# Definability and Undefinability with Real Order at the Background


Yuri Gurevich*
Microsoft Research
One Microsoft Way Redmond, WA 98052-6399, USA

Alexander Rabinovich
Department of Computer Science,
Beverly Sackler School of Exact Sciences,
Tel Aviv University, Israel 69978


## 1 Introduction

We consider the monadic second-order theory of linear order. For the sake of brevity, linearly ordered sets will be called chains.

Let $\mathcal{A} = \langle \mathbf{A} < \rangle$ be a chain. A formula $\phi(t)$ with one free individual variable $t$ defines a point-set on $\mathbf{A}$ which contains the points of $\mathbf{A}$ that satisfy $\phi(t)$. As usually we identify a subset of $\mathbf{A}$ with its characteristic predicate and we will say that such a formula defines a predicate on $\mathbf{A}$.

A formula $\chi(X)$ with one free monadic predicate variable defines the set of predicates (or family of point-sets) on $\mathbf{A}$ that satisfy $\chi(X)$. This family is said to be definable by $\chi(X)$ in $\mathcal{A}$. Suppose that $\mathcal{A}$ is a subchain of $\mathcal{B} = \langle \mathbf{B}, < \rangle$. With a formula $\chi(X,\ A)$ we associate the following family of point-sets (or set of predicates) $\{\mathbf{P} : \mathbf{P} \subseteq \mathbf{A}$ and $\chi(\mathbf{P}, \mathbf{A})$ holds in $\mathcal{B}\}$ on $\mathbf{A}$. This family is said to be definable by $\chi$ in $\mathcal{A}$ with $\mathcal{B}$ at the background.


*This author is on leave of absence from the University of Michigan. His work on the paper was supported by the grant CCR 95-04375.

Note that in such a definition bound individual (respectively predicate) variables of $\chi$ range over $\mathbf{B}$ (respectively over subsets of $\mathbf{B}$). Hence, it is reasonable to expect that the presence of a background chain $\mathcal{B}$ allows one to define point sets (or families of point-sets ) on $\mathbf{A}$ which are not definable inside $\mathcal{A}$. We show

**Theorem 1** *A family of point-sets is definable in the chain $\mathcal{N}$ of natural numbers if and only if it is definable in $\mathcal{N}$ with the chain of reals at the background.*

Hence, the quantification over the reals does not allow to define more point-set families on the natural numbers.

Let us point out some difficulties which arise when one attempts to prove Theorem 1 and the techniques used to overcome these difficulties. The *only-if* part of the theorem is easily obtained by the method of interpretation (see Theorem 5). This method provides an algorithm that for a formula $\chi(X)$ constructs a formula $\chi^*(X, A)$ such that for any chain $\mathcal{B}$ and its subchain $\mathcal{A} = \langle \mathbf{A}, < \rangle$ a family of point-sets on $\mathbf{A}$ is definable by $\chi(X)$ in $\mathcal{A}$ if and only if it is definable by $\chi^*(X,\ A)$ in $\mathcal{A}$ with $\mathcal{B}$ at the background. Hence, definability with a background is at least as powerful as definability without any background. To show the other direction of Theorem 1 one may also try to look for a translation algorithm. However,

**Theorem 2** *There is no algorithm that translates an arbitrary monadic formula $\chi(X,\ A)$ into a monadic formula $\phi(X)$ in such a way that the point-set family defined by $\phi(X)$ in $\mathcal{N}$ is equal to the point-set family defined in $\mathcal{N}$ by $\chi(X,\ A)$ with $\mathcal{R}$ at the background.*

In the above theorem and below we use the convention that $X,\ A$ are monadic predicate variables and a formula does not contain free variables except for those shown explicitly.

In order to see that Theorem 2 holds recall that the satisfiability problem for the monadic theory of real order is undecidable [6]. On the other hand, the satisfiability problem is decidable over the natural numbers [1]. Now observe that if $\phi$ is a sentence, then it is satisfiable over the reals iff $\phi \wedge X \subset A$ defines (in $\mathcal{N}$ with the reals at the background) the family of all point-sets on the natural numbers. Hence, the translation of $\phi \wedge X \subset A$ should be a formula $\phi^*(X)$ such that $\forall X.\ \phi^*(X)$ is satisfiable over the natural numbers. In this way the undecidable problem of satisfiability over the reals is reduced by the translation mapping to the decidable problem of satisfiability over the natural numbers. Hence, the translation cannot be recursive.

The proof of the *if* direction of Theorem 1 will be based on the composition theorem [6, 2]. The composition theorem is a very powerful theorem for the proofs of decidability of various theories [6, 3]. Here is the first application of the composition theorem to definability. The method of interpretation is also used extensively in our proof.

The rest of the paper is organized as follows. In Section 2 we fix notations and terminology, state some preliminary results and provide a formulation of the composition theorem which is needed later. This version of the composition theorem is weaker than the version in [6], however, its formulation is simple. We recall also the method of interpretation and show the easy direction of Theorem 1.

In Section 3 Theorem 1 is proved. The results of Section 4 imply

**Theorem 3** *For any closed subset* $\mathbf{F}$ *of the reals, a family of point-sets is definable in the subchain* $\mathcal{F} = \langle \mathbf{F}, < \rangle$ *of reals if and only if it is definable in* $\mathcal{F}$ *with the chain of reals at the background.*

In fact, we prove a somewhat stronger theorem, Theorem 13, in Section 4. We show that there is a uniform way to translate a definition (in the closed subchains of the reals) with the reals at the background into a definition without background. It is instructive to contrast Theorem 3 with

**Theorem 4** *There exists an open subset* $\mathbf{G}$ *of the reals and a formula* $\chi(X,\ G)$ *such that the family of point-sets is definable by* $\chi(X,\ G)$ *in* $\mathcal{G} = \langle \mathbf{G}, < \rangle$ *with the reals at the background is not definable in* $\mathcal{G}$*.*

Indeed, let $\mathbf{G}_0$ be the open subset $(0,1)+(1,2)+(2,3)+\ldots$ of the reals. Since there are only countably many definable point-families, there is a subset $\mathbf{P}$ of natural numbers such that the family $\{(i, i+1) : i \in \mathbf{P}\}$ is not definable in $\langle \mathbf{G}_0,\ < \rangle$. The desired $\mathbf{G}$ is the open set $\bigcup\{(2i, 2i+2) :\ i\ \notin \mathbf{P}\} \cup \bigcup\{(2i, 2i+1)\ :\ i \in \mathbf{P}\}$. Since $\mathbf{G}$ is isomorphic to $\mathbf{G}_0$, the point-set family $\{(2i, 2i+1)\ :\ i \in \mathbf{P}\}$ is not definable in $\mathbf{G}$. Yet this family is definable in $\mathbf{G}$ with the reals at the background; it consists of all maximal intervals in $\mathbf{G}$ with a positive length interval of $\mathbf{R} - \mathbf{G}$ adjacent to it at the right.

We conclude this section with

**Open Problem:** Is it true that a family of point-sets is definable in the chain $\mathcal{Q}$ of rationals if and only if it is definable in $\mathcal{Q}$ with the chain of reals at the background.

# 2 Preliminaries

## 2.1 Notations and terminology

$\mathbf{N}$ is the set of natural numbers; $\mathbf{R}$ is the set of real numbers, $\mathbf{R}^{\geq 0}$ is the set of non negative reals; $\mathbf{BOOL}$ is the set of booleans and $\Sigma$ is a non-empty finite set. A $\Sigma$-predicate or $\Sigma$-coloring over a set $\mathbf{A}$ is a function from $\mathbf{A}$ into $\Sigma$; the letters $\mathbf{P}$, $\mathbf{Q}$ range over $\Sigma$-predicates. Whenever the domain $\mathbf{A}$ and the range $\Sigma$ of $\mathbf{P}$ is clear from the context we use 'predicate' or 'coloring' for '$\Sigma$-predicate over $\mathbf{A}$'. A subset $\mathbf{B}$ of a set $\mathbf{A}$ will be identified with the corresponding boolean predicate over $\mathbf{B}$. Accordingly a point-set family is identified with the set of corresponding boolean predicates. It is well-known that if $\Sigma$ is an alphabet of size $n > 1$ and $k$ is the least positive integer such that $n < 2^k$ then $\Sigma$-colorings can be coded with $k$ boolean predicates. The restriction of a predicate $\mathbf{P}$ onto a set $\mathbf{A}$ will be denoted by $\mathbf{P} \upharpoonright \mathbf{A}$

A chain is a linearly ordered set. Calligraphic letters $\mathcal{A}$, $\mathcal{B}$ range over chains; corresponding bold upper-case letters denote the domains of chains and bold lower-case letters $\mathbf{t}$, $\mathbf{t}'$ range over the domain of a chain. In particular, $\mathcal{N}$ and $\mathcal{R}$ are the chains of natural and real numbers, respectively. By an abuse of notation we will use $\{0, 1, \ldots, k\}$ both for the subset of natural numbers and for the corresponding subchain of $\mathcal{N}$. Similarly, for $\mathbf{t}_1, \mathbf{t}_2 \in \mathbf{R}$ we use $[\mathbf{t}_1, \mathbf{t}_2)$ both for the subinterval and for the subchain of reals.

## 2.2 A variant of monadic theory of order

It will be slightly more convenient for us to deal with the following variant of second-order monadic logic of order. The language (of this variant) of monadic second-order theory of order (see e.g. [8]) has two types of variables: (1) first-order (or individual) variables, and (2) $\Sigma$-predicate (or $\Sigma$-color) variables for each finite non-empty set $\Sigma$.

The letter $t$ with subscripts and superscripts ranges over individual variables; and upper-case case Latin letters are used for predicate variables. The atomic formulas are formulas of the form: $t_1 < t_2$ and $X(t) = \sigma$, where $X$ is a $\Sigma$-predicate variable and $\sigma$ is an element of $\Sigma$. The formulas are constructed from atomic formulas by logical connectives and first and second-order quantifiers. We write $\phi(X_1, \ldots, X_n,\ t)$ to indicate that the free variables of $\phi$ are among $\{X_1, \ldots, X_n,\ t\}$.

The formulas are interpreted over chains. When a formula is interpreted over a chain $\mathcal{A} = \langle \mathbf{A}, < \rangle$, its individual variables range over the elements of $\mathbf{A}$ and the $\Sigma$-predicates variables range over the functions in $\mathbf{A} \to \Sigma$.

We write $\mathcal{A},\ \mathbf{P_1},\ \mathbf{P_2},\ \mathbf{t} \models \phi(X_1,\ X_2,\ t)$ if the assignment of $\mathbf{P}_i$ to $X_i$ and $\mathbf{t}$ to $t$ satisfies $\phi(X_1,\ X_2,\ t)$ in the chain $\mathcal{A}$. When there is no confusion we use $\mathcal{A} \models \phi(\mathbf{P}_1,\ \mathbf{P}_2,\ \mathbf{t})$ for $\mathcal{A},\ \mathbf{P}_1, \mathbf{P_2},\ \mathbf{t} \models \phi(X_1,\ X_2,\ t)$.

We use standard abbreviations, e.g., for a boolean predicate variable $X$, we write $X(t)$ for $X(t) = TRUE$, for boolean predicates $X,\ X'$ we write $X \subset X'$ for $\forall t.\ X(t) \to X'(t)$ and for $\{\sigma_1, \ldots, \sigma_k\}$-predicate $Y$ we write $Y(t_1) = Y(t)$ for $(Y(t_1) = \sigma_1 \wedge Y(t) = \sigma_1) \vee \ldots \vee (Y(t_1) = \sigma_k \wedge Y(t) = \sigma_k)$.

## 2.3 Method of Interpretation

The following theorem is immediately obtained by interpreting a subchain $\mathcal{A}$ of $\mathcal{B}$ in the chain $\mathcal{B}$ augmented by a unary predicate $\mathbf{A}$ (see [5] for the detailed description of the methods of interpretation).

**Theorem 5** *Let $\mathcal{B} = \langle \mathbf{B}, < \rangle$ be a chain and $\mathcal{A} = \langle \mathbf{A}, < \rangle$ be a subchain of $\mathcal{B}$. Let $\phi^A(X_1, \ldots, X_n,\ A)$ be the formula constructed by relativizing all first-order quantifiers of a formula $\phi(X_1, \ldots, X_n)$ to a boolean predicate variable $A$ (see Fig. 1). Then*

1. *$\mathcal{B},\ \mathbf{P}_1, \ldots, \mathbf{P}_n,\ \mathbf{A} \models \phi^A(X_1, \ldots, X_n,\ A)$ iff $\mathcal{A},\ \mathbf{P}_1 \restriction \mathbf{A}, \ldots, \mathbf{P}_n \restriction \mathbf{A} \models \phi(X_1, \ldots, X_n)$.*

**Input:** A formula $\phi$ and a monadic boolean predicate variable $A$ that does not occurs in $\phi$.

**Output:** A formula $\phi^A$ whose free variables are those of $\phi$ plus $A$.

$\phi^A$ is defined inductively on the structure of $\phi$ by the following rules:

1. If $\phi$ is without quantifiers, then $\phi^A = \phi$
2. If $\phi = \phi_1 \wedge \phi_2$, or $\phi = \phi_1 \vee \phi_2$, or $\phi = \neg\phi_1$ then $\phi^A = \phi_1^A \wedge \phi_2^A$, or $\phi^A = \phi_1^A \vee \phi_2^A$ or $\phi^A = \neg\phi_1^A$ , respectively.
3. If $\phi = \exists X.\phi_1$ or $\phi = \forall X.\ \phi_1$, where $X$ is a predicate variable, then $\phi^A = (\exists X.\phi_1^A)$ or $\phi^A = (\forall X.\phi_1^A)$, respectively.
4. If $\phi = \exists t.\phi_1$ or $\phi = \forall t.\ \phi_1$, where $t$ is a first-order variable, then $\phi^A = \exists t.(A(t) \wedge \phi_1^A)$ or $\phi^A = \forall t.(A(t) \rightarrow \phi_1^A)$, respectively.

Figure 1: Relativizing all first-order variables of $\phi$ to $A$

*2. If $X_1, \ldots, X_n$ are boolean predicate variables then*

$$\mathcal{B},\ \mathbf{P}_1, \ldots, \mathbf{P}_n, \mathbf{A} \models \phi^A(X_1, \ldots, X_n,\ , A) \wedge X_1 \subset A \wedge \ldots \wedge X_n \subset A \text{ iff } \mathcal{A},\ \mathbf{P}_1, \ldots, \mathbf{P}_n \models \phi(X_1, \ldots, X_n)$$

**Corollary 6** *If a family of point-sets is definable in $\mathcal{A}$, then it is definable in $\mathcal{A}$ with $\mathcal{B}$ at the background.*

## 2.4 Coloring Indices Theorem

**Definition 1** *(Partition) Formulas $\phi_1(X_1, \ldots, X_n)$, $\ldots$, $\phi_k(X_1, \ldots, X_n)$ form a partition if $\phi_1 \vee \phi_2 \vee \ldots \vee \phi_k$ is valid and for all $i \neq j$ the formulas $\phi_i \wedge \phi_j$ are unsatisfiable.*

We often say "a set $\{\phi_1, \ldots, \phi_k\}$ is a partition" instead of "formulas $\phi_1, \ldots, \phi_k$ form a partition".

**Definition 2** *(Lexicographic Sum) The lexicographic sum of (disjoint) chains $\mathcal{A}_i$ with respect to a chain Ind (notation $L\Sigma\langle\mathcal{A}_i : i \in Ind\rangle$) is the chain $\mathcal{A} = \langle\mathbf{A}, <\rangle$, where $\mathbf{A} = \cup_{i \in Ind}\mathbf{A}_i$ and point $\mathbf{t} \in \mathbf{A}_i$ precedes point $\mathbf{t}' \in \mathbf{A}_j$ in $\mathcal{A}$ if $i$ precedes $j$ in Ind or $i = j$ and $\mathbf{t}$ precedes $\mathbf{t}'$ in $\mathcal{A}_i$.*

We write $\mathcal{A}_0 + \mathcal{A}_1$ for $L\Sigma\langle \mathcal{A}_i : i \in \{0,\ 1\}\rangle$. We refer to $\mathcal{A}_i$ as the summand chains and to $Ind$ as the indices chain of $L\Sigma\langle \mathcal{A}_i : i \in Ind\rangle$.

**Definition 3** *(Induced predicate or induced coloring of indices) Let $\phi_1(X_1,\ldots,X_n)$, $\ldots,\phi_k(X_1,\ldots,X_n)$ be a partition. Let $\mathbf{P}_1,\ldots,\mathbf{P}_n$ be predicates over $L\Sigma\langle \mathcal{A}_i : i \in Ind\rangle$. The predicate $\mathbf{Q} : Ind \to \{1,\ldots,k\}$ is said to be induced by $\mathbf{P}_1,\ldots,\mathbf{P}_n$ with respect to the partition $\phi_1,\ldots,\phi_k$ (and $L\Sigma\langle \mathcal{A}_i : i \in Ind\rangle$) if $\mathbf{Q}$ is defined as follows:*

$$\mathbf{Q}(i) = j \text{ if and only if } \mathcal{A}_i,\ \mathbf{P}_1 \restriction \mathbf{A}_i,\ldots,\mathbf{P}_n \restriction \mathbf{A}_i \models \phi_j(X_1,\ldots,X_n).$$

*We often refer to $\mathbf{Q}$ defined as above as coloring of indices induced by $\mathbf{P}_1,\ldots,\mathbf{P}_n$ with respect to the partition $\phi_1,\ldots,\phi_k$.*

Note that the requirement that $\phi_1,\ldots,\phi_k$ is a partition guarantees that $\mathbf{Q}$ is well defined.

The following theorem is a weak version of Composition Theorem (Shelah [6]). Its proof will be sketched in the Appendix.

**Theorem 7** *(Colored Indices Theorem.) For every formula $\chi(X_1,\ldots,X_n)$ there exist a partition $\phi_1(X_1,\ldots,X_n),\phi_2(X_1,\ldots,X_n),\ldots,\phi_k(X_1,\ldots,X_n)$ and a formula $\phi(Y)$ such that $L\Sigma\langle A_i : i \in Ind\rangle, \mathbf{P}_1,\ldots\mathbf{P}_n \models \chi(X_1,\ldots,X_n)$ if and only if $Ind,\ \mathbf{Q} \models \phi(Y)$ where $\mathbf{Q}$ is induced by $\mathbf{P}_1,\ldots,\mathbf{P}_n$ with respect to the partition $\phi_1,\ldots,\phi_k$. Moreover, there exists an algorithm that constructs $\phi_1,\phi_2,\ldots,\phi_k$ and $\phi$ from $\chi$.*

The next lemma follows from the colored indices theorem and will be referred later.

**Lemma 8** *For every chain $\mathcal{B}$ and for every formula $\chi(X_1,\ldots,X_n)$ ($X_i$ are boolean predicate variables) there exist formulas $\chi^r(X_1,\ldots,X_n)$ and $\chi^l(X_1,\ldots,X_n)$ such that for every chain $\mathcal{A}$=$\langle \mathbf{A},<\rangle$ and subsets $\mathbf{P}_1,\ldots,\mathbf{P}_n$ of $\mathbf{A}$*

1. *$\mathcal{A}+\mathcal{B},\ \mathbf{P}_1,\ldots,\mathbf{P}_n \models \chi(X_1,\ldots,X_n)$ iff $\mathcal{A},\ \mathbf{P}_1,\ldots,\mathbf{P}_n \models \chi^r(X_1,\ldots,X_n)$*
2. *$\mathcal{B}+\mathcal{A},\ \mathbf{P}_1,\ldots,\mathbf{P}_n \models \chi(X,\ldots,X_n)$ iff $\mathcal{A},\ \mathbf{P}_1,\ldots\mathbf{P}_n \models \chi^l(X,\ldots,X_n)$*

*Proof:* Below only the case when $\chi$ has only one free predicate variable $X$ is considered and formula $\chi^r(X)$ is constructed. The case with $n$ free variables and the construction of $\chi^l$ is similar.

By the colored indices theorem there is a partition $\phi_1(X), \ldots, \phi_k(X)$ and a formula $\phi(Y)$ such that

$$\mathcal{A} + \mathcal{B},\ \mathbf{P} \models \chi(X) \text{ iff } \{0,\ 1\}, \mathbf{Q} \models \phi(Y),$$

where $\mathbf{Q}(0) = i$ iff $\mathcal{A}, \mathbf{P} \models \phi_i(X)$ and $\mathbf{Q}(1) = j$ iff $\mathcal{B}, \emptyset \models \phi_j(X)$.

Observe that every $\phi(Y)$ is equivalent over $\{0,\ 1\}$ to a quantifier free formula $\phi'(Y)$ of the form $\bigvee_{i=1}^{n}(Y(0) = m_i \wedge Y(1) = k_i)$.

Let $j_{\mathcal{B}}$ be such that $\mathcal{B}, \emptyset \models \phi_{j_{\mathcal{B}}}(X)$ (such $j$ is unique because $\phi_1, \ldots, \phi_k$ is a partition).

Let $\chi^r(X)$ be obtained from $\phi'(Y)$ as follows:

Step 1 - Eliminate $Y(1)$: Replace $Y(1) = k_i$ by TRUE if $k_i = j_{\mathcal{B}}$ and by FALSE otherwise.

Step 2 - Eliminate $Y(0)$: Replace $Y(0) = m_i$ by $\phi_{m_i}(X)$

It is clear that the constructed formula $\chi^r(X)$ satisfies the conclusion of the lemma. □

**Remark 9** *Note that the construction of $\chi^r$ (or $\chi^l$) from $\chi$ is not necessarily recursive, because Step 1 that eliminates occurrences of $Y(1)$ is not necessary recursive (this step is recursive whenever satisfiability of monadic formulas over $B$ is decidable). All other steps in the construction of $\chi^r$ from $\chi$ are recursive.*

*Let us demonstrate that if the monadic theory of $\mathcal{B}$ is undecidable then $\chi^r$ cannot be recursive in $\chi$. Take any sentence $\phi$ and an individual variable $t_0$ that does not occur in $\phi$. Let $\phi^*(t_0)$ be obtained from $\phi$ by relativizing its first-order quantifiers as follows: replace "$\exists t$" by "$\exists t > t_0$" and "$\forall t$" by "$\forall t > t_0$". Observe that $\mathcal{B} \models \phi$ iff $\{0\} + \mathcal{B} \models \exists t_0.(\phi^*(t_0) \wedge \neg\exists t.t < t_0)$. Hence, the translation of $X \subset X \wedge \exists t_0.(\phi^*(t_0) \wedge \neg\exists t.t < t_0)$ is equivalent to $TRUE$ over one point chain $\{0\}$ if and only if $\mathcal{B} \models \phi$. Since, by the assumption, the satisfiability problem for $\mathcal{B}$ is undecidable and the equivalence problem over one element chain is decidable, it follows that there is no recursive translation.*

# 3 Definability in $\mathcal{N}$ with the reals at the background

**Theorem 10** *For every $\chi(X,\ A)$ there exists $\psi(X)$ such that*

$$\mathcal{R},\ \mathbf{P},\mathbf{N} \models \chi(X,\ A) \wedge X \subset A \textit{ iff } \mathcal{N},\ \mathbf{P} \models \psi(X)$$

*Proof:* Below we show that for every $\chi(X,\ A)$ there exists $\psi(X)$ such that

$$\mathcal{R}^{\geq 0},\ \mathbf{P},\mathbf{N} \models \chi(X,\ A) \wedge X \subset A \text{ iff } \mathcal{N},\ \mathbf{P} \models \psi(X) \tag{1}$$

Theorem 10 follows immediately from (1) and Lemma 8(2).

First note that $\mathcal{R}^{\geq 0}$ is the lexicographic sum $L\Sigma\langle [i,\ i+1) :\ i \in \mathbf{N}\rangle$. Therefore, by the colored indices theorem there is a partition $\phi_1(X,\ A), \ldots, \phi_k(X,\ A)$ and a formula $\phi(Y)$ such that

$$\begin{array}{c}\mathcal{R}^{\geq 0},\ \mathbf{P},\mathbf{N} \models \chi(X,\ A) \wedge X \subset A \text{ iff } \mathcal{N},\ \mathbf{Q} \models \phi(Y), \text{ where} \\ \mathbf{Q} \text{ is induced by } \mathbf{P},\mathbf{N} \text{ wrt } \phi_1(X,\ A), \ldots, \phi_k(X,\ A).\end{array} \tag{2}$$

Observe that if $\mathcal{R}^{\geq 0},\ \mathbf{P},\mathbf{N} \models \chi(X,\ A) \wedge X \subset A$ then $\mathbf{P} \subseteq \mathbf{N}$. Therefore, $\mathbf{P} \cap [i,\ i+1)$ is either the empty set or the singleton set $\{i\}$.

Since, $\phi_1(X,\ A), \ldots, \phi_k(X,\ A)$ is a partition there are $j_0$ and $j_1$ such that

$$[i,\ i+1),\ \emptyset,\ \{i\} \models \phi_{j_0}(X,\ A) \text{ and } [i,\ i+1),\ \{i\},\ \{i\} \models \phi_{j_1}(X,\ A) \tag{3}$$

For $\mathbf{P} \subseteq \mathbf{N}$ define $\mathbf{Q} : \mathbf{N} \rightarrow \{1, \ldots, k\}$ as follows:

$$\mathbf{Q}(i) = \begin{cases} j_0 & \text{if } \mathbf{P} \cap [i,\ i+1) = \emptyset \\ j_1 & \text{if } \mathbf{P} \cap [i,\ i+1) = \{i\} \end{cases} \tag{4}$$

Suppose that $\mathbf{P},\ \mathbf{N}$ satisfy $\chi(X,\ A) \wedge X \subset A$. Then $\mathbf{Q}$ is induced by $\mathbf{P},\ \mathbf{N}$ wrt $\phi_1, \ldots, \phi_k$ if and only if (4) holds.

Let $\phi^*(Y)$ be the formula which is obtained from $\phi(Y)$ when the sub-formulas $Y(t) = m$ are replaced by $FALSE$ for $m \notin \{j_0,\ j_1\}$.

$$\forall i.\ \mathbf{Q}(i) \in \{j_0,\ j_1\} \text{ implies that } \mathcal{N},\ \mathbf{Q} \models \phi(Y) \text{ iff } \mathcal{N},\ \mathbf{Q} \models \phi^*(Y) \qquad (5)$$

Let $\psi(X)$ be obtained from $\phi^*(Y)$ by eliminating $Y(t)$ as follows: replace $Y(t) = j_0$ by $\neg X(t)$ and replace $Y(t) = j_1$ by $X(t)$. It is clear that for every $\mathbf{P} \subset \mathbf{N}$ and for $\mathbf{Q}$ defined as in (4)

$$\mathcal{N}, \mathbf{Q} \models \phi^*(Y) \text{ iff } \mathcal{N},\ \mathbf{P} \models \psi(X) \qquad (6)$$

Finally, to complete the proof observe that (1) follows from (2)-(6). □

**Remark 11** *(Generalization) In the above proof we used the following property of subintervals of reals:*

*Property 1: If $a_1 < b_1$ and $a_2 < b_2$ then the intervals $(a_1,\ b_1)$ and $(a_2,\ b_2)$ have the same monadic theory.*

*We used also the following property of the natural numbers:*

*Property 2: For every $a$ there exists $b > a$ such that no $c$ lies between $a$ and $b$.*

*The same proof shows the following*

***Theorem 12*** *Let $\mathcal{B}$ be a chain that satisfies Property 1 and let $\mathcal{A}$ be a subchain of $\mathcal{B}$, which satisfies Property 2. Suppose that for every $\mathbf{A}' \subset \mathbf{A}$*

$$inf\{\mathbf{t}\ :\mathbf{t} \in \mathbf{B} \text{ and } \forall \mathbf{a}' \in \mathbf{A}'.\mathbf{t} \geq \mathbf{a}'\} = inf\{\mathbf{t}\ :\mathbf{t} \in \mathbf{A} \text{ and } \forall \mathbf{a}' \in \mathbf{A}'.\mathbf{t} \geq \mathbf{a}'\}.$$

*Then a set of predicates is definable in $\mathcal{A}$ if it is definable in $\mathcal{A}$ with $\mathcal{B}$ at the background.*

*Note that the set of rationals and the set of irrationals satisfy Property 1. Property 2 is satisfied by every well founded chain.*

*Let us illustrate that Property 1 is essential. Let $\mathbf{P}$ be a non-recursive subset of natural numbers. Let $\mathbf{R}'$ be obtained from $\mathbf{R}$ (the reals) by removing $i + 1/2$ for* *all $i$ in $\mathbf{P}$. Then, having $\mathbf{R}'$ at the background allows one to define $\mathbf{P}$ which is not definable in $\mathcal{N}$.*

*Similarly, the requirement that $inf\{\mathbf{t} \ : \mathbf{t} \in \mathbf{B}$ and $\forall \mathbf{a}' \in \mathbf{A}'.\mathbf{t} \geq \mathbf{a}'\} = inf\{\mathbf{t} \ : \mathbf{t} \in \mathbf{A}$ and $\forall \mathbf{a}' \in \mathbf{A}'.\mathbf{t} \geq \mathbf{a}'\}$ for every $\mathbf{A}' \subset \mathbf{A}$ is essential. Let $\mathbf{P}$ be a non-recursive subset of the positive natural numbers. Let $\mathbf{A} = \mathbf{N} \cup \{k - \frac{1}{2^{n+2}} \ : \ 0 < n \in \mathbf{N}$ and $0 < k \in \mathbf{P}\} \cup \{k + \frac{1}{2} - \frac{1}{2^{n+2}} \ : \ 0 < n \in \mathbf{N}$ and $k \notin \mathbf{P}\}$. Notice that the order type of $\mathcal{A}$ is $\omega^2$. It is easy to check that $\mathbf{P}$ is not definable in $\mathcal{A}$ but it is definable in $\mathcal{A}$ with $\mathcal{R}$ at the background.*

# 4 Definability in closed subchains of Reals

A subchain $\mathcal{F} = \langle \mathbf{F}, < \rangle$ of the reals is said to be closed if $\mathbf{F}$ is a closed subset of the reals. The main result of this section is

**Theorem 13** *For every formula $\chi(X, \ A)$ there exists a formula $\chi^c(X)$ such that for every closed subchain $\mathcal{F} = \langle \mathbf{F}, < \rangle$ of the reals a set of predicates is definable by $\chi(X, \ A)$ in $\mathcal{F}$ with $\mathcal{R}$ at the background if and only if it is definable by $\chi^c(X)$ in $\mathcal{F}$.*

This theorem immediately implies Theorem 3 stated in the Introduction.

In the first subsection the representation of closed subchains of reals as a lexicographic sum is given. Relying on this representation, we provide a proof of Theorem 13 in the second subsection.

## 4.1 Representation of closed sets as a lexicographic sum

Let $\mathbf{F}$ be a subset of $\mathbf{R}$. Real numbers $\mathbf{t}$ and $\mathbf{t}'$ are said to be $\mathbf{F}$-equivalent (notations $\mathbf{t} \sim_{\mathbf{F}} \mathbf{t}'$) if one of the following conditions holds: Let $\mathbf{t}_1 = min(\mathbf{t}, \ \mathbf{t}')$ and let $\mathbf{t}_2 = max(\mathbf{t}, \ \mathbf{t}')$. Then

1. $[\mathbf{t}_1, \ \mathbf{t}_2] \subseteq \mathbf{F}$.
2. $[\mathbf{t}_1, \ \mathbf{t}_2] \subseteq \mathbf{R} \setminus \mathbf{F}$.
3. There exists $\mathbf{t}_3 \in [\mathbf{t}_1, \ \mathbf{t}_2)$ such that $[\mathbf{t}_1, \ \mathbf{t}_3] \subseteq \mathbf{F}$ and $(\mathbf{t}_3, \ \mathbf{t}_2] \subseteq \mathbf{R} \setminus \mathbf{F}$.

**Example 1** *Let* $\mathbf{F}$ *be* $\{0\} \cup \{\frac{1}{i} \ : \ i \text{ positive integer}\}$. *Then* $\mathbf{t}_1 \sim_{\mathbf{F}} \mathbf{t}_2$ *if* $\mathbf{t}_1 = \mathbf{t}_2 = 0$ *or* $\mathbf{t}_1, \ \mathbf{t}_2 \geq 1$ *or* $\mathbf{t}_1, \mathbf{t}_2 < 0$ *or* $\mathbf{t}_1, \mathbf{t}_2 \in [\frac{1}{i+1}, \ \frac{1}{i})$.

It is clear that $\sim_{\mathbf{F}}$ is an equivalence relation.

Let $\mathbf{I_t}$ be the set of all reals equivalent to $\mathbf{t}$ and let $\mathbf{F_t}$ be the set of all points in $\mathbf{F}$ equivalent to $\mathbf{t}$.

Suppose that $\mathbf{F}$ is a closed subset of the reals. Observe that if there exists $\mathbf{t}_1 \in \mathbf{F}$ such that $\mathbf{t}_1 < \mathbf{t}$ then there exists $\mathbf{t}_2 \in \mathbf{F}$ such that $\mathbf{t} \in \mathbf{I}_{\mathbf{t}_2}$; the desired $\mathbf{t}_2 = sup\{\mathbf{t}_1 \in \mathbf{F} \ : \mathbf{t}_1 < t\}$. It is also easy to see that for every $\mathbf{t} \in \mathbf{F}$ either $\mathbf{I_t} = \mathbf{F_t}$ or there are $\mathbf{t}_1 < \mathbf{t}_2$ both in $\mathbf{F}$ such that (1) $\mathbf{t}_1 = sup(\mathbf{F_t})$; (2) $(\mathbf{t}_1, \ \mathbf{t}_2) \cap \mathbf{F} = \emptyset$ and (3) $\mathbf{I_t} = \mathbf{F_t} \cup (\mathbf{t}_1, \ \mathbf{t}_2)$.

Let $\mathcal{I}_\mathbf{t}$ and $\mathcal{F}_\mathbf{t}$ be the subchains of reals over the sets $\mathbf{I_t}$ and $\mathbf{F_t}$. The above observations imply

**Lemma 14** *Assume that* $\mathcal{F} = \langle \mathbf{F}, \ < \rangle$ *is a subchain of* $\mathcal{R}$ *and* $\mathbf{F}$ *is a closed subset of reals. Then there exists a subchain* $Ind$ *of* $\mathcal{F}$ *such that*

1. $\mathcal{F} = L\Sigma\langle \mathcal{F}_i \ : \ i \in Ind \rangle$
2. $\mathcal{R} = \begin{cases} L\Sigma\langle \mathcal{I}_i \ : \ i \in Ind\rangle & \text{If there exists no minimal element in } F \\ (\infty, \ \mathbf{t}) + L\Sigma\langle \mathcal{I}_i \ : \ i \in Ind\rangle & \text{If } \mathbf{t} \text{ is the minimal element of } F \end{cases}$
3. *For every* $i \in Ind$ *either* $\mathcal{I}_i = \mathcal{F}_i$ *or there exists an open subinterval* $\mathcal{B}_i$ *of reals such that* $\mathcal{I}_i = \mathcal{F}_i + \mathcal{B}_i$.
4. *There are formulas* $full(t)$, $equiv(t_1, \ t_2)$ *(these formulas are independent from* $\mathcal{F}$*) such that*
   (a) $\mathcal{F} \models full(\mathbf{t})$ *iff* $\mathcal{F}_\mathbf{t} = \mathcal{I}_\mathbf{t}$.
   (b) $\mathcal{F} \models equiv(\mathbf{t}_1, \ \mathbf{t}_2)$ *if* $\mathbf{t}_1 \sim_F \mathbf{t}_2$.

## 4.2 Proof of Theorem 13

Let us first generalize Lemma 8 as follows:

**Lemma 15** *For every chain* $\mathcal{B}$ *there exist functions* $Right_\mathcal{B}$ *and* $Left_\mathcal{B}$ *that map formulas with free boolean predicate variables* $\{X_1, \dots X_n, A\}$ *to formulas with free variable* $\{X_1, \dots, X_n\}$ *and satisfy the following conditions: for every chain* $\mathcal{A}$

*1.* $\chi^*(X_1,\ldots,X_n) = Right_{\mathcal{B}}(\chi(A,\ X_1,\ldots,X_n))$ *implies*

$$\mathcal{A}+\mathcal{B},\ \mathbf{A},\ \mathbf{P}_1,\ldots,\mathbf{P}_n \models \chi(A,\ X_1,\ldots,X_n)\wedge X_1\subset A\ldots\wedge X_n\subset A \text{ iff}$$
$$\mathcal{A},\ \mathbf{P}_1,\ldots\mathbf{P}_n \models \chi^*(X_1,\ldots,X_n).$$

*2.* $\chi^*(X_1,\ldots,X_n) = Left_{\mathcal{B}}(\chi(A,\ X_1,\ldots,X_n))$ *implies*

$$\mathcal{B}+\mathcal{A},\ \mathbf{A},\ \mathbf{P}_1,\ldots,\mathbf{P}_n \models \chi(A,\ X_1,\ldots,X_n)\wedge X_1\subset A\wedge\ldots\wedge X_n\subset A \text{ iff}$$
$$\mathcal{A},\ \mathbf{P}_1,\ldots,\mathbf{P}_n \models \chi^*(X_1,\ldots,X_n).$$

*3. Moreover, if* $\mathcal{B}_1$ *and* $\mathcal{B}_1$ *have the same monadic theory then* $Right_{\mathcal{B}_1} = Right_{\mathcal{B}_2}$ *and* $Left_{\mathcal{B}_1} = Left_{\mathcal{B}_2}$.

*Proof:* Below we explain only how $Right_{\mathcal{B}}$ maps the formulas with free variables $\{X,\ A\}$. The case when the set of free variables is $\{X_1,\ldots X_n,\ A\}$ and the construction of $Left_{\mathcal{B}}$ is similar.

Step 1. Apply the colored indices theorem and find a partition $\phi_1(A,\ X)$, $\ldots,\phi_k(A,\ X)$ and a formula $\phi(Y)$ such that

$$\mathcal{A}+\mathcal{B},\ \mathbf{A},\ \mathbf{P}\models\chi(A,\ X)\wedge X\subset A \text{ if and only if}$$
$$\{0,1\},\ \mathbf{Q}\models\phi(Y),\ \text{where } \mathbf{Q} \text{ is induced by } \mathbf{A},\mathbf{P}.$$

Step 2. Let $\phi'(Y)$ be a formula $\bigvee_i^n(Y(0)=m_i\wedge Y(1)=k_i)$ which is equivalent to $\phi(Y)$ over $\{0,1\}$.

Step 3. Let $j_{\mathcal{B}}$ be such that $\mathcal{B},\emptyset,\emptyset\models\phi_{j_{\mathcal{B}}}(A,\ X)$.

Step 4. Let $\phi^*(Y)$ be obtained from $\phi'(Y)$ by eliminating $Y(1)$ as follows: replace $Y(1)=m$ by $TRUE$ for $m=j_{\mathcal{B}}$ and by $FALSE$ otherwise.

Step 5. Let $\phi_i^*(X)$ be obtained from $\phi_i(A,\ X)$ when $A(t)$ are replaced by $TRUE$.

Step 6. Let $\chi^*(X)$ be obtained from $\phi^*(Y)$ by eliminating $Y(0)$ as follows: replace $Y(0)=m$ by $\phi_m^*(X)$.

Observe that only step 4, depends on $\mathcal{B}$ and for chains $\mathcal{B}_1$ and $\mathcal{B}_2$ with the same monadic theory, the same formulas are constructed.

Let us check that the constructed formula $\chi^*(X)$ satisfies the the conclusion of the lemma.

*The if direction:* Assume that $\mathcal{A}, \mathbf{P} \models \chi^*(X)$. Let $\mathbf{Q}$ be induced by $\mathbf{A}, \mathbf{P}$ wrt $\phi_1, \ldots, \phi_k$ and $\mathcal{A} + \mathcal{B}$.. Then

$$\mathbf{Q}(1) = j_{\mathcal{B}}, \quad \text{(because } \mathbf{P} \subset \mathbf{A}) \tag{7}$$

and $\mathbf{Q}(0) = i$ iff $\mathcal{A},\ \mathbf{A},\ \mathbf{P} \models \phi_i(A,\ X)$.

By construction of step 5 it follows that

$$\mathbf{Q}(0) = m \text{ if and only if } \mathcal{A},\ \mathbf{P} \models \phi^*_m(X) \tag{8}$$

From (8), (7) and the constructions in steps 4 and 6, it follows that

$$\{0,\ 1\},\ \mathbf{Q} \models \phi'(Y) \tag{9}$$

Hence, by step 2

$$\{0,\ 1\},\ \mathbf{Q} \models \phi(Y) \tag{10}$$

and by step 1,

$$\mathcal{A} + \mathcal{B},\ \mathbf{A},\ \mathbf{P} \models \chi(A,\ X) \wedge X \subset A$$

*The only-if direction:* Assume that $\mathcal{A} + \mathcal{B},\ \mathbf{A},\ \mathbf{P} \models \chi(A,\ X) \wedge X \subset A$.

Let $\mathbf{Q}$ be induced by $\mathbf{A}, \mathbf{P}$. Then

$$\mathbf{Q}(1) = j_{\mathcal{B}}, \text{ (because } \mathbf{P} \subset \mathbf{A}) \text{ and} \tag{11}$$

$$\mathbf{Q}(0) = m \text{ iff } \mathcal{A},\ \mathbf{A},\ \mathbf{P} \models \phi_m(A,\ X). \tag{12}$$

$$\{0,\ 1\},\ \mathbf{Q} \models \phi(Y), \text{ by the colored indices theorem} \tag{13}$$

From (11), (12), (13) and the construction of $\chi^*$ it follows that $\mathcal{A}, \mathbf{P} \models \chi^*(X)$.
□

Let us proceed now with the proof of Theorem 13.

By Lemma 8, one can construct for every formula $\theta(X_1,\ X_2)$ a formula $\theta^l(X_1,\ X_2)$ such that for all $\mathbf{t}$ and all subsets $\mathbf{P}_1,\ \mathbf{P}_2$ of $[\mathbf{t}, \infty)$

$$(-\infty,\ \mathbf{t}) + [\mathbf{t}, \infty), \mathbf{P}_1, \mathbf{P}_2 \models \theta(X_1,\ X_2) \wedge X_1 \subset X_2 \text{ if and only if}$$

$$[\mathbf{t}, \infty), \mathbf{P}_1, \mathbf{P}_2 \models \theta^l(X_1,\ X_2) \wedge X_1 \subset X_2$$

Hence, if a set $\mathbf{F}$ has a minimal element $\mathbf{t}$, then for every $\chi(X,\ A)$ there exists $\chi^{lb}(X,\ A)$ such that

$$\mathcal{R}, \mathbf{P}, \mathbf{F} \models \chi(X,\ A) \wedge X \subset A \text{ iff } [\mathbf{t}, \infty), \mathbf{P}, \mathbf{F} \models \chi^{lb}(X,\ A) \wedge X \subset A \tag{14}$$

Let *lbound* be the sentence $\exists t \forall t'.t' \geq t$ and let $\psi(X,\ A)$ be defined as $(\neg lbound \rightarrow \chi(X,\ A)) \wedge (lbound \rightarrow \chi^{lb}(X,\ A))$

From (14), definition of $\psi$ and Lemma 14(2) it follows that for every closed $\mathbf{F}$

$$\begin{aligned} \mathcal{R}, \mathbf{P}, \mathbf{F} \models \chi(X,\ A) \wedge X \subset A \text{ iff} \\ L\Sigma\langle \mathcal{I}_i\ :\ i \in Ind\rangle, \mathbf{P}, \mathbf{F} \models \psi(X,\ A) \wedge X \subset A. \end{aligned} \tag{15}$$

where $Ind$ and $\mathcal{I}_i$ are as in Lemma 14.

Below we are going to construct $\chi^c(X)$ from $\psi(X,\ A)$ such that

$$L\Sigma\langle \mathcal{I}_i\ :\ i \in Ind\rangle,\ \mathbf{P}, \mathbf{F} \models \psi(X,\ A) \wedge X \subset A \text{ iff } \mathcal{F},\ \mathbf{P} \models \chi^c(X) \tag{16}$$

Theorem 13 is immediately obtained from (15) and (16).

Formula $\chi^c(X)$ is constructed by the following steps:

Step 1: Apply the colored indexes theorem to formula $\psi(X,\ A) \wedge X \subset A$ and find a partition $\phi_1(X,\ A), \dots, \phi_k(X,\ A)$ and a formula $\phi(Y)$ such that $L\Sigma\langle A_i : i \in Ind\rangle,\ \mathbf{P}_1,\ \mathbf{P}_2 \models \psi(X,\ A) \wedge X \subset A$ if and only if $Ind,\ \mathbf{Q} \models \phi(Y)$ where $\mathbf{Q}$ is induced by $\mathbf{P}_1,\ \mathbf{P}_2$ with respect to the partition $\phi_1(X,\ A), \dots, \phi_k(X,\ A)$.

In particular, let $\mathbf{F}$ be a closed subset of the set $\mathbf{R}$ of reals and let $Ind,\ \mathcal{F}_i$ and $\mathcal{I}_i$ be as in Lemma 14. Then

$$\begin{aligned} L\Sigma\langle \mathcal{I}_i\ : i \in Ind\rangle,\ \mathbf{P},\ \mathbf{F} \models \psi(X,\ A) \wedge X \subset A \text{ if and only if } Ind,\ \mathbf{Q} \models \phi(Y), \\ \text{for } \mathbf{Q} \text{ defined as } \mathbf{Q}(i) = j \text{ if } \mathcal{I}_i,\ \mathbf{P} \restriction \mathcal{I}_i,\ \mathbf{F}_i \models \phi_j(X,\ A). \end{aligned}$$

Step 2: Let $equiv(t_1,\ t_2)$ be the formula from Lemma 14(4). We say that a predicate $\mathbf{Z}$ respects *equiv* if it satisfies the formula

$$Respect(Z) \triangleq \forall t_1.\forall t_2.(equiv(t_1,\ t_2) \rightarrow Z(t_1) = Z(t_2)).$$

Let $\phi'(Y)$ be obtained from $\phi(Y)$ when (A) atomic formulas $t_1 < t_2$ are replaced by $t_1 < t_2 \wedge \neg equiv(t_1,\ t_2)$, and (B) monadic quantifiers are replaced by the quantification over predicates that respects $equiv$, i.e., by replacing "$\forall Z.$" (respectively "$\exists Z.$") by "$\forall Z. Respect(Z) \rightarrow$ " (respectively "$\exists Z. Respect(Z) \wedge$").

Let $\varphi(Y)$ be $\phi'(Y) \wedge Respect(Y)$.

Observe that $\varphi(Y)$ on $\mathcal{F}$ "simulates" $\phi(Y)$ on $Ind$ in the following sense:

$$\text{if } Ind, \mathbf{Q} \models \phi(Y) \text{ and } \forall i \in Ind \forall \mathbf{t} \in \mathbf{F}_i.\ \mathbf{Q}'(\mathbf{t}) = \mathbf{Q}(i) \\ \text{then } \mathcal{F}, \mathbf{Q}' \models \varphi(Y).$$

$$\text{if } \mathcal{F}, \mathbf{Q}' \models \varphi(Y) \text{ and } \mathbf{Q}' \text{ is the restriction of } \mathbf{Q}' \text{ on } Ind \\ \text{then } Ind, \mathbf{Q} \models \phi(Y) \text{ and } \forall i \in Ind \forall \mathbf{t} \in \mathbf{F}_i.\ \mathbf{Q}'(\mathbf{t}) = \mathbf{Q}(i).$$

Step 3: Let full-color$_m(t,\ X)$ be formula $\exists Z.(\phi_m(X,\ Z) \wedge \forall t'.(Z(t') \leftrightarrow equiv(t,\ t'))$.

Suppose that $\mathbf{t} \in \mathbf{F}_i$ and $\mathcal{F}_i = \mathcal{I}_i$. Then

$$\mathcal{F},\ \mathbf{P}, \mathbf{t} \models \text{full-color}_m(t,\ X) \text{ iff } \mathbf{Q}(i) = m,$$

where $\mathbf{Q}$ is the coloring induced by $\mathbf{F}$, $\mathbf{P}$ wrt $\phi_1(X,\ A), \ldots, \phi_k(X,\ A)$ and $L\Sigma\langle \mathcal{I}_i\ :\ i \in Ind\rangle$.

Let $\phi_i^{\neg full}(X,\ Z)$ be obtained from $\phi_i(X,\ A)$ as follows: first let $\phi_i^*(X)$ be the result of applying $Right_{\mathcal{B}}$ to $\phi_i(X,\ A)$, where $\mathcal{B}$ is any open subinterval of reals and $Right_{\mathcal{B}}$ is the function from Lemma 15; second relativize all first-order variables in $\phi_i^*(X)$ to $Z$. (see Fig. 1 in the proof of Theorem 5).

Let notfull-color$_m(t,\ X)$ be formula $\exists Z.(\phi_m^{\neg full}(X,\ Z) \wedge \forall t' Z(t') \leftrightarrow equiv(t,\ t'))$.

Suppose that $\mathbf{t} \in \mathbf{F}_i$ and $\mathcal{F}_i \neq \mathcal{I}_i$. Then

$$\mathcal{F},\ \mathbf{P}, \mathbf{t} \models \text{notfull-color}_m(t,\ X) \text{ iff } \mathbf{Q}(i) = m,$$

where $\mathbf{Q}$ is the coloring induced by $\mathbf{F}$, $\mathbf{P}$ wrt $\phi_1(X,\ A), \ldots \phi_k(X,\ A)$ and $L\Sigma\langle \mathcal{I}_i\ :\ i \in Ind\rangle$.

Step 4: Finally, let $\chi^c(X)$ be obtained from the formula $\varphi(Y)$ constructed in Step 2, by eliminating $Y(t)$ as follows: replace "$Y(t) = m$" by "$(full(t) \rightarrow$ full-color$_m(t,\ X)) \wedge (\neg full(t) \rightarrow$ notfull-color$_m(t,\ X))$, where $full(t)$ is the formulas from Lemma 14(4).

It is a routine task to verify that $\chi^c(x)$ indeed satisfies (16) and this completes the proof of Theorem 13.

**Remark 16** *Slightly modifying the above proof one can show*

***Theorem 17*** *Let $\mathcal{A}_i$, $\mathcal{B}_i$ be chains for every $i$ in a chain $Ind$. Let $\mathcal{C}$ be defined as $L\Sigma\langle\mathcal{A}_i + \mathcal{B}_i \ : \ i \in Ind\rangle$ and let $\mathcal{A} = L\Sigma\langle\mathcal{A}_i \ : \ i \in Ind\rangle$. Suppose that*

1. *All $\mathcal{B}_i$ have the same monadic theory and*
2. *There exists a formula $equiv(t_1, t_2)$ such that $\mathcal{A} \models equiv(\mathbf{t}_1, \mathbf{t}_2)$ iff $\mathbf{t}_1, \mathbf{t}_2 \in \mathcal{A}_i$ for some $i \in Ind$.*

*Then for every formula $\chi(X, \ A)$ there exists a formula $\chi^*(X)$ such that*

$$\mathcal{C}, \ \mathbf{P}, \ \mathbf{A} \models \chi(X, \ A) \wedge X \subset A \ \textit{iff} \ \mathcal{A}, \mathbf{P} \models \chi^*(X).$$

*Assumption 1 can be replaced by a weaker assumption*

3. *There exist a partition $\phi_1(t), \ldots, \phi_k(t)$ and chains $\mathcal{D}_1, \ldots, \mathcal{D}_k$ such that if $\mathcal{A} \models \phi_m(\mathbf{t})$ and $\mathbf{t} \in \mathbf{A}_i$ then $\mathcal{B}_i$ and $\mathcal{D}_m$ have the same monadic theory.*

*Notice that this theorem does not generalize the theorem about closed subchains because every $\mathcal{B}_i$ is non-empty. However, one can formulate a theorem that generalizes both these theorems.*

# References


[1] J. R. Büchi. On a decision method in restricted second order arithmetic In *Proc. International Congress on Logic, Methodology and Philosophy of Science*, E. Nagel at al. eds, Standford University Press, pp 1-11, 1960.

[2] Y. Gurevich, Modest theory of short chains, I, Journal of Symbolic Logic 44 (1979), 481–490.

[3] Y. Gurevich. Monadic second order theories. In J. Barwise and S. Feferman eds., Model theoretical logics, pp. 479-506, Springer Verlag, 1986.

[4] Y. Gurevich and S. Shelah. Monadic theory of order and topology in ZFC. Annals of Mathematical Logic 23 (1982), 179–198.

[5] M. O. Rabin. Decidable theories. In J. Barwise editor *Handbook of Mathematical Logic*, North-Holland, 1977.

[6] S. Shelah. The monadic theory of order. *Ann. of Math.*, **102**, pp 349-419, 1975.

[7] W. Thomas (1997). Ehrenfeucht Games, the composition method, and the monadic theory of ordinal words. In *Structures in Logic and Computer Science: A Selection of Essays in Honor of A. Ehrenfeucht*, *Lecture Notes in Computer Science* **1261**:118-143, Springer-Verlag.

[8] B. A. Trakhtenbrot and Y. M. Barzdin. Finite Automata, North Holland Amsterdam, 1973.

# A A proof of Colored Indices Theorem

We presume that the reader knows subsections 3.1 and 3.2 of [3].

Letter $\xi$ ranges over finite sequences of natural numbers (such sequences are called alternation types in [3]).

In [3], for every $\xi$ and a natural number $l$ a finite set called $\xi - l$- Box is defined. Then to every element $b$ of $\xi - l$- Box a formula $\phi_b(X_1, \dots, X_l)$ is assigned (Theorem 1.2 in [2]). The construction of boxes and the formulas assigned to the elements of the boxes is recursive. It will be convenient to view the elements of every $\xi - l$-Box as ordered in some way and we will refer to an element of a box by its place in this order. Let us mention the following

**Fact 1:**

1. The set of formulas assigned to the elements of $\xi - l$- Box forms a partition.
2. For every formula $\theta(X_1, \dots, X_l)$ whose free variables are among $\{X_1, \dots, X_l\}$ there are $\xi$ and a subset $B$ of $\xi - l$-Box such that $\theta$ is equivalent to the disjunction $\vee_{i \in B} \phi_i(X_1, \dots, X_l)$. Moreover $\xi$ and $B$ are recursive in $\theta$.

**Remark 18** *The above fact is easily extracted from the results in [6, 3, 2] for the standard monadic second-order logic with only boolean predicate variables. It has a proof similar to the proof of Theorem 1.2 in [2] or Lemma 2.1 in [6]. We are not going to reprove it here for the following reasons: (1) The definitions of boxes and the corresponding formulas require a considerable amount of notations which are not needed elsewhere. (2) There is a well-written paper by W. Thomas [7] which provides a clear exposition of the Composition Theorem with many detailed examples; in [7] boxes are called types and the corresponding formulas are called Hintika formulas.*

*Fact 1 also holds for the variant with $\Sigma$-predicate variables; a $\Sigma$-predicate can be coded by a tuple of boolean predicates and the proof for our variant is an immediate adaptation of the proofs for the boolean case.*

Let $\xi$ be a finite sequence of natural numbers, let $l$ be a natural number, $n$ be the number of elements in $\xi - l$- Box and let $\mathbf{X}=\langle \mathbf{X}_1, \dots \mathbf{X}_l\rangle$ be a tuple of predicates on $L\Sigma\langle \mathcal{A}_i \ : \ i \in Ind\rangle$. The $n$-tuple $P(\xi, l, \mathbf{X}) = \langle P_b(\xi, l, \mathbf{X}) \ : \ 0 < b \leq n\rangle$ of boolean predicates on $Ind$ is defined as follows:

$$P_b(i) \text{ holds iff } \mathcal{A}_i,\ \mathbf{X}_1 \restriction \mathbf{A}_i, \dots \mathbf{X}_l \restriction \mathbf{A}_i \models \phi_b(X_1, \dots, X_l).$$

Recall (Definition 3 Section 2.4) that the coloring $\mathbf{Q}$ induced by $\mathbf{X}_1, \dots \mathbf{X}_l$ with respect to partition $\{\phi_b(X_1, \dots, X_l) \ : \ b \in \xi - l - Box\}$ is defined by

$$\mathbf{Q}(i) = b \text{ iff } \mathcal{A}_i,\ \mathbf{X}_1 \restriction \mathbf{A}_i, \dots \mathbf{X}_l \restriction \mathbf{A}_i \models \phi_b(X_1, \dots, X_l).$$

Hence $\mathbf{Q}$ codes the tuple $P(\xi, l, \mathbf{X})$.

**Theorem 19** *(Composition Theorem) There exists a recursive function $H(\xi,\ l)$ that outputs a finite sequence of natural numbers and for every $b \in \xi - l$- Box there is a boolean combination $\phi_b^*(X_1, \dots, X_n)$ of the formulas $\phi_t$, where $n$ is the number of elements in $\xi - l$- Box. and $t$ is in $H(\xi,\ l)$-$n$-Box such that*

$$L\Sigma\langle \mathcal{A}_i \ : \ i \in Ind\rangle, \mathbf{X} \models \phi_b(X_1, \dots, X_l) \textit{ iff}$$
$$Ind, P(\xi, l, \mathbf{X}) \models \phi_b^*(X_1, \dots, X_n).$$

*Moreover, $\phi_b^*(X_1, \dots, X_n)$ is computable from $b$.*

**Remark 20** *Actually, Theorem 3.2.3 in [3] states that (under the above assumption) the set of formulas*

$$\{\phi_b(X_1,\ldots X_l)\ :\ b\in\xi-l-Box\ \textit{and}\ L\Sigma\langle\mathcal{A}_i\ :\ i\in Ind\rangle,\mathbf{X}\models\phi_b(X_1,\ldots,X_l)\}$$

*is recursive in the set*

$$\{\phi_i(X_1,\ldots X_n)\ : i\in H(\xi,l)-n-Box\ \textit{and}\ \ Ind, P(\xi,l,\mathbf{X})\models\phi_i(X_1,\ldots,X_n)\}.$$

*This together with the fact that every box has a finite number of elements imply the conclusion of Theorem 19.*

Finally, Colored Indices Theorem is derived as follows:

Given a formula $\chi(X_1,\ldots,X_l)$. Let $\xi$ and $B\subset\xi-l$-Box be such that $\chi(X_1,\ldots X_l)$ is equivalent to $\vee_{b\in B}\phi_b(X_1,\ldots,X_l)$ (see Fact 1).

Let $n$ be the number of the elements in $\xi-l$-Box and let $\phi^*_b(X_1,\ldots X_n)$ be obtained from $\phi_b(X_1,\ldots,X_l)$ as in Theorem 19.

Let $\phi(Y)$ be obtained from $\vee_{b\in B}\phi^*_b(X_1,\ldots,X_n)$ when the sub-formulas $X_b(t)$ are replaced by $Y(t)=b$.

It is routine to check that the partition $\{\phi_b\ :\ b\in\xi-l-Box\}$ and $\phi(Y)$ satisfy the conclusion of Colored Indices Theorem.