chapter in this volume.

[12] Gabbay, D. M., Kurucz, Á., Wolter, F., and Zakharyaschev, M., Many-dimensional modal logics: theory and applications. Elsevier, 2003.

[13] Gyenis, Z., On atomicity of free algebras in certain cylindric-like varieties. Logic Journal of the IGPL 19,1 (2011), 44-52.

[14] Henkin, L., Monk, J. D., and Tarski, A., Cylindric Algebras Parts I and II, North-Holland, 1985.

[15] Jeřábek, E., The ubiquity of conservative translations. Preprint, Aug. 2011. arXiv:1108.6263v1

[16] Jónsson, B., and Tarski, A., On two properties of free algebras. Math. Scand. 9 (1961), 95-101.

[17] Kurucz, Á., Representable cylindric algebras and many-dimensional modal logics. Chapter in this volume.

[18] Madarász, J. X., and Németi, I., Free Boolean algebras with closure operators and a conjecture of Henkin, Monk, and Tarski. Studia Sci. Math. Hungar. 38 (2001), 273-278.

[19] Madarász, J. X., and Sayed Ahmed, T., Amalgamation, interpolation and epimorphisms in algebraic logic. Chapter in this volume.

[20] Maddux, R. D., Some sufficient conditions for the representability of relation algebras. Algebra Universalis 8 (1978), 162-171.

[21] Maddux, R. D., The equational theory of CA_3 is undecidable. Journal of Symbolic Logic 45(1980), 311-316.

[22] Maddux, R. D., Some varieties containing relation algebras. Trans. Amer. Math. Soc. 272 (1982), 501-526.

[23] McCune, W., Solution of the Robbins problem. J. Automated Reasoning, 19,3 (1997), 263-276.

[24] Monteiro, L. F., Abad, M., Savini, S., and Sewald, J., Free monadic Tarski algebras. Algebra Universalis 37 (1997), 106-118.

[25] Németi, I., Algebraic proofs of "$\mathfrak{F}r_\beta\mathsf{CA}_\alpha$ ($\alpha \geq 4$) is not atomic". Manuscript, Mathematical Institute of the Hungarian Academy of Science, Budapest, 1984.

[26] Németi, I., Logic with three variables has Gödel's incompleteness property - thus free cylindric algebras are not atomic. Mathematical Institute of the Hungarian Academy of Sciences, Preprint No 49/1985, 1985. `http://www.math-inst.hu/~nemeti/NDis/NPrep85.pdf`

[27] Németi, I., Free algebras and decidability in algebraic logic. Dissertation with the Hungarian Academy of Sciences, Budapest, 1986. In Hungarian. English summary is [30]. `http://www.math-inst.hu/~nemeti/NDis/NDis86.pdf`

[28] Németi, I., Decidability of relation algebras with weakened associativity. Proc. Amer. Math. Soc. 100,2 (1987), 340-344.

[29] Németi, I., Decidable versions of first order logic and cylindric-relativized set algebras. In: Logic Colloquium'92 (Proc. Veszprém, Hungary 1992), eds: Csirmaz, L., Gabbay, D. M., and de Rijke, M., Studies in Logic, Language and Computation, CSLI Publications, 1995. pp.177-241.

[30] Németi, I., Free algebras and decidability in algebraic logic. Summary in English. 12pp. `http://www.math-inst.hu/~nemeti/NDis/NSum.pdf`

[31] Németi, I., Formalizing set theory in weak fragments of algebraic logic (updated in June 2011). 4pp. `http://www.math-inst.hu/~nemeti/NDis/formalizingsettheory.pdf`

[32] Németi, I., and Simon, A., Relation algebras from cylindric and polyadic algebras. Logic Journal of the IGPL 5,4 (1997), 575-588.

[33] Sain, I., Definability issues in universal logic. Chapter in this volume.

[34] Simon, A., Connections between quasi-projective relation algebras and cylindric algebras. Algebra Universalis 56,3-4 (2007), 263-301.

[35] Simon, A., Finite schema completeness for typeless logic and representable cylindric algebras. In: Algebraic Logic (Proc. Conf. Budapest 1988) Colloq. Math. Soc. J. Bolyai Vol 54, North-Holland, Amsterdam, 1991. pp.665-670.

[36] Simon, A., Representing all cylindric algebras by twisting. On a problem of Henkin. Chapter in this volume.

[37] Tarski, A., A formalization of set theory without variables. Journal of Symbolic Logic 18 (1953), p.189.

[38] Tarski, A., An undecidable system of sentential calculus. Journal of Symbolic Logic 18 (1953), p.189.

[39] Tarski, A., and Givant, S. R., Formalizing set theory without variables. AMS Colloquium Publications Vol. 41, AMS, Providence, Rhode Island, 1987.

[40] Venema, Y., Many-dimensional modal logic. Doctoral Dissertation, University of Amsterdam, Amsterdam, 1992.

[41] Venema, Y., Cylindric modal logic. Chapter in this volume.

Hajnal Andréka and István Németi
Rényi Mathematical Research Institute
Budapest, Reáltanoda st. 13-15
H-1053 Hungary
hajnal.andreka@renyi.mta.hu, istvan.nemeti@renyi.mta.hu

Paper appeared in: Cylindric-like algebras and algebraic logic, Andréka, H., Ferenczi, M., Németi, I., eds., Bolyai Society Mathematical Studies 22, Springer Verlag, Berlin, 2013. Pp.15-35.