Definability and Undecidability

André Nies
The University of Chicago

June 26, 1999

Definability: examples

Structure	Subset	Formula
$(\mathbb{R}, +, \cdot)$	$\{r : r \geq 0\}$	$\exists y \; \boldsymbol{x} = y \cdot y$
$(\mathbb{Z}, +, \cdot)$	$\{n : n \geq 0\}$	$\exists y_1 \exists y_2 \exists y_3 \exists y_4 \; \boldsymbol{x} = y_1 \cdot y_1 + y_2 \cdot y_2 + y_3 \cdot y_3 + y_4 \cdot y_4$

Some concepts from logic

- S symbol set, e.g. $\{f, g, R\}$ (R binary relation symbol, f, g unary function symbols)
- First-order formulas over S:

$$R\boldsymbol{x}\boldsymbol{x},\ \forall x \exists y\ (Rxfy\ \wedge\ \neg fy = g\boldsymbol{z}).$$

- An S-structure is of the form

$$\boldsymbol{A} = (A, f^A, g^A, R^A).$$

- $\boldsymbol{A} \models \varphi(a)$ denotes

"a satisfies φ in $\boldsymbol{A}$".

Definability

- $M \subseteq \boldsymbol{A}$ is **definable** in first-order language if

$$M = \{a \in A : \boldsymbol{A} \models \varphi(a)\}$$

  for a formula $\varphi(x)$ in the language of $\boldsymbol{A}$.

- Similarly: definability for relations $R \subseteq A^n$.
- There are only countably many definable relations on A.

The significance of definability

A property of an element of $\boldsymbol{A}$ is definable

$$\Leftrightarrow$$

the property can be formulated without reference to objects external to $\boldsymbol{A}$.

In other words, one can restrict oneself to the universe given by $\boldsymbol{A}$. Not permitted are:

- quantifying over subsets of $\boldsymbol{A}$
- infinite disjunctions
- ...

Defining $\mathbb{N}$ in $\mathbb{Q}$

Theorem 1 (Julia Robinson) $\mathbb{N} \subseteq \mathbb{Q}$ *is definable in* $(\mathbb{Q}, +, \times)$.

A *second*-order definition: $q \in \mathbb{N} \;\Leftrightarrow\; \forall X \subseteq \mathbb{Q}$

$(*)\quad 0 \in X \wedge \forall m[m \in X \Rightarrow m{+}1 \in X] \;\Rightarrow\; \forall r\, r \in X$

J.R. shows: it suffices to require $(*)$ for all the sets

$$X_{a,b} = \{w : abw^2 + 2 \in Z_{a,b}\}.$$

$Z_{a,b}$: the range of the quadratic form

$$x^2 + ay^2 - bz^2$$

Thus
$q \in \mathbb{Q} - \mathbb{Z} \;\Rightarrow\; \exists a \exists b[q \notin X_{a,b} \;\wedge\; X_{a,b}$ inductive$]$.
Now we can express $(*)$ by a first–order formula. $\diamondsuit$

Parameters

$U \subseteq \boldsymbol{A}^k$ is **parameter-definable**

$M = \{a_1, \ldots, a_k \in A^k :$

$$\boldsymbol{A} \models \varphi(a_1, \ldots, a_k; p_1, \ldots, p_n)\}$$

for a formula $\varphi(x)$ and a list of parametern $p_1, \ldots, p_n$.

A further example

Algebraic curves over $\mathbb{R}$ are parameter-definable relations in $(\mathbb{R}, +, \times)$. E.g.

$$\{x, y : y^2 = x^3 + ax^2 + bx + c\}$$

is a parameter-definable relation. It is definable if $a, b, c \in \mathbb{Q}$. The defining formula has no quantifiers.

Invariance and L_∞

- $M \subseteq \boldsymbol{A}$ is **invariant** if $\Phi(M) = M$ for every automorphism Φ of $\boldsymbol{A}$.
- Definable $\Rightarrow$ invariant.

L_∞ (or $L_{\omega_1,\omega}$) consists of formulas which can be built up with $\exists x, \neg$ and countable conjunctions and disjunctions. This language is considerably more expressive than first-order language. For instance, every $M \subseteq \mathbb{N}$ is L_∞-definable in $(\mathbb{N}, +, 1)$ via

$$\varphi(\boldsymbol{x}) \Leftrightarrow \bigvee_{n \in M} \boldsymbol{x} = \underbrace{1 + \ldots + 1}_{n}.$$

Easy: L_∞-definable in $\boldsymbol{A}$ $\Rightarrow$ invariant in $\boldsymbol{A}$.
For countable $\boldsymbol{A}$ the converse is true.

Theorem 2 (Scott) *If $\boldsymbol{A}$ is countable and $U \subseteq \boldsymbol{A}$, then*

$$U \text{ invariant} \;\Rightarrow\; U\ L_\infty\text{-definable.}$$

Undecidability

The **halting problem**: Let (M_n) be an effective list of all Turingmachines. A diagonal argumnent proves that
$K = \{n : \text{the computation of } M_n$
$\quad \text{with input } n \text{ stops}\}$
is undecidable.

For which structures $\boldsymbol{A}$ is $\mathrm{Th}(\boldsymbol{A})$ decidable ?

$(\mathbb{N}, +, \times)$	no	halting problem
$(\mathbb{Q}, +, \times)$	no	$\mathbb{N} \models \varphi \Leftrightarrow$ $\mathbb{Q} \models \varphi^{[\{x:\alpha(x)\}]}$
$(\mathbb{R}, +, \times)$	yes	quantifier-elimination

Coding

$(\mathbb{N}, +, \times)$ can be **coded** in $\boldsymbol{A}$ if there are parameter-definable relations as follows:

- a set D,
- an equivalence relation $\equiv$
- ternary relations R, S

such that

$$(D, R, S)/_{\equiv} \;\cong\; (\mathbb{N}, +, \times).$$

Example: $(\mathbb{N}, +, \times)$ can be coded in $(\mathbb{Q}, +, \times)$ (even without parameters).

Theorem 3 (Rabin) *If* $(\mathbb{N}, +, \times)$ *can be coded in* $\boldsymbol{A}$*, then* $Th(A)$ *is undecidable.*

As a consequence, $\mathbb{N} \subseteq \mathbb{R}$ is *not* parameter-definable in $(\mathbb{R}, +, \times)$.

Computability theory

$W \subseteq \mathbb{N}$ is **computably enumerable (c.e.)** if there is a Turingmachine T such that

$$W = \{n : \text{ the computation of } T \text{ on} input\ n \text{ stops}\}.$$

C.e. sets are of central importance in mathematical logic and also occur in other areas of mathematics (e.g. the Higman Embedding Theorem).
The most elementary way to compare c.e. sets is **inclusion**.

$$\mathcal{E} = (\text{c.e. sets}, \subseteq)$$

is a distributive lattice with least and greatest elements. In $\mathcal{E}$,

$$\text{computable} \;\Leftrightarrow\; \text{complemented}$$

Reducibilities

Let $X, Y \subseteq \mathbb{N}$.
$X \leq_m Y \;\Leftrightarrow$

$$X = f^{-1}(Y) \text{ for a computable } f.$$

$X \leq_T Y \;\Leftrightarrow$

X is computed by an oracle-TM T where Y is the oracle set.

$$\mathcal{R}_T = (\text{C.e. sets}, \leq_T)/_{\equiv_T}$$

is the partial order of Turing degrees. (Degree: equivalence class of sets of the same complexity.) Again it has least and greatest elements. Least: degree consisting of the computable sets. Greatest: degree of the halting problm.
In a similar way one obtains $\mathcal{R}_m$, the degree structure on c.e. sets based on m-reducibility.
R_T is dense, while $\mathcal{R}_m - \{0\}$ has minimal elements.

Undecidability and beyond

Harrington, Slaman 1985:

Theorem 4 $(\mathbb{N}, +, \times)$ *is parameter–definable in* $\mathcal{R}_T$. *Th*$(\mathbb{N})$ *can be interpreted in* *Th*(R_T).

A simpler proof was found by Slaman/Woodin.

N, 1994:

Theorem 5 *The same for* $\mathcal{R}_m$.

Harrington, 1995:

Theorem 6 *The same for* $\mathcal{E}$.

Definability results

If A is a structure from computability theory, definability of a set $M \subseteq \boldsymbol{A}$ often means:

> Subsets which at first were introduced using concepts external for $\boldsymbol{A}$ can actually be recognized internally.

Theorem 7 (Harrington)

$$\mathcal{C} = \{A \ c.e. : A \ m - complete\}$$

is definable.

$A \in \mathcal{C} \;\Leftrightarrow\; \exists C \; \forall B \; \exists R$ comp.
$C \cap R$ non-comp. $\wedge \; A \cap R = B \cap R$.

(Recall: computable $\Leftrightarrow$ complemented in $\mathcal{E}$.)

The formula is $\exists\forall\exists$.

The jump-operator

Let (M_n) be an effective list of all oracle- Turing-machines.

- $K^X = \{n : \text{the computation of } M_n$ on input n with oracle X stops$\}$
  Then $K^X >_T X$.
- If $\boldsymbol{x}$ is a Turing–degree, let $\boldsymbol{x}'$ be the degree of K^X, for a set $X \in \boldsymbol{x}$.
- $\boldsymbol{x}^{(n)}$ is the result of applying the jump-operator to x for n times.
- $$\begin{aligned} Low_n &= \{\boldsymbol{x} : \boldsymbol{x}^{(n)} = 0^{(n)}\}, \\ High_n &= \{\boldsymbol{x} : \boldsymbol{x}^{(n)} = 0^{(n+1)}\} \end{aligned}$$

Definability in $\mathcal{R}_T$

N,Shore, Slaman (1996) prove definability results using coding methods.
We write $\boldsymbol{x} \sim_2 \boldsymbol{y}$ for $\boldsymbol{x}^{(2)} = \boldsymbol{y}^{(2)}$.

Theorem 8 *Let $\mathcal{C} \subseteq \mathcal{R}_T^n$ be closed under $\sim_2$. If the index set of $\mathcal{C}$ is arithmetical, then $\mathcal{C}$ is definable.*

The index set of $\mathcal{C}$ is $\{e : \deg_T(W_e) \in \mathcal{C}\}$, where $(W_e)_{e \in \mathbb{N}}$ is an effective list of c.e. sets.

Corollary: definable are

- Low_n, $High_n$ for $n \geq 2$
- The relation $\sim_2$ itself.

Theorem 9 *$High_1$ is definable in $\mathcal{R}_T$.*

Formula:

$$\boldsymbol{x} \in High_1 \Leftrightarrow \mathcal{R}_T \models \forall \boldsymbol{y}\ \exists \boldsymbol{z} \leq \boldsymbol{x}[\boldsymbol{z} \sim_2 \boldsymbol{y}].$$