Mutual definability does not imply definitional equivalence, a simple example.*

by

H. Andréka, J. X. Madarász, and I. Németi

Abstract

We give two theories, Th_1 and Th_2, which are explicitly definable over each other (i.e. the relation symbols of one theory are explicitly definable in the other, and vice versa), but are not definitionally equivalent. The languages of the two theories are disjoint.

1 Motivation

Intuitively, two theories (of first-order logic) are definitionally equivalent if they are (two presentations of) the same theory; i.e. if they differ only in the choices of their basic vocabularies. Definitional equivalence of theories is an important concept of logic, hence it is worthwhile to analyse it. Cf. [3], [4], [5], [7]. The question naturally comes up whether it is sufficient for definitional equivalence if the two theories are definable over each other. Definability of theories is stronger than interpretability of theories in that in definability the universes of the models remain the same, while in interpretability the universe of a model can get restricted to a definable subset.

*Research supported by Hungarian National Foundation for Scientific Research grants No's T43242, T35192 and COST grant 274.

2 Recalling the notions of definability we use

We recall the basic notions of definability we use from [4] and [3]. Let R be a relation symbol, and let L be a first-order language. An *explicit definition of R in terms of L* is a sentence of the form $\forall x_1 \ldots x_n[R(x_1,\ldots,x_n) \leftrightarrow \varphi]$ where φ is a formula of L with all its free variables among $x_1,\ldots,x_n$ ([4, p.59]). Let Th_1 and Th_2 be theories maybe on different first-order languages. An *explicit definition of Th_1 over Th_2* is a conjunction Δ of explicit definitions of the relation symbols of Th_1 in terms of the language of Th_2 such that the models of Th_1 are exactly the reducts of the models of $Th_2 \cup \Delta$ (to the language of Th_1). Thus, we get the models of Th_1 from those of Th_2 by first defining the relations of Th_1 via using Δ, and then forgetting the relations not present in the language of Th_2.

The explicit definition Δ of Th_1 over Th_2 induces a function f mapping the class $\mathsf{Mod}(Th_2)$ of all models of Th_2 onto the class $\mathsf{Mod}(Th_1)$ of all models of Th_1 such that for all models $\mathfrak{M}$ of Th_2 the universes of $\mathfrak{M}$ and $f(\mathfrak{M})$ coincide; and further Δ is a uniform definition of the relations of $f(\mathfrak{M})$ from those of $\mathfrak{M}$.

We now turn to definitional equivalence. Two theories Th_1 and Th_2 are said to be *definitionally equivalent* if they have a common definitional extension; here a definitional extension of a theory Th is a theory $Th \cup \Delta$ where Δ is a conjunction of explicit definitions of relation symbols not occurring in Th, and two theories on the same language are considered to be the same if they have the same models ([4, pp.60, 61]). Thus Th_1 and Th_2 are definitionally equivalent if there are explicit definitions Δ_1 and Δ_2 such that $Th_1 \cup \Delta_1$ and $Th_2 \cup \Delta_2$ are on the same language and have the same models on this common language.

It can be seen that Th_1 and Th_2 are definitionally equivalent iff there is a bijection f between $\mathsf{Mod}(Th_1)$ and $\mathsf{Mod}(Th_2)$ such that for all models $\mathfrak{M}$ of Th_1, the universes of $\mathfrak{M}$ and $f(\mathfrak{M})$ coincide and there are two uniform definitions, one of the relations of $f(\mathfrak{M})$ from those of $\mathfrak{M}$, and the other one of the relations of $\mathfrak{M}$ from those of $f(\mathfrak{M})$. Indeed, this is how definitional equivalence is defined in [3, p.56].

Definability theory gets used in an essential way in our work of formalizing and analyzing relativity theory in first-order logic (cf. e.g. [2], [6]). In that work, we needed to extend definability theory in such a way that we allow to define new elements also, not only define new relations on already existing elements. I.e., using the above notation, we do not require the models $\mathfrak{M}$

and $f(\mathfrak{M})$ to have the same universe. For details see [6, §4.3] and [1].

3 The example

The language of Th_1 consists of two binary relation symbols $R, \equiv$, while that of Th_2 consists of two binary relation symbols S and $\sim$. Th_1 states that R is an arbitrary binary relation and $\equiv$ an equivalence relation such that each block (i.e. equivalence class) of $\equiv$ contains infinitely many points, but at most one point in the field of R (the field of R is the union of the domain and the range of R). Th_2 states the same of S and $\sim$, but it states in addition that S is a *symmetric* relation. In more detail:

$$
\begin{aligned}
Th_1 \;&=\; Th_1(R, \equiv) \;= \\
&\{\forall xyz[(x \equiv y \to y \equiv x) \wedge (x \equiv y \wedge y \equiv z \to x \equiv z)]\} \cup \\
&\{\forall y \exists x_1 \ldots x_n (y \equiv x_1 \wedge \cdots \wedge y \equiv x_n \wedge x_1 \neq x_2 \wedge \cdots \wedge x_{n-1} \neq x_n) : n \in \omega\} \cup \\
&\{\neg \exists xy[(\exists z R(x,z) \vee \exists z R(z,x)) \wedge (\exists z R(y,z) \vee \exists z R(z,y)) \wedge x \equiv y \wedge x \neq y \ \} \ . \\
Th_2 \;&=\; Th_1(S, \sim) \cup \{\forall xy(S(x,y) \to S(y,x))\}.
\end{aligned}
$$

Having defined our two theories, now we give the two explicit definitions of one over the other. The definition of $S, \sim$ over $R, \equiv$ is very simple: for S we just take the symmetric closure of R and for $\sim$ we take $\equiv$.

$$\Delta(S, \sim) = \{\forall xy[S(x,y) \leftrightarrow (R(x,y) \vee R(y,x)] \wedge \forall xy[x \sim y \leftrightarrow x \equiv y]\}.$$

Claim 1. $\Delta(S, \sim)$ is an explicit definition of Th_2 over Th_1, i.e. the models of Th_2 are exactly the appropriate reducts of the models of $Th_1 \cup \Delta(S, \sim)$.

Proof. To prove Claim 1 we have to prove that if we take any model of Th_1 and define S and $\sim$ in it the given way (and then forget $R, \equiv$) then we obtain a model of Th_2, and conversely, every model of Th_2 can be obtained from a model of Th_1 the above way.

Indeed, let $\mathfrak{M}$ be any model of Th_1 and define $S, \sim$ according to $\Delta(S, \sim)$. Then S will be a symmetric relation with the same field as that of R. Since we took $\sim$ to be $\equiv$ and each block of $\equiv$ contained at most one point from the field of R, then each block of $\sim$ will contain at most one point from the field of S. Also, each block of $\sim$ is infinite, since this was true for $\equiv$ (since $\mathfrak{M}$ is a model of Th_1).

Conversely, let $\mathfrak{M}$ be any model of Th_2 and define $\mathfrak{M}'$ to be the same model except the names of $S, \sim$ are $R, \equiv$ respectively. Then clearly $\mathfrak{M}'$ is a model of Th_1 and when defining $S, \sim$ in it according to $\Delta(S, \sim)$ we get back $\mathfrak{M}$. $\square$

The definition of Th_1 over Th_2 is a little more involved, as could be expected. See Figures 1,2. We begin with some auxiliary definitions. In the following, "$\exists!$" denotes "there is a *unique*", i.e. the formula $\exists! x \varphi(x)$ abbreviates $\exists x[\varphi(x) \wedge \forall y(\varphi(y) \rightarrow y = x)]$.

$End(x)$ denotes the formula $\exists! z S(x,z) \wedge \neg S(x,x)$.

$Next(x)$ denotes the formula $\exists z(S(x,z) \wedge End(z))$.

$Old(x)$ denotes the formula $\neg End(x) \wedge \neg Next(x)$.

$Middle(x)$ denotes the formula $Next(x) \wedge \exists! z(Old(z) \wedge S(x,z))$.

$root(x,y)$ denotes the formula $x = y \vee [Middle(x) \wedge S(x,y) \wedge Old(y)] \vee [End(x) \wedge \exists z(S(x,z) \wedge Middle(z) \wedge S(z,y)) \wedge Old(y))]$.

$$\Sigma(R, \equiv) \;=$$

$$\{\ \forall xy[R(x,y) \leftrightarrow Old(x) \wedge Old(y) \wedge \exists u v v' v'' \\ (Next(u) \wedge End(v') \wedge End(v'') \wedge v' \neq v'' \wedge \\ S(x,y) \wedge S(x,u) \wedge S(y,v) \wedge S(u,v) \wedge S(v,v') \wedge S(v,v''))],$$

$$\forall xy[x \equiv y \leftrightarrow \exists x' z y'(x \sim x' \wedge root(x',z) \wedge root(y',z) \wedge y' \sim y)]\}.$$

Claim 2. $\Sigma(R, \equiv)$ is an explicit definition of Th_1 over Th_2, i.e. the models of Th_1 are exactly the appropriate reducts of the models of $Th_2 \cup \Sigma(R, \equiv)$.

Proof. Let $\mathfrak{M} = \langle M, S, \sim \rangle$ be any model of Th_2 and define $R, \equiv$ according to $\Sigma(R, \equiv)$. I.e. $\langle M, S, \sim, R, \equiv \rangle \models Th_2 \cup \Sigma(R, \equiv)$. We want to show that $\langle M, R, \equiv \rangle \models Th_1$.

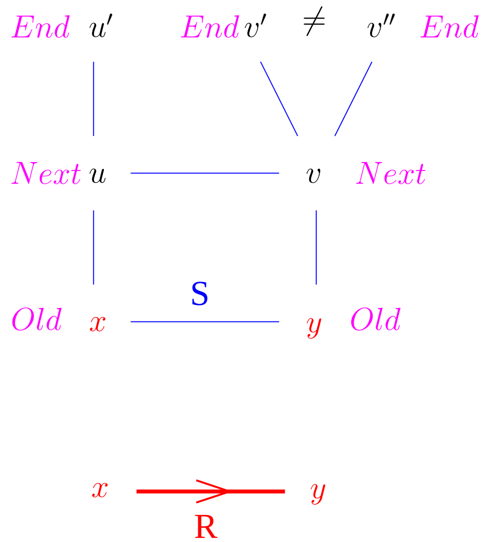

Figure 1: Illustration of the definition of R in terms of S. Notice that the existence of u' follows from $Next(u)$, and $Next(v)$ follows e.g. from $S(v, v') \wedge End(v')$.

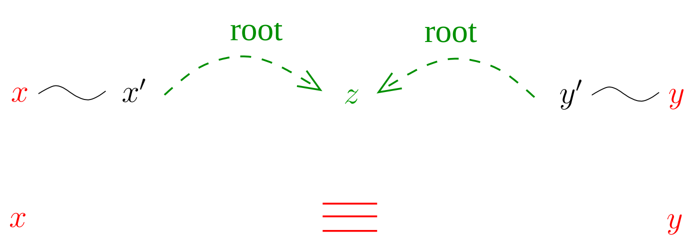

Figure 2: Illustration of the definition of $\equiv$ in terms of $\sim$ and S. See also Figure 4 on p.8.

First we show that $\equiv$ is an equivalence relation. Clearly, $\equiv$ is reflexive and symmetric. To show that it is transitive, assume that $x \equiv y$ and $y \equiv u$, we want to show that $x \equiv u$. Now, $x \equiv y$ and $y \equiv u$ mean that there are x', z, y', y'', v, u' such that $x \sim x', \mathit{root}(x', z), \mathit{root}(y', z), y' \sim y, y \sim y'', \mathit{root}(y'', v), \mathit{root}(u', v)$ and $u' \sim u$. If $x' = y'$ or $y'' = u'$ or $z = v$ then we are done. So assume that $x' \neq y' \wedge y'' \neq u' \wedge z \neq v$. We will derive a contradiction. In deriving the contradiction we will use the following three properties of *root* which we do not prove, since their verification is not difficult.

(1) $\mathit{root}(x, y) \wedge x \neq y \Rightarrow x, y$ are in the field of S.

(2) $\mathit{root}(x, y) \wedge \mathit{root}(x, z) \wedge x \neq y \wedge x \neq z \Rightarrow y = z$.

(3) $\mathit{root}(x, y) \wedge \mathit{root}(y, z) \Rightarrow x = y \vee y = z$.

By $y' \neq x'$ we have that either $y' \neq z$ or $y' = z \wedge z \neq x'$. Then y' is in the field of S in both cases, by (1). We get similarly that y'' is in the field of S. By $y' \sim y''$ and $\mathfrak{M} \models Th_2$ then $y' = y''$, since each block of $\sim$ contains at most one element from the field of S. By $\mathit{root}(y', z) \wedge \mathit{root}(y', v) \wedge z \neq v$ and (2) we get that either $y' = z$ or $y' = v$. By symmetry we may assume that $y' = z$. By $\mathit{root}(x', z) \wedge \mathit{root}(z, v) \wedge z \neq v$ and (3) then we get $x' = z$. This contradicts our hypothesis $x' \neq y'$ and $z = y'$. We have derived a contradiction, which proves that $\equiv$ is an equivalence relation.

Each block of $\equiv$ is infinite since $\sim \subseteq \equiv$ and each block of $\sim$ is infinite.

Finally we show that each block of $\equiv$ contains at most one element from the field of R. Assume therefore that $x \equiv y$ and both x and y are in the field of R. We have to show that $x = y$. Now, $x \equiv y$ means that there are x', z, y' such that $x \sim x'$, $\mathit{root}(x', z)$, $\mathit{root}(y', z)$ and $y' \sim y$. By $R \subseteq S$ (which follows from $\Sigma(R, \equiv)$) we have that x, y are in the field of S, too. Thus if $x' = y'$ then we are done since then $x \sim y'$ and each block of S contains at most one point from the field of S. Assume therefore that $x' \neq y'$, we will derive a contradiction. By $x' \neq y'$ we have that either $x' \neq z$ or $y' \neq z$. By symmetry we may assume that $x' \neq z$. By $\mathit{root}(x', z)$ and (1) then we have that x' is in the field of S. By $x \sim x'$ then we have $x = x'$ since each block of $\sim$ contains at most one element from the field of S. Now we have $\mathit{Old}(x)$ because x is in the field of R. But then $\mathit{Old}(x')$ by $x = x'$. This contradicts $\mathit{root}(x', z) \wedge x' \neq z$ since the following can be checked:

(4) $\mathit{root}(x, y) \wedge x \neq y \Rightarrow \neg \mathit{Old}(x)$.

By now we have shown that $\langle M, R, \equiv \rangle$ is a model of Th_1.

Conversely, let $\mathfrak{M} = \langle M, R, \equiv \rangle$ be any model of Th_1. We want to find a model $\mathfrak{M}'$ of Th_2 such that if we define $R, \equiv$ in $\mathfrak{M}'$ according to $\Sigma(R, \equiv)$ (and forget $S, \sim$), then we get our original $\mathfrak{M}$. In other words, we want to find $S, \sim$ such that $\langle M, S, \sim, R, \equiv \rangle \models Th_2 \cup \Sigma(R, \equiv)$. This is where we will use the equivalence relations $\equiv, \sim$.

We now "construct" the relations S and $\sim$ on M. For each $x \in M$ such that x is in the field of R let u_x, u'_x, v_x, v'_x and v''_x be distinct (distinct from x, too) elements in the $\equiv$-block of x. There are such elements because each block of $\equiv$ is infinite. We define S as follows (see Figure 3):

$$S' = \bigcup \{\{(x, u_x), (u_x, u'_x), (x, v_x), (v_x, v'_x), (v_x, v''_x)\} : x \text{ is in the field of } R\} \cup$$
$$\{(u_x, v_y) : (x, y) \in R\} \cup R.$$
$$S = S' \cup S'^{-1}.$$

We define the equivalence relation $\sim$ such that it is a refinement of $\equiv$. More specifically, let $\sim$ be an equivalence relation with the following properties:

Assume that x is in the field of R. Then the $\equiv$-block of x is partitioned to six $\sim$-blocks, each of them infinite, and each of them containing exactly one element from $\{x, u_x, u'_x, v_x, v'_x, v''_x\}$. Note that in this case x is the only element of the $\equiv$-block of x which is in the field of R. See Figure 4.

A $\equiv$-block which does not contain an element from the field of R is also a $\sim$-block.

Having defined $S, \sim$ we have to show that $\langle M, S, \sim, R, \equiv \rangle \models Th_2 \cup \Sigma(R, \equiv)$.

Clearly, S is symmetric and each block of $\sim$ is infinite. It is easy to check that each block of $\sim$ contains at most one element from the field of S. Thus Th_2 holds. To check that $\Sigma(R, \equiv)$ also holds, first we check some auxiliary statements. Let $Field(S), Field(R)$ denote the fields of S and R, respectively. We will use the following statements (5)-(9).

(5) $Field(S) = \bigcup \{\{x, u_x, u'_x, v_x, v'_x, v''_x\} : x \in Field(R)\}$.

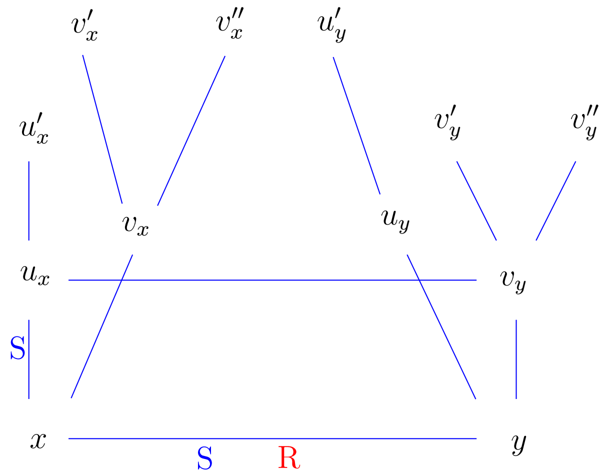

Figure 3: Illustration of the construction of S.

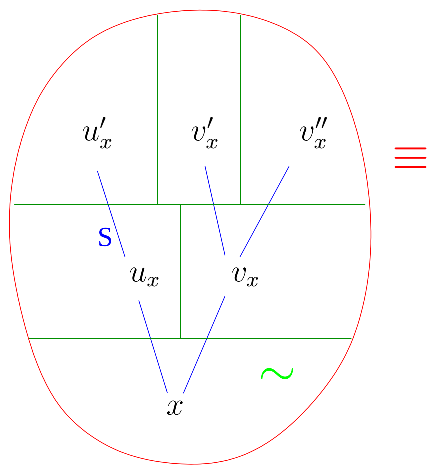

Figure 4: Illustration of the construction of $\sim$. A $\equiv$-block which contains an element from the field of R is partitioned into 6 infinite $\sim$-blocks.

(6) $End(x)$ if and only if $x \in \bigcup\{\{u'_y, v'_y, v''_y\} : y \in Field(R)\}$.

(7) $Next(x)$ iff $Middle(x)$ iff $x \in \bigcup\{\{u_y, v_y\} : y \in Field(R)\}$.

(8) $(Old(x) \wedge x \in Field(S))$ iff $x \in Field(R)$.

(9) $root(x,y) \wedge x \neq y \Rightarrow x \in \{u_y, u'_y, v_y, v'_y, v''_y\}$.

Statement (5) follows immediately from the definition of S. To prove (6), assume that $End(x)$. Then $x \in Field(S)$ and it can be checked by inspecting the definition of S that all elements from $\bigcup\{\{y, u_y, v_y\} : y \in Field(R)\}$ are in S-relation with at least two other elements. This shows implication $\Rightarrow$ in (6). The other direction follows by checking that all elements of form u'_y, v'_y, v''_y are indeed in S-relation with a unique, distinct element. The proofs of (7)-(9) are similar, we omit them.

To prove that $\Sigma(R, \equiv)$ holds, let $R', \equiv'$ be the unique relations for which $\Sigma(R', \equiv')$ holds, we want to prove that $R = R'$ and $\equiv = \equiv'$.

$R \subseteq R'$ follows from (6)-(8) and the construction of S, as follows. Assume $R(x,y)$, we have to show $R'(x,y)$. The latter means $Old(x) \wedge Old(y) \wedge \exists uvv'v''(\dots)$. Now, $Old(x), Old(y)$ follow from (8), and for u, v, v', v'' take u_x, v_y, v'_y, v''_y. The rest follows from the construction of S and (7),(6).

To show $R' \subseteq R$, assume $R'(x,y)$. This implies that $Old(x), Old(y)$ and there are u, v, v', v'' such that $Next(u), End(v'), End(v''), v' \neq v''$ and $S(x,y), \dots$. By $Old(y), S(y,v), End(v'), End(v''), S(v,v'), S(v,v''), v' \neq v''$, (6)-(8) and the definition of S we get that $v = v_y, v' = v'_y$ and $v'' = v''_y$. Similarly we get that $u \in \{u_x, v_x\}$. By $S(u,v)$ then we must have $u = u_x$ because S does not contain pairs of form (v_x, v_y). But then $S(u_x, v_y)$ and the definition of S show that $R(x,y)$. We have seen that $R = R'$.

Assume $x \equiv y$. If the $\equiv$-block of x does not contain an element from $Field(R)$, then $x \sim y$ by the definition of $\sim$ and so $x \equiv' y$ by the definition of $\equiv'$. Assume that $x \equiv z \in Field(R)$. Then by the definition of $\sim$ we have that x, y are in the $\sim$-block of one of $\{z, u_z, \dots, v''_z\}$. I.e. there are $x', y' \in \{z, u_z, \dots, v''_z\}$ such that $x \sim x'$ and $y \sim y'$. By (9) we then have $root(x', z)$ and $root(y', z)$. Hence $x \equiv' y$ by the definition of $\equiv'$. We have seen that $\equiv \subseteq \equiv'$. To show the other inclusion, notice first that $\sim \subseteq \equiv$ because $\sim$ is a refinement of $\equiv$. Also, $root(x,y) \Rightarrow x \equiv y$ by (9) since we chose $u_y, \dots, v''_y$ from the $\equiv$-block of y. Now, $x \equiv' y$ implies the existence of x', z, y' such that $x \sim x', root(x', z), root(y', z), y' \sim y$. By the above then $x \equiv x' \equiv z \equiv y' \equiv y$, and thus $x \equiv y$ since $\equiv$ is an equivalence relation. We have seen that $\equiv = \equiv'$. □

Claim 3. Th_1 and Th_2 are not definitionally equivalent. I.e. Th_1 and Th_2 do not have a common definitional extension.

Proof. Assume that Th_1 and Th_2 are definitionally equivalent, say Th_3 is a common definitional extension of both Th_1 and Th_2. This means that there are definitions $\Delta'(S,\sim)$ and $\Sigma'(R,\equiv)$ such that $Th_1 \cup \Delta'(S,\sim)$ is equivalent to $Th_2 \cup \Sigma'(R,\equiv)$. Then there is a bijection f between $\mathsf{Mod}(Th_1)$ and $\mathsf{Mod}(Th_2)$ such that $\mathfrak{M}$ and $f(\mathfrak{M})$ have the same universes, and moreover, they have the same automorphisms. E.g., for any $\mathfrak{M} \in \mathsf{Mod}(Th_1)$ we can take the unique expansion $\mathfrak{M}' \models Th_3$, and let $f(\mathfrak{M})$ be the reduct of $\mathfrak{M}'$ to the language of Th_2.

Let X, Y be two disjoint infinite sets of the same cardinality, $a \in X, b \in Y, M = X \cup Y, S = \{(a,a),(a,b),(b,a)\}, \sim = (X \times X) \cup (Y \times Y)$ and $\mathfrak{M} = \langle M, S, \sim \rangle$. Then $\mathfrak{M} \models Th_2$ and the automorphisms of $\mathfrak{M}$ are exactly the bijections of M that leave a, b fixed and map X to X and Y to Y. There is exactly one other model $\mathfrak{M}'$ of Th_2 with universe M and the same automorphisms, namely $\mathfrak{M} = \langle M, \{(a,b),(b,a),(b,b)\}, \sim \rangle$. On the other hand, there are more than two models of Th_1 with the same properties. E.g. $\mathfrak{M}_1 = \langle M, \{(a,b)\}, \sim \rangle$, $\mathfrak{M}_2 = \langle M, \{(b,a)\}, \sim \rangle$, $\mathfrak{M}_3 = \langle M, \{(a,b),(b,b)\}, \sim \rangle$ are three different models of Th_1 with universe and automorphisms the same as those of $\mathfrak{M}$. Hence there cannot be a bijection f as we described between the models of Th_1 and of Th_2. □

Acknowledgements. We are grateful to the anonymous referee for his very useful suggestions.

References

[1] Andréka, H., Madarász, J. X. and Németi, I.: *Defining new universes in many-sorted logic.* Preprint, Rényi Institute of Mathematics, Budapest, 2002.

[2] Andréka, H., Madarász, J. X. and Németi, I.: *Logical analysis of relativity theories.* In: First-order Logic Revisited (Hendricks et al. eds), Logos Verlag, Berlin, 2004. pp.7-36.

[3] Henkin, L., Monk, J. D. and Tarski, A.: *Cylindric Algebras.* North-Holland, Amsterdam, 1971 and 1985.

[4] Hodges, W.: *Model Theory.* Cambridge University Press, Cambridge, 1993.

[5] Hodges, W., Hodkinson, I. and Macpherson, H.: *Omega-categoricity, relative categoricity and coordinatisation.* Annals of Pure and Applied Logic 46,4 (1990), 169-199.

[6] Madarász, J. X.: *Logic and relativity (in the light of definability theory).* PhD Dissertation, Eötvös Loránd University, Budapest, 2002. www.math-inst.hu/pub/algebraic-logic/Contents.html.

[7] Makkai, M.: *Duality and Definability in First Order Logic.* Mamoirs of the AMS, Vol. 105, No. 503, American Mathematical Society, 1993.

Rényi Institute of Mathematics
Budapest, Reáltanoda u. 13-15
H-1053 Hungary.
[andreka,madarasz,nemeti]@renyi.hu