A Model Theoretic Proof of Completeness of an Axiomatization of Monadic Second-Order Logic on Streams

Colin Riba

A Model Theoretic Proof of Completeness of an Axiomatization of Monadic Second-Order Logic on Streams

Colin Riba
ENS de Lyon, Université de Lyon, LIP*
colin.riba@ens-lyon.fr
http://perso.ens-lyon.fr/colin.riba/

April 28, 2012

We discuss the completeness of an axiomatization of Monadic Second-Order Logic (MSO) on infinite words (or *streams*). By using model-theoretic tools, we give an alternative proof of D. Siefkes' result that a fragment with full comprehension and induction of second-order Peano's arithmetic is complete w.r.t. the validity of MSO-formulas on streams. We rely on Feferman-Vaught Theorems and the Ehrenfeucht-Fraïssé method for Henkin models of second-order arithmetic. Our main technical contribution is an infinitary Feferman-Vaught Fusion of such models. We show it using Ramseyan factorizations similar to those for standard infinite words. We also discuss a Ramsey's theorem for MSO-definable colorings, and show that in linearly ordered Henkin models, Ramsey's theorem for additive MSO-definable colorings implies Ramsey's theorem for all MSO-definable colorings.

*UMR 5668 CNRS ENS Lyon UCBL INRIA

Contents

1. Introduction

We discuss the completeness of an axiomatization of Monadic Second-Order Logic (MSO) on infinite words (or *streams*). MSO on streams is known to be decidable since the celebrated work of Büchi [Büc62]. The usual route is to translate MSO-formulas to finite state automata running on streams. Such automata provide an established framework for the specification and verification of non-terminating programs, while MSO is a yardstick language for expressing properties about them. We refer to e.g. [Gur85, Tho96] for surveys and to [GTW02, PP04] for comprehensive treatments of the subject.

D. Siefkes has shown in [Sie70] that a fragment of second-order Peano's arithmetic containing the comprehension axiom scheme and the induction axiom is complete with respect to the standard model: every MSO-formula true on streams is provable. The approach taken there was to formalize the translation of MSO-formulas to Büchi automata. This requires to represent automata in the logic and to formalize the correctness proof of the translation in the corresponding deduction system.

In this paper, we give an alternative proof of Siefkes' completeness result by using model-theoretic tools. This leads to a more abstract proof which does not require explicit manipulation of automata in the logic. To our knowledge, such approaches to MSO have not been much explored compared to the great body of work on automata and corresponding algebraic structures [GTW02, PP04].

We follow the method of [GtC09], where complete axiomatizations of variants of MSO on finite trees are presented. Starting from Henkin completeness, we show that all models of our axiomatization are equivalent w.r.t. the validity of MSO-formulas. As in [GtC09], we use Feferman-Vaught Theorems obtained by the Ehrenfeucht-Fraïssé method [She75].

In contrast with [She75, Gur85], works like [GtC09] or the present one have to handle non-standards models of second-order arithmetic. As far as Henkin completeness is concerned, a model $\mathcal{M}$ of MSO can be seen as a structure with two domains: a domain $\mathcal{M}^\iota$ of individuals and a domain $\mathcal{M}^o \subseteq \mathcal{P}(\mathcal{M}^\iota)$ of sets of individuals (called *predicates* in this paper). Besides non-standards individuals (whose order type is very different from ω), the main difficulty is that $\mathcal{M}^o$ is in general strictly contained in $\mathcal{P}(\mathcal{M}^\iota)$: there might not be "enough" predicates. The interested reader can look at e.g. [BS73] for a discussion on this topic.

Even if we avoid the technicalities of explicitly manipulating Büchi automata in MSO, we think that much intuitions are gained by having them in mind. A crucial observation due to K. Doets [Doe89] makes apparent in (possibly non-standard) models a structure similar to standard infinite words. Our main technical contribution is a kind of Feferman-Vaught Infinitary Fusion for such models. Intuitively, it is a model-theoretic counterpart to a run of a Büchi automaton on a standard infinite word. The point is to ensure that such a "run" always exists as a predicate of a given model. For this, we use Ramseyan factorizations similar to those of infinite words (see e.g. [PP04]).

The paper is organized as follows. In Section 2, we describe our formal system for MSO, as well as the class of models we are interested in. These models are motivated by usual results on Henkin completeness for second-order logic that we briefly recall. We present in Section 3 the notions on the Ehrenfeucht-Fraïssé method that we will need.

We use it to prove a Feferman-Vaught Finite Sums Lemma for linearly ordered structures with parameters, which is discussed in Section 4. We then give the main argument for completeness in Section 5. It relies on an infinitary version of the Finite Sums Lemma, that we call "Infinite Fusion" and which is shown in Section 6. The Infinite Fusion Lemma uses a weak form of Ramsey's theorem (for additive MSO-definable colorings). We discuss it Section 7.

2. A Deduction System for Monadic Second-Order Logic on Streams

2.1. Language

We consider a formulation of *Monadic Second-Order Logic* (MSO) based on a two-sorted language: There is one sort ι intended to range over *individuals* and one sort o intended to range over *monadic* (or one-place) *predicates on individuals*. We assume given two countable sets $\mathcal{V}_\iota = \{x, y, z, \dots\}$ and $\mathcal{V}_o = \{X, Y, Z, \dots\}$ of respectively individual and predicate *variables*. The formulas of MSO are then defined by the following grammar:

$$\phi, \psi \in \Lambda \quad ::= \quad Xx \quad | \quad x < y \quad | \quad \neg\phi \quad | \quad \phi \vee \psi \quad | \quad \exists X\, \phi \quad | \quad \exists x\, \phi$$

The set $\mathrm{FV}(\phi)$ of free (individual and predicate) variables of a formula ϕ is defined as usual. A *sentence* (or *closed formula*) is a formula with no free variable, *i.e.* a formula ϕ such that $\mathrm{FV}(\phi) = \emptyset$. Formulas are identified modulo renaming of their bound variables. The capture-avoiding substitution of y for x in ϕ is written $\phi[y/x]$.

The other logical connectives are defined as usual:

$$\begin{array}{rcl} \phi \to \psi & := & \neg\phi \vee \psi \\ \phi \wedge \psi & := & \neg(\neg\phi \vee \neg\psi) \\ \phi \longleftrightarrow \psi & := & (\phi \to \psi) \wedge (\psi \to \phi) \end{array} \qquad \begin{array}{rcl} \forall X\, \phi & := & \neg\exists X\, \neg\phi \\ \forall x\, \phi & := & \neg\exists x\, \neg\phi \end{array}$$

2.2. Deduction for Second-Order Logic

We now discuss formal deduction for second-order logic. As usual, the rules for second-order logic are those of the (two-sorted) classical predicate calculus together with the comprehension axiom scheme (see e.g. [Sha91]). There are several different formulations equivalent w.r.t. provability. The following Natural Deduction system is a possible choice.

The *deduction relation* is writen $\Gamma \vdash \phi$, where Γ is a (possibly empty) finite unordered list of (possibly not closed) formula, and ϕ is a (possibly not closed) formula. It is inductively defined by the following rules.

- Rules for propositional logic:

$$\frac{}{\Gamma \vdash \phi \vee \neg\phi} \qquad \frac{}{\Gamma, \phi \vdash \phi} \qquad \frac{\Gamma \vdash \phi \qquad \Gamma \vdash \neg\phi}{\Gamma \vdash \psi}$$

$$\frac{\Gamma \vdash \phi}{\Gamma \vdash \phi \vee \psi} \qquad \frac{\Gamma \vdash \psi}{\Gamma \vdash \phi \vee \psi} \qquad \frac{\Gamma \vdash \phi \vee \psi \qquad \Gamma, \phi \vdash \varphi \qquad \Gamma, \psi \vdash \varphi}{\Gamma \vdash \varphi}$$

- Rules for predicate logic (where $\mathcal{X}, \mathcal{Y} \in \mathcal{V}_\iota$ or $\mathcal{X}, \mathcal{Y} \in \mathcal{V}_o$)

$$\frac{\Gamma \vdash \phi[\mathcal{Y}/\mathcal{X}]}{\Gamma \vdash \exists \mathcal{X}\, \phi} \qquad\qquad \frac{\Gamma \vdash \exists \mathcal{X}\, \phi \qquad \Gamma, \phi \vdash \psi}{\Gamma \vdash \psi}\ (\mathcal{X} \notin \mathrm{FV}(\Gamma, \psi))$$

- Comprehension scheme (for all formula ϕ):

$$\frac{}{\Gamma \vdash \exists X\, \forall x\, (Xx \longleftrightarrow \phi)} \qquad (X \notin \mathrm{FV}(\phi))$$

2.3. Models of Second-Order Logic

We discuss the class of structures (or models) we will use to interpret the language of MSO presented in Section 2.1. These structures are motivated by known results on Henkin completeness that we briefly recall.

Structures, Assignments and Satisfiability. We are interested in *(Henkin) structures* $\mathcal{M} = (\mathcal{M}^\iota, \mathcal{M}^o, <_\mathcal{M})$ where $\mathcal{M}^\iota$ is a *non-empty* set of *individuals*, $\mathcal{M}^o \subseteq \mathcal{P}(\mathcal{M}^\iota)$ is a *non-empty* set of *predicates* and $<_\mathcal{M}$ is a binary relation on $\mathcal{M}^\iota$. It is convenient to call $\mathcal{M}^\iota$ and $\mathcal{M}^o$ respectively the *individual* and *predicate domains* of $\mathcal{M}$.

An $\mathcal{M}$-*assignment* is a map $\rho : (\mathcal{V}_\iota \cup \mathcal{V}_o) \to (\mathcal{M}^\iota \cup \mathcal{M}^o)$ which respects the sorts, *i.e.* such that $\rho(x) \in \mathcal{M}^\iota$ and $\rho(X) \in \mathcal{M}^o$ if $x \in \mathcal{V}_\iota$ and $X \in \mathcal{V}_o$. Given $x \in \mathcal{V}_\iota$ and $a \in \mathcal{M}^\iota$, we write $\rho[a/x]$ for the assignment which maps x to a and is equal to ρ everywhere else. The assignment $\rho[A/X]$ (where $X \in \mathcal{V}_o$ and $A \in \mathcal{M}^o$) is defined similarly.

Given a structure $\mathcal{M}$, an $\mathcal{M}$-assignment ρ and a formula ϕ, we define the *satisfaction relation* $\mathcal{M}, \rho \models \phi$ by induction on ϕ as usual:

$$\begin{array}{lll}
\mathcal{M}, \rho \models Xx & \text{iff} & \rho(x) \in \rho(X) \\
\mathcal{M}, \rho \models x < y & \text{iff} & \rho(x) <_\mathcal{M} \rho(y) \\
\mathcal{M}, \rho \models \neg\phi & \text{iff} & \mathcal{M}, \rho \not\models \phi \\
\mathcal{M}, \rho \models \phi \vee \psi & \text{iff} & \mathcal{M}, \rho \models \phi \text{ or } \mathcal{M}, \rho \models \psi \\
\mathcal{M}, \rho \models \exists X\, \phi & \text{iff} & \text{there is some } A \in \mathcal{M}^o \text{ such that } \mathcal{M}, \rho[A/X] \models \phi \\
\mathcal{M}, \rho \models \exists x\, \phi & \text{iff} & \text{there is some } a \in \mathcal{M}^\iota \text{ such that } \mathcal{M}, \rho[a/x] \models \phi
\end{array}$$

We say that ϕ is *valid* in $\mathcal{M}$ (notation $\mathcal{M} \models \phi$) if $\mathcal{M}, \rho \models \phi$ for every ρ. A set of formulas Δ is valid in $\mathcal{M}$ (notation $\mathcal{M} \models \Delta$) if $\mathcal{M} \models \phi$ for every $\phi \in \Delta$.

It is sometimes convenient to consider formulas with a fixed assignment of their free variables to some structure $\mathcal{M}$. These formulas are called *formulas with parameters in* $\mathcal{M}$. We define them as pairs of a formula ϕ and a *finite partial* $\mathcal{M}$-assignment $\nu : (\mathcal{V}_\iota \cup \mathcal{V}_o) \rightharpoonup (\mathcal{M}^\iota \cup \mathcal{M}^o)$. The set of free variables of the formula with parameters (ϕ, ν) is $\mathrm{FV}(\phi, \nu) := \mathrm{FV}(\phi) \setminus \mathrm{dom}(\nu)$. We will often write $\phi[\nu(\mathcal{X})/\mathcal{X} \mid \mathcal{X} \in \mathrm{dom}(\nu)]$ for the formula with parameters (ϕ, ν).

The satisfaction of a formula with parameters (ϕ, ν) in a structure $\mathcal{M}$ and assignment ρ (notation $\mathcal{M}, \rho \models (\phi, \nu)$) is defined as the satisfaction of ϕ in $\mathcal{M}$ and assignment $\rho[\nu(\mathcal{X})/\mathcal{X} \mid \mathcal{X} \in \mathrm{dom}(\nu)]$. The corresponding validity relation $\mathcal{M} \models (\phi, \nu)$ holds if $\mathcal{M}, \rho \models (\phi, \nu)$ for every ρ.

Second-Order Henkin Structures. Deduction *without the comprehension scheme* is correct in any structure $\mathcal{M}$: if $\vdash \phi$ is derivable without using the comprehension then ϕ is valid in $\mathcal{M}$. The following notions are useful to handle the comprehension scheme. A set of individuals $A \in \mathcal{P}(\mathcal{M}^\iota)$ is *definable* if there is a formula ϕ and an $\mathcal{M}$-assignment ρ such that

$$A \quad = \quad \{a \in \mathcal{M}^\iota \mid \mathcal{M}, \rho[a/x] \models \phi\}$$

Of course, all $A \in \mathcal{M}^o$ are definable. The converse is more interesting, since $\mathcal{M}$ satisfies every instance of the comprehension scheme if and only if $\mathcal{M}^o$ is the set of all definable $A \in \mathcal{P}(\mathcal{M}^\iota)$. In this case, we call $\mathcal{M}$ a *second-order* (Henkin) structure.

Remark 2.1 *(i) We say that $\mathcal{M}$ is* full *if $\mathcal{M}^o = \mathcal{P}(\mathcal{M}^\iota)$. Full structures are second-order.*

(ii) Finite boolean combinations of definable predicates are definable. Hence, the predicate domain of a second-order structure is closed under finite boolean operations.

Henkin Completeness. Usual Henkin completeness holds for deduction w.r.t. validity in all second-order Henkin structures (see e.g. [Sha91]):

Theorem 2.2 (Henkin Completeness) *Let Δ be a set of sentences and ϕ be a sentence. Assume that for all second-order Henkin structure $\mathcal{M}$, if $\mathcal{M} \models \Delta$ then $\mathcal{M} \models \phi$. Then there is a finite set $\Gamma \subseteq \Delta$ such that $\Gamma \vdash \phi$.*

Remark 2.3 (On Henkin Completeness) *It may be worth recalling some points on Henkin completeness w.r.t. Henkin structures. The most general notion of models for the language of MSO is that of* general models *$\mathcal{M} = (\mathcal{M}^\iota, \mathcal{M}^o, \varepsilon_{\mathcal{M}}, <_{\mathcal{M}})$ where $\mathcal{M}^\iota$ and $\mathcal{M}^o$ are* two arbitrary non-empty sets *and $\varepsilon_{\mathcal{M}} \subseteq \mathcal{M}^\iota \times \mathcal{M}^o$. Satisfaction in general models is defined as for Henkin structures, but with the clause for Xx replaced by*

$$\mathcal{M}, \rho \models Xx \qquad \textit{iff} \qquad \rho(x)\, \varepsilon_{\mathcal{M}}\, \rho(X)$$

A general model is second-order *if it satisfies every instance of the comprehension scheme. Henkin completeness w.r.t. validity in general second-order models is just usual Henkin-completeness for two-sorted first-order logic:*

- *Let Δ be a set of sentences and ϕ be a sentence. Assume that for all second-order general model $\mathcal{M}$, if $\mathcal{M} \models \Delta$ then $\mathcal{M} \models \phi$. Then there is a finite set $\Gamma \subseteq \Delta$ such that $\Gamma \vdash \phi$.*

Theorem 2.2 is obtained from completeness w.r.t. general models thanks to the following well-known fact (see e.g. [Sha91]). Given a general model $\mathcal{M} = (\mathcal{M}^\iota, \mathcal{M}^o, \varepsilon_{\mathcal{M}}, <_{\mathcal{M}})$, define the Henkin structure $\mathcal{M}_{\mathcal{P}}$ as $(\mathcal{M}^\iota, \mathcal{M}^o_{\mathcal{P}}, <_{\mathcal{M}})$ where $\mathcal{M}^o_{\mathcal{P}}$ is the set of the extensions w.r.t. $\varepsilon_{\mathcal{M}}$ of the predicates of $\mathcal{M}$:

$$\mathcal{M}^o_{\mathcal{P}} \quad := \quad \{\{a \in \mathcal{M}^\iota \mid a\, \varepsilon_{\mathcal{M}}\, A\} \mid A \in \mathcal{M}^o\}$$

Then, $\mathcal{M}$ and $\mathcal{M}_{\mathcal{P}}$ are equivalent in the sense that for every formula ϕ, we have $\mathcal{M} \models \phi$ if and only if $\mathcal{M}_{\mathcal{P}} \models \phi$. In particular, $\mathcal{M}$ is second-order if and only if $\mathcal{M}_{\mathcal{P}}$ is second-order.

2.4. Equality

Monadic Second-Order Logic has a definable equality (see e.g. [Sha91]):

$$x \doteq y \;:=\; \forall X\,(Xx \to Xy)$$

Thanks to the comprehension scheme, $\doteq$ is an equivalence relation:

$$\vdash \forall x\,(x \doteq x) \qquad \vdash \forall xy\,(x \doteq y \to y \doteq x) \qquad \vdash \forall xyz\,(x \doteq y \to y \doteq z \to x \doteq z)$$

which moreover satisfies Leibniz's scheme: for all formula ϕ,

$$\vdash \forall xy\,(x \doteq y \to \phi[x/z] \to \phi[y/z])$$

Remark 2.4 *Given a second-order structure $\mathcal{M}$, we have $\mathcal{M}^\iota, \emptyset \in \mathcal{M}^o$ since $\mathcal{M}^\iota$ is definable by the formula $(x \doteq x)$.*

Second-Order Structures with Correct Equality. It is well-known (see e.g. [Sha91]) that the equality $\doteq$ may not be correct: Given a structure $\mathcal{M}$, it is possible that $\mathcal{M} \models (a \doteq b)$ but $a \neq b$, even if $\mathcal{M}$ is second-order. We say that a structure $\mathcal{M}$ has *correct equality* if $\mathcal{M} \models (a \doteq b)$ implies $a = b$ for all $a, b \in \mathcal{M}^\iota$.

Remark 2.5 *(i) Full structures have correct equality.*

(ii) Consider an arbitrary structure $\mathcal{M}$ with correct equality. Note that every singleton $\{a\}$ with $a \in \mathcal{M}^\iota$ is definable (by the formula with parameters $(x \doteq y, [a/x])$). According to Remark 2.1.(ii), it follows that if $\mathcal{M}$ is second-order, then $\mathcal{M}^o$ contains all the finite subsets of $\mathcal{M}^\iota$.

(iii) In particular, finite second-order structures with correct equality are full.

As far as Henkin completeness is concerned, it is always possible to assume that a second-order structure has correct equality. The ideas are similar to those of Remark 2.3.

Consider a second-order structure $\mathcal{M} = (\mathcal{M}^\iota, \mathcal{M}^o, <_\mathcal{M})$ and define $\doteq_\mathcal{M} \subseteq \mathcal{M}^\iota \times \mathcal{M}^\iota$ as $\{(a, b) \mid \mathcal{M} \models a \doteq b\}$. Then $\doteq_\mathcal{M}$ is an equivalence relation since $\mathcal{M}$ is second-order. By definition of $\doteq$, predicates $A \in \mathcal{M}^o$ are preserved by $\doteq_\mathcal{M}$: if $a \in A$ and $a \doteq_\mathcal{M} b$ then $b \in A$. Moreover, Leibniz's scheme implies that $<_\mathcal{M}$ is also preserved by $\doteq_\mathcal{M}$. It follows that we can define a structure $\mathcal{M}_{\doteq}$ by quotienting each component of $\mathcal{M}$ by $\doteq_\mathcal{M}$. It is well-known (see e.g. [Sha91]) that $\mathcal{M}$ and $\mathcal{M}_{\doteq}$ are equivalent in the sense of Remark 2.3. In particular, $\mathcal{M}_{\doteq}$ is second-order, and we have the following strengthening of Henkin completeness:

Corollary 2.6 *Let Δ be a set of sentences and ϕ be a sentence. Assume that for all second-order Henkin structure $\mathcal{M}$ with correct equality, if $\mathcal{M} \models \Delta$ then $\mathcal{M} \models \phi$. Then there is a finite set $\Gamma \subseteq \Delta$ such that $\Gamma \vdash \phi$.*

2.5. Axiomatization

The *standard model* is $\mathsf{N} := (\mathbb{N}, \mathcal{P}(\mathbb{N}), <_{\mathbb{N}})$, where $<_{\mathbb{N}}$ is the usual order on natural numbers. Recall that thanks to the celebrated result of Büchi [Büc62], the *monadic theory* of N (*i.e.* the set of MSO-sentences ϕ such that $\mathsf{N} \models \phi$) is decidable.

In this section, we describe a set MSO^ω of MSO-sentences which *completely axiomatizes* the monadic theory of N: for all MSO-sentence ϕ, if $\mathsf{N} \models \phi$ then $\mathsf{MSO}^\omega \vdash \phi$. The axiomatization we consider is an adaptation of that of [Sie70] to the language of MSO presented in Section 2.1. This is essentially a fragment of second-order Peano's arithmetic with full comprehension and induction.

For the completeness proof of MSO^ω, we shall also discuss variations on Ramsey's theorem and the axiom of choice in Sections 5, 6 and 7. We moreover work in the weaker axiomatization MSO (not every non-zero individual has a predecessor) when discussing Ramsey's Theorem in Section 7.

Definition 2.7 (MSO^ω and MSO) *MSO^ω is the set of the following sentences:*

- *Linear Order axioms:*

$$\forall x\, \neg(x < x) \qquad \forall xyz\, (x < y \to y < z \to x < z)$$

$$\forall xy\, (x < y \lor x \doteq y \lor y < x)$$

- *Unboundedness axiom:*

$$\forall x\, \exists y\, (x < y)$$

- *Induction axiom:*

$$\forall X\ [\forall x\, (\forall y\, (y < x \to Xy) \to Xx) \to \forall x\, Xx]$$

- *Predecessor axiom:*

$$\forall x\, (\exists y(y < x) \to \exists y[y < x \land \neg\exists z\, (y < z \land z < x)])$$

MSO is MSO^ω minus the predecessor axiom.

A formula ϕ is derivable in MSO^ω *(resp.* derivable in MSO*) if $\mathsf{MSO}^\omega \vdash \phi$ (resp. $\mathsf{MSO} \vdash \phi$) is derivable using the deduction system of Section 2.2.*

A second-order structure with correct equality $\mathcal{M}$ is a model of MSO^ω *(resp.* a model of MSO*) if $\mathcal{M} \models \mathsf{MSO}^\omega$ (resp. $\mathcal{M} \models \mathsf{MSO}$).*

In this paper, we give a model-theoretic proof of Siefkes' completeness result:

Theorem 2.8 (Completeness of MSO^ω [Sie70]) *For all sentence ϕ, if $\mathsf{N} \models \phi$ then $\mathsf{MSO}^\omega \vdash \phi$.*

Following the method of [GtC09], our route to Theorem 2.8 is to use usual Henkin completeness (as formulated in Corollary 2.6), and to show that all models of MSO^ω are equivalent w.r.t. the validity of MSO-formulas. This is the main result of the paper.

Theorem 2.9 (Main Theorem) *Let $\mathcal{M}$ be a model of MSO^ω and ϕ be a sentence. We have $\mathcal{M} \models \phi$ if and only if $\mathsf{N} \models \phi$.*

Theorem 2.9 is proved in Section 5. As [GtC09], we rely on Feferman-Vaught Theorems proved by the Ehrenfeucht-Fraïssé method.

We now discuss some aspects of the different axioms of MSO^ω. All structures considered here are second-order and have correct equality.

Orders. We use the following defined formula:

$$x \leq y \;:=\; x < y \vee x \doteq y$$

Hence, in a structure $\mathcal{M}$ with correct equality, given $a, b \in \mathcal{M}^\iota$ we have $\mathcal{M} \models a \leq b$ if and only if ($a = b$ or $a <_\mathcal{M} b$).

A structure $\mathcal{M}$ is *linearly ordered* if is satisfies the Linear Order axioms. The first two sentences say that $<_\mathcal{M}$ is strict and transitive. Note that $<_\mathcal{M}$ is thus antisymmetric: if $a <_\mathcal{M} b$ then $b \not<_\mathcal{M} a$. The third sentence says that $<_\mathcal{M}$ is total. Since $\mathcal{M}$ is assumed to have correct equality, it is equivalent to requiring that for all $a, b \in \mathcal{M}^\iota$ we have either $a <_\mathcal{M} b$ or $a = b$ or $b <_\mathcal{M} a$.

Induction. The induction axiom holds in the standard model N but is false for instance in the full structure of real numbers.[1]

Assume that $\mathcal{M}$ satisfies the induction axiom. The contrapositive of this axiom says that each non-empty predicate $A \in \mathcal{M}^o$ has minimal elements:

$$\mathcal{M} \models \exists x A x \to \exists x \left[A x \wedge \forall y (y < x \to \neg A y)\right]$$

If moreover $\mathcal{M}$ is linearly ordered, then A has a unique least element.

Recursion. Assume that $\mathcal{M}$ satisfies the induction axiom, and moreover that $<_\mathcal{M}$ is strict and transitive. Thanks to the comprehension scheme, $\mathcal{M}$ satisfies the following course-of-values Recursion Theorem.

Given a formula ϕ with parameters in $\mathcal{M}$ and variables $x \in \mathcal{V}_\iota$ and $X \in \mathcal{V}_o$, write $\phi[X < x]$ when ϕ is independent from X for the y such that $\neg(y < x)$:

$$\mathcal{M} \models \forall x \forall X Y \left[\forall y < x \,(X y \longleftrightarrow Y y) \to (\phi[X < x] \longleftrightarrow \phi[Y < x])\right]$$

Theorem 2.10 (Recursion Theorem) *Given $\phi[X < x]$ as above, using the comprehension scheme let $A \in \mathcal{M}^o$ be such that for all $a \in \mathcal{M}^\iota$,*

$$\mathcal{M} \models A a \longleftrightarrow \forall X \left[\forall x \leq a (X x \longleftrightarrow \phi[X < x]) \to X a\right]$$

Then A is correct for ϕ:

$$\mathcal{M} \models \forall x (A x \longleftrightarrow \phi[A < x])$$

and moreover, if $B \in \mathcal{M}^o$ is also correct for ϕ, then B and A are equal:

$$\mathcal{M} \models \left[\forall x (B x \longleftrightarrow \phi[B < x]) \to \forall x (A x \longleftrightarrow B x)\right]$$

[1] The monadic theory of $\mathbb{R}$ is undecidable (see [Gur85] for references).

Proof. In the case of MSO^ω, this is Theorem 1.b.1 of [Sie70]. The proof is easily adapted ordered structures satisfying the induction axiom (see Appendix A). $\square$

Successors and Predecessors. If $\mathcal{M}$ is linearly ordered and satisfies the induction axiom, then every $a \in \mathcal{M}^\iota$ which is not maximal has a *successor*, *i.e.* there is a unique least $b >_\mathcal{M} a$. However, a non minimal $a \in \mathcal{M}^\iota$ may not have a *predecessor*, *i.e.* a greatest $b <_\mathcal{M} a$.[2] The predecessor axiom ensures that every non-minimal individual has a predecessor.

Remark 2.11 (Backward Bounded Induction) *If $\mathcal{M}$ is linearly ordered and satisfies the axiom of induction, then the predecessor axiom is equivalent to the following principle of* Backward Bounded Induction*:*

Given a non-empty predicate $B \in \mathcal{M}$, if B is bounded*, i.e. if there is some $b \in \mathcal{M}^\iota$ such that $a \leq_\mathcal{M} b$ for all $a \in B$, then B has a* maximal element *i.e. there is an (unique) $a \in B$ such that $c \leq_\mathcal{M} a$ for all $c \in B$.*

Unboundedness. The axiom of Unboundedness is a kind of infinity axiom. Given a structure $\mathcal{M}$, we say that $U \in \mathcal{M}^o$ is *unbounded in* $\mathcal{M}$ if for all $a \in \mathcal{M}$ there is some $b \in U$ such that $a <_\mathcal{M} b$. If $<_\mathcal{M}$ is strict and transitive, then U must be infinite. Note however that the converse does not hold, even for models of MSO^ω.

Remark 2.12 (Non-Standard Models of MSO^ω) *A model $\mathcal{M}$ of MSO^ω can be* non-standard *(i.e. non-isomorphic to the standard model N) for two reasons: (i) because its predicate domain $\mathcal{M}^o$ is different from $\mathcal{P}(\mathcal{M}^\iota)$ or (ii) because its individual domain is not isomorphic to $\mathbb{N}$. Let us discuss these two points in view of Theorem 2.9.*

(i) It is well-known that if $\mathcal{M}$ is full (i.e. $\mathcal{M}^o = \mathcal{P}(\mathcal{M}^\iota)$), then $\mathcal{M}^\iota$ is isomorphic to $\mathbb{N}$ (see e.g. [Sha91]). Hence non-standard models $\mathcal{M}$ always have $\mathcal{M}^o \subsetneq \mathcal{P}(\mathcal{M}^\iota)$.

(ii) Thanks to the Löwenheim-Skolem Theorem (see e.g. [vD04]), we can always assume that an MSO^ω-model $\mathcal{M}$ has a countable individual domain $\mathcal{M}^\iota$. However, the order structure of $\mathcal{M}$ can be very different from that of $\mathbb{N}$. For instance, if $\mathcal{M}$ is a non-standard model of second-order Peano's arithmetic, then it is also a model of MSO^ω. But $\mathcal{M}$ is also a non-standard model of First-Order Peano's Arithmetic, and it is well-known (see e.g. [BBJ07]) that its order type is that of: $\mathbb{N}$ followed by $\mathbb{Q}$ copies of $\mathbb{Z}$. In particular, segments of the form $[a, b) = \{c \in \mathcal{M}^\iota \mid a \leq_\mathcal{M} c <_\mathcal{M} b\}$ may be infinite.

3. The Ehrenfeucht-Fraïssé Method

The proof of Theorem 2.9 is based on the Ehrenfeucht-Fraïssé method (see e.g. [EF99]). We present here the notions on this method that we will need.

[2] Besides completeness w.r.t. N, recall that the monadic theory of the ordinal ω_2 is independent from ZFC (see [She75] or [GS83] for related results).

For the remaining of the paper, we fix enumerations of the individual and predicate variables. Let $\mathcal{V}_\iota = \{x_1, \dots, x_p, \dots\}$ and $\mathcal{V}_o = \{X_1, \dots, X_q, \dots\}$. We say that ϕ is a *p-q-formula* if $\mathrm{FV}(\phi) \subseteq \{x_1, \dots, x_p, X_1, \dots, X_q\}$.

Unlike the rest of the paper, the results discussed in this section are insensitive on whether we are dealing with Henkin structures, general models, or second-order version thereof. For convenience, we will only consider Henkin structures which are *not necessarily second-order*. In this context, two formulas ϕ and ψ are *logically equivalent* if $(\phi \longleftrightarrow \psi)$ is valid in all such structures.

3.1. Logical Equivalence Up To Bounded Quantifier Depth

The first step is to classify formulas according to their quantifier-depth.

Definition 3.1 (Quantifier-Depth) *The* quantifier depth $\mathrm{qd}(\phi)$ *of a formula ϕ is defined by induction on ϕ as follows:*

$$\begin{array}{rcl@{\qquad}rcl} \mathrm{qd}(Xx) & := & 0 & \mathrm{qd}(\exists x\, \phi) & := & \mathrm{qd}(\phi) + 1 \\ \mathrm{qd}(x < y) & := & 0 & \mathrm{qd}(\exists X\, \phi) & := & \mathrm{qd}(\phi) + 1 \\ \mathrm{qd}(\neg\phi) & := & \mathrm{qd}(\phi) & \mathrm{qd}(\phi \vee \psi) & := & \max(\mathrm{qd}(\phi), \mathrm{qd}(\psi)) \end{array}$$

We let $\Lambda^{p,q}_n$ be the set of p-q-formulas of quantifier depth $\leq n$ and write Λ_n for $\Lambda^{0,0}_n$.

A remarkable property of languages without function symbols, such as the language of MSO, is the following standard observation (see e.g. [EF99]).

Lemma 3.2 (First Finiteness Lemma) *Up to logical equivalence, there are only finitely many p-q-formulas of quantifier depth $\leq n$.*

Proof. By induction on $n \in \mathbb{N}$, show that for all $p, q \in \mathbb{N}$ there are finitely many equivalence classes of $\Lambda^{p,q}_n$. Note that logical equivalence on $\Lambda^{p,q}_n$ can be characterized by maps from p-q-atoms and $(\Lambda^{p+1,q}_m \cup \Lambda^{p,q+1}_m)_{m<n}$-classes to booleans. Observe now that there are finitely many p-q-atoms and that $(\Lambda^{p+1,q}_m \cup \Lambda^{p,q+1}_m)_{m<n}$ has finitely many equivalence classes by induction hypothesis. □

Recall that logical equivalence is defined as validity of equivalence in all (possibly non second-order) structures. Requiring instead validity of equivalence in *all second-order structures* has no impact on finiteness: This amounts to add the comprehension axiom scheme, and adding axioms can only *reduce* the number of equivalence classes.

3.2. Structures with Parameters

A *structure with parameters* is a structure $\mathcal{M}$ together with $a_1, \dots, a_p \in \mathcal{M}^\iota$ and $A_1, \dots, A_q \in \mathcal{M}^o$. We write $\overline{a}$ for a finite sequence of individuals of length $|\overline{a}|$ and similarly for $\overline{A}$. If $|\overline{a}| = p$ and $|\overline{A}| = q$ then we say that $(\mathcal{M}, \overline{a}, \overline{A})$ is a *p-q-structure*.

If ϕ is a p-q-formula, we write $(\mathcal{M}, \overline{a}, \overline{A}) \models \phi$ for $\mathcal{M} \models \phi[\overline{a}/\overline{x}][\overline{A}/\overline{X}]$. Two p-q-structures $(\mathcal{M}, \overline{a}, \overline{A})$ and $(\mathcal{N}, \overline{b}, \overline{B})$ are *n-equivalent* (notation $\equiv^{p,q}_n$) if they satisfy the same p-q-formulas of q.d. $\leq n$. We write $\equiv_n$ instead of $\equiv^{p,q}_n$ when p, q are clear from the context.

Lemma 3.3 (Second Finiteness Lemma) *Given $n, p, q \in \mathbb{N}$, there are only finitely many $\equiv_n^{p,q}$ equivalences classes.*

Proof. Each n-equivalence class of a p-q-structure can be characterized by an union of equivalence classes $\Lambda_n^{p,q}$ modulo logical equivalence. But there are only finitely many equivalence classes of the latter thank to the First Finiteness Lemma 3.2. $\square$

The Finiteness Lemmas allow to characterize the n-equivalence class of a p-q-structure by a single p-q-formula.

Corollary 3.4 *For all $n \in \mathbb{N}$ and all p-q-structure $(\mathcal{M}, \overline{a}, \overline{A})$, there is a formula $\phi \in \Lambda_n^{p,q}$ such that for all p-q-structure $(\mathcal{N}, \overline{b}, \overline{B})$, we have $(\mathcal{N}, \overline{b}, \overline{B}) \models \phi$ if and only if $(\mathcal{M}, \overline{a}, \overline{B}) \equiv_n (\mathcal{N}, \overline{b}, \overline{B})$. Such a ϕ is an n-*characteristic *of $(\mathcal{M}, \overline{a}, \overline{B})$.*

Moreover, there is a finite set *$\Phi_n^{p,q} \subseteq \Lambda_n^{p,q}$ of n-characteristics which contains an n-characteristic of each p-q-structure.*

3.3. Ehrenfeucht-Fraïssé Games

Ehrenfeucht-Fraïssé games are a convenient characterization of $\equiv_n$-equivalence for languages satisfying the First Finiteness Lemma. There are different possible formulations for second-order logic, see e.g. [Ros82, EF99, GtC09]. Our presentation is inspired from [GtC09], which is itself that of [EF99] adapted to non-full models.

Definition 3.5 (Ehrenfeucht-Fraïssé Games) *Given two structures $(\mathcal{M}, \overline{a}, \overline{A})$ and $(\mathcal{N}, \overline{b}, \overline{B})$ and $n \in \mathbb{N}$, the* Ehrenfeucht-Fraïssé Game $\mathrm{EF}_n((\mathcal{M}, \overline{a}, \overline{A}), (\mathcal{N}, \overline{b}, \overline{B}))$ *is an n-round game played between two players called "Spoiler" and "Duplicator".*

At each round, Spoiler plays first and chooses either an individual or a predicate in one of the two structures. Duplicator then responds in the other structure by choosing an individual if Spoiler chose an individual or a predicate if Spoiler goosed a predicate. After n rounds, Spoiler and Duplicator have build a finite relation

$$\{(a'_1, b'_1), \dots, (a'_p, b'_p), (A'_1, B'_1), \dots, (A'_q, B'_q)\}$$

with $n = p + q$, $\overline{a}' \in \mathcal{M}^\iota$, $\overline{b}' \in \mathcal{N}^\iota$, $\overline{A}' \in \mathcal{M}^o$ and $\overline{B}' \in \mathcal{N}^o$. Then Duplicator wins if and only if $(\mathcal{M}, \overline{aa}', \overline{AA}') \equiv_0 (\mathcal{N}, \overline{bb}', \overline{BB}')$.

Our presentation differs from [GtC09, EF99] on the following point. In these works, Duplicator wins if the finishing tuple is a finite partial isomorphism between the two structures. In our case, equality is not a quantifier-free formula, and it seems simpler to have instead a coarser wining condition based on $\equiv_0$-equivalence.

Ehrenfeucht-Fraïssé games characterize $\equiv_n$-equivalence:

Theorem 3.6 *Given two structures $(\mathcal{M}, \overline{a}, \overline{A})$ and $(\mathcal{N}, \overline{b}, \overline{B})$ and $n \in \mathbb{N}$, Duplicator has a wining strategy in $\mathrm{EF}_n((\mathcal{M}, \overline{a}, \overline{A}), (\mathcal{N}, \overline{b}, \overline{B}))$ if and only if $(\mathcal{M}, \overline{a}, \overline{A}) \equiv_n (\mathcal{N}, \overline{b}, \overline{B})$.*

Theorem 3.6 is proved by induction on $n \in \mathbb{N}$. The base case is trivial, and the induction step is provided by the following Lemma 3.7.

We will actually only use Ehrenfeucht-Fraïssé games in the form of Lemma 3.7. It is convenient to use the Spoiler/Duplicator terminology in that Lemma: we call "Spoiler moves" the object universally quantified in the items (i)–(iv) below, and call "Duplicator moves" the corresponding existentially quantified objects.

Lemma 3.7 *Consider two structures $(\mathcal{M}, \overline{a}, \overline{A})$ and $(\mathcal{N}, \overline{b}, \overline{B})$ and $n \in \mathbb{N}$. Then we have $(\mathcal{M}, \overline{a}, \overline{A}) \equiv_{n+1} (\mathcal{N}, \overline{b}, \overline{B})$ if and only if*

(i) for all $a \in \mathcal{M}^\iota$ there is $b \in \mathcal{N}^\iota$ such that $(\mathcal{M}, \overline{a}a, \overline{A}) \equiv_n (\mathcal{N}, \overline{b}b, \overline{B})$, and

(ii) for all $b \in \mathcal{N}^\iota$ there is $a \in \mathcal{M}^\iota$ such that $(\mathcal{M}, \overline{a}a, \overline{A}) \equiv_n (\mathcal{N}, \overline{b}b, \overline{B})$, and

(iii) for all $A \in \mathcal{M}^o$ there is $B \in \mathcal{N}^o$ such that $(\mathcal{M}, \overline{a}, \overline{A}A) \equiv_n (\mathcal{N}, \overline{b}, \overline{B}B)$, and

(iv) for all $B \in \mathcal{N}^o$ there is $A \in \mathcal{M}^o$ such that $(\mathcal{M}, \overline{a}, \overline{A}A) \equiv_n (\mathcal{N}, \overline{b}, \overline{B}B)$.

Proof. ($\Longrightarrow$) Assume that $(\mathcal{M}, \overline{a}, \overline{A}) \equiv_{n+1} (\mathcal{N}, \overline{b}, \overline{B})$. We only detail the case of Spoiler choosing an individual $a \in \mathcal{M}^\iota$, the others being similar. By Corollary 3.4 let ϕ be the n-characteristic of $(\mathcal{M}, \overline{a}a, \overline{A})$. Since $(\mathcal{M}, \overline{a}, \overline{A})$ satisfies the formula $\exists x\, \phi$ of quantifier depth $n+1$, by assumption $(\mathcal{N}, \overline{b}, \overline{B})$ satisfies also $\exists x\, \phi$. Hence there is a $b \in \mathcal{N}^\iota$ such that $(\mathcal{N}, \overline{b}b, \overline{B}) \models \phi$. But this implies $(\mathcal{N}, \overline{b}b, \overline{B}) \equiv_n (\mathcal{M}, \overline{a}a, \overline{A})$ by definition of ϕ.

($\Longleftarrow$) We have to show that $(\mathcal{M}, \overline{a}, \overline{A}) \equiv_{n+1} (\mathcal{N}, \overline{b}, \overline{B})$. Since formulas of quantifier $n+1$ are boolean combination of formulas of the form $\exists x\, \phi$ or $\exists X\, \phi$ with $\mathrm{qd}(\phi) = n$, it is sufficient to show that the two structures agree on formulas of the form $\exists x\, \phi$ or $\exists X\, \phi$ with $\mathrm{qd}(\phi) = n$. We only consider the case of $\exists X\, \phi$, that of $\exists x\, \phi$ being similar.

Assume that $(\mathcal{M}, \overline{a}, \overline{A}A) \models \phi$ for some $A \in \mathcal{M}^o$. By assumption there is some $B \in \mathcal{N}^o$ such that $(\mathcal{N}, \overline{b}, \overline{B}B) \equiv_n (\mathcal{M}, \overline{a}, \overline{A}A)$, hence $(\mathcal{N}, \overline{b}, \overline{B}B) \models \phi$. □

4. Finite Sums of Segments

We now discuss how to restrict structures into *segments* that can be concatenated. This will be done for second-order linearly ordered structures with correct equality. The Ehrenfeucht-Fraïssé method allows to give simple proofs that concatenation of segments preserves $\equiv_n$-equivalence. This leads to a partial sum operation on $\equiv_n$-classes. We follow well-known patterns of Feferman-Vaught Theorems [She75, Gur85, GtC09].

4.1. Restrictions and Relativizations

Segments will be obtained from structures by restrictions and relativizations. The material presented in this section is independent from the Ehrenfeucht-Fraïssé techniques, and in particular from the Finiteness Lemmas of Section 3.

The *restriction* of a structure $\mathcal{M}$ to some *non-empty* predicate $A \in \mathcal{M}^o$ is the structure $\mathcal{M}{\restriction}A$ defined as expected: its individual domain is $\mathcal{M}^\iota \cap A$, its predicate domain is $\{B \cap A \mid B \in \mathcal{M}^o\}$ and its relation $<_{\mathcal{M}{\restriction}A}$ is the restriction of $<_{\mathcal{M}}$ to A: $<_{\mathcal{M}{\restriction}A} := <_{\mathcal{M}} \cap (A \times A)$. It is convenient to write the individual and predicate domains of $\mathcal{M}{\restriction}A$ respectively as $\mathcal{M}^\iota{\restriction}A$ and $\mathcal{M}^o{\restriction}A$.

Restrictions of Structures with Parameters. We shall also need the less usual restriction of *structures with parameters*. Let $p, q \in \mathbb{N}$. Consider a structure $\mathcal{M}$ with individual parameters $\overline{a} = a_1 \dots a_p$ and predicate parameters $\overline{A} = A_1 \dots A_q$. Let $A \in \mathcal{M}^o$ be non-empty and such that $a_1, \dots, a_p \in A$. We define the *restriction of* $(\mathcal{M}, \overline{a}, \overline{A})$ *to* A to be the structure:

$$(\mathcal{M}, \overline{a}, \overline{A}){\restriction}A \quad := \quad (\mathcal{M}{\restriction}A, a_1 \dots a_p, (A_1 \cap A) \dots (A_q \cap A))$$

Remark *It seems is unclear how to deal in general with individual parameters. Our assumption that $\overline{a} \in A$ may not be the weakest possible, but it is sufficient for the results of this paper.*

We now discuss a simple technical fact which will be useful in the following Transfer Property 4.2.

Let us look at the satisfiability of a formula of the form $\exists X\, \psi$ in $(\mathcal{M}, \overline{a}, \overline{A}){\restriction}A$. By definition, we have $(\mathcal{M}, \overline{a}, \overline{A}){\restriction}A \models \exists X\, \psi$ if and only if there is some $C \in \mathcal{M}^o{\restriction}A$ such that $(\mathcal{M}, \overline{a}, \overline{A}){\restriction}A, [C/X] \models \psi$. Now, C is of the form $B \cap A$ for some $B \in \mathcal{M}^o$. For the forthcoming Transfer Property 4.2, we would like to deduce $(\mathcal{M}, \overline{a}, \overline{A}B){\restriction}A \models \psi$. This requires an induction on formulas that we perform in the following Lemma.

Lemma 4.1 *Let $p, q, k \in \mathbb{N}$. Consider a structure $\mathcal{M}$ with individual parameters $\overline{a} = a_1 \dots a_p$ and predicate parameters $\overline{A} = A_1 \dots A_q$. Let $A \in \mathcal{M}^o$ be non-empty and such that $a_1, \dots, a_p \in A$.*

Furthermore, let $B_1, \dots, B_k \in \mathcal{M}^o$, and for each $1 \leq i \leq k$, let $C_i := B_i \cap A \in \mathcal{M}^o{\restriction}A$. Then, for all p-$(q+k)$-formula ϕ we have

$$(\mathcal{M}, \overline{a}, \overline{A}){\restriction}A, [\overline{C}/\overline{X}] \models \phi \quad \textit{if and only if} \quad (\mathcal{M}, \overline{a}, \overline{AB}){\restriction}A \models \phi$$

Proof. By induction on ϕ. If ϕ is a negation or a disjunction then the result follows by induction hypothesis. We consider the other cases for ϕ.

- **Cases of $x_i < x_j$ and of $X_i x_j$ with $i \leq q$.** Both cases are trivial since the truth value of ϕ is independent from $\overline{B}$ and $\overline{C}$.
- **Case of $X_i x_j$ with $i \geq q+1$.** Then ϕ holds in $(\mathcal{M}, \overline{a}, \overline{AB}){\restriction}A$ if and only if $a_j \in B_{i-q}$. Since $a_j \in A$, this is equivalent to $a_j \in B_{i-q} \cap A = C_{i-q}$, hence to $(\mathcal{M}, \overline{a}, \overline{A}){\restriction}A, [\overline{C}/\overline{X}] \models \phi$.
- **Case of $\exists X\, \psi$.** Then ϕ holds in $(\mathcal{M}, \overline{a}, \overline{A}){\restriction}A, [\overline{C}/\overline{X}]$ if and only if there is $C \in \mathcal{M}^o{\restriction}A$ such that $(\mathcal{M}, \overline{a}, \overline{A}){\restriction}A, [\overline{C}/\overline{X}, C/X] \models \psi$. But $C \in \mathcal{M}^o{\restriction}A$ if and

only if $C = B \cap A$ for some $B \in \mathcal{M}^o$, and by induction hypothesis, we get that $(\mathcal{M}, \overline{a}, \overline{A})\restriction A, [\overline{C/X}, C/X] \models \psi$ if and only if $(\mathcal{M}, \overline{a}, \overline{ABB})\restriction A \models \psi$. By a second application of the induction hypothesis, this is equivalent to $(\mathcal{M}, \overline{a}, \overline{AB})\restriction A, [C/X] \models \psi$, hence to $(\mathcal{M}, \overline{a}, \overline{AB})\restriction A \models \exists X\, \psi$,

- **Case of $\exists x\, \psi$.** By direct application of the induction hypothesis since both structures have the same individual domain $\mathcal{M}^\iota \cap A$. □

Relativization of Formulas. An analogous operation can be defined on formulas. Let ϕ and φ be two formulas with no free variables in common, and let y be a variable not appearing free in ϕ. The *relativization* of ϕ to $\varphi[y]$, notation $\phi\restriction\varphi[y]$, is defined by induction on ϕ as follows:

$$\begin{array}{rcll}
\phi\restriction\varphi[y] & := & \phi & \text{if } \phi \text{ is atomic} \\
(\phi \vee \psi)\restriction\varphi[y] & := & (\phi\restriction\varphi[y]) \vee (\psi\restriction\varphi[y]) & \\
(\neg\phi)\restriction\varphi[y] & := & \neg(\phi\restriction\varphi[y]) & \\
(\exists X\, \phi)\restriction\varphi[y] & := & \exists X\, (\phi\restriction\varphi[y]) & \text{if } X \notin \mathrm{FV}(\varphi) \\
(\exists x\, \phi)\restriction\varphi[y] & := & \exists x\, (\varphi[x/y] \wedge \phi\restriction\varphi[y]) & \text{if } x \notin \mathrm{FV}(\varphi) \cup \{y\}
\end{array}$$

If (ϕ, ν) is a formula with parameters in a structure $\mathcal{M}$, and if $A \in \mathcal{M}^o$ contains all individual parameters of ϕ, then $(\phi, \nu)\restriction A$ is defined as $((\phi\restriction(Xx)[x]), \nu[A/X])$ where $X, x \notin \mathrm{FV}(\phi, \nu)$.

The Transfer Property. We now check that restriction and relativization are equivalent w.r.t. satisfaction. This in particular implies that restriction preserves the comprehension scheme.

Proposition 4.2 (Transfer) *Let $p, q \in \mathbb{N}$ and $(\mathcal{M}, a_1, \dots, a_p, A_1, \dots, A_q)$ be a structure with parameters. Let φ be a formula with parameters in $\mathcal{M}$ and whose free variable are disjoint from $\{x_1, \dots, x_p, X_1, \dots, X_q\}$. Given $x_0 \notin \{x_1, \dots, x_p\}$, let $A \in \mathcal{M}^o$ be non-empty and such that*

$$(\mathcal{M}, \overline{a}, \overline{A}) \models \forall x\, (Ax \longleftrightarrow \varphi[x/x_0])$$

Assume that $a_1, \dots, a_p \in A$. Let ϕ be a formula with $\mathrm{FV}(\phi) \subseteq \{x_1, \dots, x_p, X_1, \dots, X_q\}$. Then we have

$$(\mathcal{M}, \overline{a}, \overline{A})\restriction A \models \phi \quad \textit{if and only if} \quad (\mathcal{M}, \overline{a}, \overline{A}) \models \phi\restriction\varphi[x_0]$$

Proof. By induction on ϕ. If ϕ is a negation or a disjunction then the result follows by induction hypothesis.

- If ϕ is atomic then $\phi\restriction\varphi[x_0]$ is ϕ and there are two cases. If $\phi = (x_i < x_j)$, then $(\mathcal{M}, \overline{a}, \overline{A}) \models \phi$ if and only if $a_i <_{\mathcal{M}} a_j$. Since $a_i, a_j \in A$, this is equivalent to $a_i <_{\mathcal{M}\restriction A} a_j$ hence to $(\mathcal{M}, \overline{a}, \overline{A})\restriction A \models x_i < x_j$.

  Otherwise $\phi = (X_i x_j)$. Then $(\mathcal{M}, \overline{a}, \overline{A}) \models \phi$ if and only if $a_j \in A_i$. Since $a_j \in A$ this is equivalent to $a_j \in (A_i \cap A)$, hence to $(\mathcal{M}, \overline{a}, \overline{A})\restriction A \models X_i x_j$.

- If $\phi = (\exists X\, \psi)$, then $\phi\restriction\varphi[x_0]$ is $\exists X\,(\psi\restriction\varphi[x_0])$. Then $(\mathcal{M}, \overline{a}, \overline{A}) \models \phi\restriction\varphi[x_0]$ if and only if there is $B \in \mathcal{M}^o$ such that $(\mathcal{M}, \overline{a}, \overline{A}B) \models \psi\restriction\varphi[x_0]$. Since we can assume $X \notin \mathrm{FV}(\varphi)$, by induction hypothesis this is equivalent to $(\mathcal{M}, \overline{a}, \overline{A}B)\restriction A \models \psi$. By our technical Lemma 4.1, this is equivalent to $(\mathcal{M}, \overline{a}, \overline{A})\restriction A, [C/X] \models \psi$ where $C := B \cap A \in \mathcal{M}^o\restriction A$. Then we are done since the latter is equivalent to $(\mathcal{M}, \overline{a}, \overline{A})\restriction A \models \exists X\, \psi$.

- If $\phi = (\exists x\, \psi)$, then $\phi\restriction\varphi[x_0]$ is $\exists x\,(\varphi[x/x_0] \wedge \psi\restriction\varphi[x_0])$. We can assume $x \notin \mathrm{FV}(\varphi)$. Since the free variables of φ are disjoint from $\overline{x}, \overline{X}$, we have $(\mathcal{M}, \overline{a}, \overline{A}) \models \phi\restriction\varphi[x_0]$ if and only if there is some $a \in A$ such that $(\mathcal{M}, \overline{a}a, \overline{A}) \models \psi\restriction\varphi[x_0]$, which is equivalent to $(\mathcal{M}, \overline{a}a, \overline{A})\restriction A \models \psi$ since by induction hypothesis $(\mathcal{M}, \overline{a}b, \overline{A})\restriction A \models \psi$ is equivalent to $(\mathcal{M}, \overline{a}b, \overline{A}) \models \psi\restriction\varphi[x_0]$ for all $b \in A$. □

An important consequence of Proposition 4.2 is that $\mathcal{M}\restriction A$ is second-order if $\mathcal{M}$ is second-order. Indeed, consider a definable $B \in \mathcal{P}(\mathcal{M}^\iota\restriction A)$. Hence, there is a formula ϕ and an $(\mathcal{M}\restriction A)$-assignment ρ such that

$$B \quad = \quad \{a \in \mathcal{M}^\iota\restriction A \mid \mathcal{M}\restriction A, \rho \models \phi\}$$

Assume that $\mathrm{FV}(\phi) = \{x_1, \dots, x_p, X_1, \dots, X_q\}$ and let $a_i := \rho(x_i)$ $(1 \le i \le p)$ and $A_j := \rho(X_j)$ $(1 \le j \le q)$. Since $\overline{a} \in A$ and $A_j = A_j \cap A$ $(1 \le j \le q)$, let

$$B' \quad := \quad \{a \in \mathcal{M}^\iota \mid (\mathcal{M}, \overline{a}, \overline{A})\restriction A \models \phi\}$$

so that $B = B' \cap A$. By the Transfer Property 4.2, we obtain that

$$B' \quad = \quad \{a \in \mathcal{M}^\iota \mid (\mathcal{M}, \overline{a}, \overline{A}) \models \phi\restriction A\}$$

Since $\mathcal{M}$ is second-order, we have $B' \in \mathcal{M}^o$, hence $B \in \mathcal{M}^o\restriction A$.

4.2. Finite Sums of Segments

All structures considered in this section are second-order, linearly ordered and with correct equality.

A *segment* of such a structure $\mathcal{M}$ is a predicate of one of the following form:

$$\begin{array}{lcll} [-,b) & := & \{c \in \mathcal{M}^\iota \mid c <_\mathcal{M} b\} & \text{where } b \text{ is not minimal} \\ [a,b) & := & \{c \in \mathcal{M}^\iota \mid a \le_\mathcal{M} c <_\mathcal{M} b\} & \text{where } a <_\mathcal{M} b \\ [a,-) & := & \{c \in \mathcal{M}^\iota \mid a \le_\mathcal{M} c\} & \end{array}$$

These predicates are defined by the formula with parameters:

$$(y < z, [b/z]) \qquad \text{and} \qquad ((x \le y \wedge y < z), [a/x, b/z]) \qquad \text{and} \qquad (x \le y, [a/x])$$

Recall that $\mathcal{M}\restriction[a,b)$ is second-order since $\mathcal{M}$ is second-order. Note that it is also a linear order with correct equality since $\mathcal{M}$ is a linear order with correct equality.

Two consecutive segments, say $(\mathcal{M}, \overline{a}, \overline{A})\restriction[a,b)$ and $(\mathcal{M}, \overline{b}, \overline{A})\restriction[b,c)$ can be concatenated to $(\mathcal{M}, \overline{a}\overline{b}, \overline{A})\restriction[a,c)$. Using the Ehrenfeucht-Fraïssé method, we now show that concatenation of segments preserves $\equiv_n$-equivalence.

Similar operations have already been defined for full models (see e.g. [She75]) as well as for Henkin models [GtC09]. Our operation differs from [GtC09] in the treatment of predicate parameters: since we only need the concatenation of consecutive segments which are restrictions of the same structure $\mathcal{M}$, we can share the predicate parameters in the two components. This simplifies both the statement and the proof of the Lemma.

In order to smoothly handle segments with different kinds of end-points, it is convenient to use the following notation in the Finite Fusion Lemma. Given a second-order structure $\mathcal{M}$ with correct equality, let $\mathcal{M}^\infty := \mathcal{M}^\iota \cup \{-\infty, +\infty\}$, where $-\infty, +\infty \notin \mathcal{M}^\iota$. Then, let $a <_{\mathcal{M}^\infty} b$ iff either $a, b \in \mathcal{M}^\iota$ and $a <_{\mathcal{M}} b$, or $a = -\infty$ and $b \in \mathcal{M}^\iota$, or $a \in \mathcal{M}^\iota$ and $b = +\infty$.

Lemma 4.3 (Finite Sums of Segments) *Consider two second-order linearly ordered structures $\mathcal{M}$ and $\mathcal{N}$, both with correct equality. Assume given $t_0, t_1, t_2 \in \mathcal{M}^\infty$ and $u_0, u_1, u_2 \in \mathcal{N}^\infty$ such that*

$$t_0 <_{\mathcal{M}^\infty} t_1 <_{\mathcal{M}^\infty} t_2 \qquad \textit{and} \qquad u_0 <_{\mathcal{N}^\infty} u_1 <_{\mathcal{N}^\infty} u_2$$

Let $n \in \mathbb{N}$.

Let $p, q, p' \in \mathbb{N}$ and assume given the following parameters:

- *$a_1, \ldots, a_p \in \mathcal{M}^\iota\restriction[t_0, t_1)$, and $a'_1, \ldots, a'_{p'} \in \mathcal{M}^\iota\restriction[t_1, t_2)$, and*
- *$b_1, \ldots, b_p \in \mathcal{N}^\iota\restriction[u_0, u_1)$, and $b'_1, \ldots, b'_{p'} \in \mathcal{N}^\iota\restriction[u_1, u_2)$, and*
- *$A_1, \ldots, A_q \in \mathcal{M}^o\restriction[t_0, t_2)$ and $B_1, \ldots, B_q \in \mathcal{N}^o\restriction[u_0, u_2)$.*

If

$$(\mathcal{M}, \overline{a}, \overline{A})\restriction[t_0, t_1) \equiv_n (\mathcal{N}, \overline{b}, \overline{B})\restriction[u_0, u_1) \quad \textit{and} \quad (\mathcal{M}, \overline{a}', \overline{A})\restriction[t_1, t_2) \equiv_n (\mathcal{N}, \overline{b}', \overline{B})\restriction[u_1, u_2)$$

then

$$(\mathcal{M}, \overline{a}\overline{a}', \overline{A})\restriction[t_0, t_2) \equiv_n (\mathcal{N}, \overline{b}\overline{b}', \overline{B})\restriction[u_0, u_2)$$

Proof. By induction on $n \in \mathbb{N}$.

Consider first the base case $n = 0$. We just have to show that $(\mathcal{M}, \overline{a}\overline{a}', \overline{A})\restriction[t_1, t_2)$ and $(\mathcal{N}, \overline{b}\overline{b}', \overline{B})\restriction[u_0, u_2)$ agree on atomic formulas ϕ with individual variables in $\{x_1, \ldots, x_p, x_{p+1}, \ldots, x_{p+p'}\}$ and predicate variables in $\{X_1, \ldots, X_q\}$. We consider the different cases for ϕ:

- **Case of $x_i < x_j$ with $i, j \leq p$ or $i, j \geq p + 1$.**
  If $i, j \leq p$, then $(\mathcal{M}, \overline{a}\overline{a}', \overline{A})\restriction[t_0, t_2) \models \phi$ if and only if $(\mathcal{M}, \overline{a}, \overline{A})\restriction[t_0, t_1) \models \phi$. The same holds for $\mathcal{N}$, and we are done since by assumption $(\mathcal{M}, \overline{a}, \overline{A})\restriction[t_0, t_1)$ and $(\mathcal{N}, \overline{b}, \overline{B})\restriction[u_0, u_1)$ are $\equiv_0$-equivalent. The case of $i, j \geq p + 1$ is similar.

- **Case of $x_i < x_j$ with $i \leq p$ and $j \geq p+1$.**
  In this case ϕ holds in both $(\mathcal{M}, \overline{aa'}, \overline{A})\restriction[t_0, t_2)$ and $(\mathcal{N}, \overline{bb'}, \overline{B})\restriction[u_0, u_2)$

- **Case of $x_i < x_j$ with $j \leq p$ and $i \geq p+1$.**
  In this case ϕ is false in both $(\mathcal{M}, \overline{aa'}, \overline{A})\restriction[t_0, t_2)$ and $(\mathcal{N}, \overline{bb'}, \overline{B})\restriction[u_0, u_2)$

- **Case of $X_i x_j$.**
  If $j \leq p$, then x_j is instantiated with a parameter in $\overline{a}$. It follows that ϕ is true in $(\mathcal{M}, \overline{aa'}, \overline{A})\restriction[t_0, t_2)$ if and only if it is true in $(\mathcal{M}, \overline{a}, \overline{A})\restriction[t_0, t_1)$. The same holds for $\mathcal{N}$, and we are done since $(\mathcal{M}, \overline{a}, \overline{A})\restriction[t_0, t_1) \equiv_0 (\mathcal{N}, \overline{b}, \overline{B})\restriction[u_0, u_1)$. The case of $j \geq p+1$ is similar.

We now consider the inductive step: we show the property for $n+1$ assuming it for n. We use the ($\Longleftarrow$) implication of Lemma 3.7, and consider the different possible moves of Spoiler. We then let Duplicator answer in the corresponding components according to the ($\Longrightarrow$) implication of Lemma 3.7 in these components.

- **Spoiler plays an individual, say** $a \in \mathcal{M}^\iota\restriction[t_0, t_1)$ (the other cases for individuals are similar).
  By Lemma 3.7 ($\Longrightarrow$), there is $b \in \mathcal{N}^\iota\restriction[u_0, u_1)$ such that $(\mathcal{M}, \overline{a}a, \overline{A})\restriction[t_0, t_1) \equiv_n (\mathcal{N}, \overline{b}b, \overline{B})\restriction[u_0, u_1)$. We can now conclude by induction hypothesis, since in the other component we have $(\mathcal{M}, \overline{a}', \overline{A})\restriction[t_1, t_2) \equiv_n (\mathcal{N}, \overline{b}', \overline{B})\restriction[u_1, u_2)$.

- **Spoiler plays a predicate, say** $A \in \mathcal{M}^\iota\restriction[t_0, t_2)$ (the other case for predicates is similar).
  By Lemma 3.7 ($\Longrightarrow$), there are $B \in \mathcal{N}^o\restriction[u_0, u_1)$ and $B' \in \mathcal{N}^o\restriction[u_1, u_2)$ such that $(\mathcal{M}, \overline{a}, \overline{A}A)\restriction[t_0, t_1)$ is $\equiv_n$-equivalent to $(\mathcal{N}, \overline{b}, \overline{B}B)\restriction[u_0, t_2)$ and $(\mathcal{M}, \overline{a}', \overline{A}A)\restriction[t_1, t_2)$ is $\equiv_n$-equivalent to $(\mathcal{N}, \overline{b}', \overline{B}B')\restriction[u_1, u_2)$. Take $B'' := B \cup B'$. Note that $B'' \in \mathcal{N}^o$ by Remark 2.1.(ii).

  Since $B''\restriction[u_0, u_1) = B$ and $B''\restriction[u_1, u_2) = B'$, we have $(\mathcal{M}, \overline{a}, \overline{A}A)\restriction[t_0, t_1) \equiv_n (\mathcal{N}, \overline{b}, \overline{B}B'')\restriction[u_0, u_1)$ and $(\mathcal{M}, \overline{a}', \overline{A}A)\restriction[t_1, t_2) \equiv_n (\mathcal{N}, \overline{b}', \overline{B}B'')\restriction[u_1, u_2)$ and we conclude by induction hypothesis. □

5. Completeness of MSO$^\omega$ w.r.t. the Standard Model

Thanks to the the Ehrenfeucht-Fraïssé method, we actually prove the following formulation of Theorem 2.9:

Theorem 5.1 *Let $\mathcal{M}$ be a model of MSO$^\omega$ and $n \in \mathbb{N}$. For all sentence $\phi \in \Lambda_n$, we have $\mathcal{M} \models \phi$ if and only if $\mathsf{N} \models \phi$.*

In this section, we present a proof of this result. It relies on some infinite combinatorics, namely a weak form of Ramsey's theorem (for additive colorings), a weak form of the axiom of choice similar to the "Splicing" axiom scheme of [Sie70, BS73], and an infinite extension of the Finite Sums Lemma, that we call "Infinite Fusion".

Ramsey's Theorem and the Splicing axiom scheme are shown to hold for MSO^ω in [Sie70]. We prove the Infinite Fusion Lemma in Section 6, where we also discuss the Splicing axiom scheme.

We discuss Ramsey's Theorem in more details in Section 7, where we show that any model of MSO satisfies Ramsey's theorem for MSO-definable colorings, and moreover that in second-order linearly ordered structures with correct equality, Ramsey's theorem for additive MSO-definable colorings implies Ramsey's theorem for all MSO-definable colorings.

5.1. Doets' Lemma

Our way to Theorem 5.1 starts from the crucial observation that bounded segments of models of MSO^ω are $\equiv_n$-equivalent to finite linear orders. To our knowledge, this is due to [Doe89] for the Π^1_1-case (first-order logic with universal prenex quantification on predicates). It may be worth recalling here that according to Remark 2.12, a bounded segment of an arbitrary model of MSO^ω may not be finite.

In our context, a *finite linear order* is a structure of the form $\mathsf{N}{\restriction}[m_0, m_1)$ with $m_0 < m_1 \in \mathbb{N}$. Note that if $m_1 - m_0 = k_1 - k_0$ (where $m_0 < m_1$ and $k_0 < k_1$), then $\mathsf{N}{\restriction}[m_0, m_1) \equiv_n \mathsf{N}{\restriction}[k_0, k_1)$ for all $n \in \mathbb{N}$.

Lemma 5.2 (Doets' Lemma) *Let $\mathcal{M}$ be a model of MSO^ω and $n \in \mathbb{N}$. For all $a <_{\mathcal{M}} b$, there is a finite linear order $\mathcal{L}$ such that $\mathcal{M}{\restriction}[a, b) \equiv_n \mathcal{L}$.*

Proof. Fix $a \in \mathcal{M}^\iota$. By Corollary 3.4, there is a finite set $\Phi \subseteq \Lambda_n$ which contains an n-characteristic of each finite linear order and moreover such that each $\phi \in \Phi$ is an n-characteristic of some finite linear order.

Let now $b >_{\mathcal{M}} a$. Then $\mathcal{M}{\restriction}[a, b)$ is $\equiv_n$-equivalent to a finite linear order if and only if $\mathcal{M}{\restriction}[a, b) \models \bigvee_{\phi\in\Phi} \phi$. By the Transfer Property 4.2, this is equivalent to $\mathcal{M} \models (\bigvee_{\phi\in\Phi} \phi){\restriction}[a, b)$. It follows that "$\mathcal{M}{\restriction}[a, b)$ is $\equiv_n$-equivalent to a finite linear order" is expressible by a formula of the form $\psi{\restriction}[a, b)$.

Hence we are done if we show that $\mathcal{M} \models \forall y\, (a < y \to \psi{\restriction}[a, y))$. We use the axiom of induction. We will leave implicit the further applications of Transfer (Proposition 4.2).

Let $b >_{\mathcal{M}} a$ such that $\mathcal{M} \models \psi{\restriction}[a, c)$ for all $a <_{\mathcal{M}} c <_{\mathcal{M}} b$. If there is no such c, then we are done since $\mathcal{M}{\restriction}[a, b)$ is the singleton $\{a\}$, hence $\equiv_n$-equivalent to $\mathsf{N}{\restriction}[0, 1)$. Otherwise, there is a greatest $a <_{\mathcal{M}} c <_{\mathcal{M}} b$, so that $\mathcal{M}{\restriction}[c, b)$ is the singleton $\{c\}$. Since $\mathcal{M}{\restriction}[a, c)$ is $\equiv_n$-equivalent to a finite linear order, say $\mathsf{N}{\restriction}[0, n)$, we conclude by the Finite Sums Lemma 4.3 that $\mathcal{M}{\restriction}[a, b) \equiv_n \mathsf{N}{\restriction}[0, n + 1)$. $\square$

5.2. Ramseyan Factorizations

Let $\mathcal{M}$ be a model of MSO^ω. In order to obtain $\mathcal{M} \equiv_n \mathsf{N}$ from Doets' Lemma 5.2, we would like to perform a kind of infinite sum of the $(\mathcal{M}{\restriction}[a, b))_{a<_{\mathcal{M}}b}$. For this we rely on Ramsey's Theorem, which is provable in MSO^ω [Sie70]. As usual with MSO on streams (see e.g. [She75, PP04]), we only need a weak form of Ramsey's Theorem, that we call "Ramseyan $\equiv_n$-Factorizations".

The terminology "Ramseyan $\equiv_n$-Factorization" is inspired from the "Ramseyan Factorizations" of ω-words discussed e.g. in [PP04]. Recall from Corollary 3.4 that if $\mathcal{M}$ is a linearly ordered second-order structure with correct equality, then for all $n \in \mathbb{N}$ and all $a <_\mathcal{M} b$, there is a $\phi \in \Phi^{0,0}_n$ such that $\mathcal{M} \models \phi{\restriction}[a,b)$. Then we say that $\mathcal{M}$ has Ramseyan $\equiv_n$-Factorizations if there is $\phi \in \Phi^{0,0}_n$ and an unbounded predicate $U \in \mathcal{M}^o$ which homogeneous for ϕ. We actually need a slightly stronger statement involving formulas with predicate parameters.

It is convenient to use the following notation. Given a structure $\mathcal{M}$ and a predicate $U \in \mathcal{M}^o$, we let $[U]^2 \subseteq \mathcal{M}^\iota \times \mathcal{M}^\iota$ be the set of pairs $(a,b) \in U \times U$ such that $a <_\mathcal{M} b$.

Definition 5.3 (Ramseyan Factorizations) *Let $\mathcal{M}$ be a second-order linearly ordered structure with correct equality and let $n,q \in \mathbb{N}$. We say that $\mathcal{M}$* has Ramseyan $\equiv^q_n$-factorizations *when the following holds.*

For all $A_1,\dots,A_q \in \mathcal{M}^o$ and all unbounded $U \in \mathcal{M}^o$, there is an unbounded predicate $V \subseteq U$ and a $\phi \in \Phi^{0,q}_n$ such that for all $(a,b) \in [V]^2$ we have $(\mathcal{M},,\overline{A}) \models \phi{\restriction}[a,b)$.

That models of MSO$^\omega$ have Ramseyan $\equiv^q_n$-factorizations for every $n,q \in \mathbb{N}$ follows from the fact that Ramsey's theorem is derivable in MSO$^\omega$. We come back to this point in Section 7.

Theorem 5.4 *Models of MSO$^\omega$ have Ramseyan $\equiv^q_n$-factorizations for every $n,q \in \mathbb{N}$.*

Proof. By Theorem I.1.c.3 of [Sie70]. See also Section 7.1.2. □

Ramseyan factorizations give the following consequence of Doets' Lemma.

Corollary 5.5 *Let $\mathcal{M}$ be a model of MSO$^\omega$ and $n \in \mathbb{N}$. There is an unbounded $U \in \mathcal{M}^o$ and a finite linear order $\mathcal{L}$ such that for all $(a,b) \in [U]^2$ we have $\mathcal{M}{\restriction}[a,b) \equiv_n \mathcal{L}$.*

5.3. Infinite Fusion

Let $\mathcal{M}$ be a model of MSO$^\omega$ and $n \in \mathbb{N}$. Consider the unbounded predicate $U \subseteq \mathcal{M}^\iota$ and the finite linear order $\mathcal{L}$ given by Corollary 5.5. We can assume that the least element u of U is not the least element of $\mathcal{M}^\iota$. Then by Doets' Lemma 5.2 there is $m \in \mathbb{N}$ such that $\mathcal{M}{\restriction}[-,u) \equiv_n \mathsf{N}[0,m)$. Moreover, there is some $l \in \mathbb{N}$ such that $\mathcal{L} \equiv_n \mathsf{N}[m+kl, m+(k+1)l)$ for all $k \in \mathbb{N}$.

Hence, the discussion up to now has lead us to the following point: There are unbounded $U \in \mathcal{M}^o$ and $V \in \mathcal{P}(\mathbb{N})$ together with $u \in U$ and $v \in V$ such that

$$\mathcal{M}{\restriction}[-,u) \equiv_n \mathsf{N}{\restriction}[-,v)$$

and for all $(u_0,u_1) \in [U]^2$ and all $(v_0,v_1) \in [V]^2$

$$\mathcal{M}{\restriction}[u_0,u_1) \equiv_n \mathsf{N}{\restriction}[v_0,v_1)$$

We can conclude that $\mathcal{M} \equiv_n \mathsf{N}$ from these assumptions thanks to the following Infinite Fusion Lemma. We prove it in Section 6, and this will achieve the proof of Theorem 2.9.

Lemma 5.6 (Infinite Fusion) *Let $\mathcal{M}$, $\mathcal{N}$ be models of MSO^ω and let $n \in \mathbb{N}$. Let $U \in \mathcal{M}^o$ and $V \in \mathcal{N}^o$ be unbounded, and assume that their respective least elements u and v are not the least elements of respectively $\mathcal{M}^\iota$ and $\mathcal{N}^\iota$. Assume that*

$$\mathcal{M}{\restriction}[-,u) \equiv_n \mathcal{N}{\restriction}[-,v)$$

and that for all $(u_0,u_1) \in [U]^2$ and all $(v_0,v_1) \in [V]^2$ we have

$$\mathcal{M}{\restriction}[u_0,u_1) \equiv_n \mathcal{N}{\restriction}[v_0,v_1)$$

Then

$$\mathcal{M} \equiv_n \mathcal{N}$$

6. The Infinite Fusion Lemma

In this section, we prove the Infinite Fusion Lemma 5.6. We shall actually prove it for linearly-ordered second-order structures with correct equality, which have Ramseyan $\equiv_n$-factorizations and satisfy an additional axiom scheme that we call "Idempotent $\equiv^q_n$-Splicing". The Idempotent $\equiv^q_n$-Splicing axiom is a variation of the Splicing axiom scheme of [Sie70, BS73]. We already pointed in Section 5.2 that models of MSO^ω have Ramseyan factorizations (see also Section 7).

We discuss the Idempotent Splicing axiom scheme in Section 6.1, where we show, using results of [Sie70] that it holds for models of MSO^ω. We then discuss the Infinite Fusion Lemma in Section 6.2

6.1. Splicing

The Idempotent $\equiv^q_n$-Splicing axiom is a weak form of the axiom of choice similar to the "Splicing" axiom scheme of [Sie70, BS73].

We way that $a, b \in \mathcal{M}^\iota$ are *consecutive in U* if $a, b \in U$, $a <_\mathcal{M} b$ and there is no $c \in U$ such that $a <_\mathcal{M} c$ and $c <_\mathcal{M} b$.

Definition 6.1 (Splicing) *Let $\mathcal{M}$ be a second-order linearly ordered structure with correct equality. We say that $\mathcal{M}$ satisfies the* Splicing axiom scheme *when the following holds.*

For all $q \in \mathbb{N}$, all $A_1, \dots, A_q \in \mathcal{M}^o$, all $U \in \mathcal{M}^o$, and all 0-$(q+1)$-formula ϕ, if for all $(a,b) \in [U]^2$ we have $(\mathcal{M},,\overline{A}A^{a,b}) \models \phi{\restriction}[a,b)$ for some $A^{a,b} \in \mathcal{M}^o$, then there is a predicate $A \in \mathcal{M}^o$ such that for all a, b consecutive *in U we have $(\mathcal{M},,\overline{A}A) \models \phi{\restriction}[a,b)$.*

The Splicing axiom scheme was considered in [Sie70] where it is shown to hold in MSO^ω. In [BS73] it is moreover shown that Splicing may fail in MSO-models which are not models of MSO^ω (*i.e.* in MSO-models which does not satisfy the predecessor axiom).

Theorem 6.2 *Let $\mathcal{M}$ be a model of MSO^ω with correct equality. Then $\mathcal{M}$ satisfies the Splicing axiom scheme.*

Proof. By Theorem I.5.b.1 of [Sie70]. See also Theorem 4.1 of [BS73] (without predicate parameters) or Appendix B. □

For the Infinite Fusion Lemma, we will rather consider the following consequence of Splicing for idempotent colorings.

Definition 6.3 (Indempotent Splicing) *Let $\mathcal{M}$ be a second-order linearly ordered structure with correct equality and let $n, q \in \mathbb{N}$. We say that $\mathcal{M}$ satisfies the* Idempotent $\equiv^q_n$-Splicing axiom *when the following holds.*

Given and $A_1, \dots, A_q \in \mathcal{M}^o$, let $\phi \in \Phi^{0,(q+1)}_n$ and $U \in \mathcal{M}^o$ such that

- *(i) $(\mathcal{M}A, , \overline{A}) \models \exists X\, \phi[X/X_{q+1}]{\restriction}[a, b)$ for all $(a, b) \in [U]^2$, and*
- *(ii) there is a 0-$(q + 1)$ second-order linearly ordered structure $(\mathcal{N}, , , \overline{B}B)$ with correct equality and $b_0, b_1, b_3 \in \mathcal{N}^\iota$ with $b_0 <_\mathcal{N} b_1 <_\mathcal{N} b_3$ such that $(\mathcal{N}, , \overline{B}B) \models \phi{\restriction}[b_i, b_j)$ for all $(i, j) \in [\{0, 3\}]^2$.*

Then there is a predicate $A \in \mathcal{M}^o$ such that for all $(a, b) \in [U]^2$ we have $(\mathcal{M}, , \overline{A}A) \models \phi{\restriction}[a, b)$.

In Definition 6.3 above, condition (i) is actually the premise of Definition 6.1. Condition (ii) intuitively says that ϕ defines an idempotent coloring. Note that given $n, q \in \mathbb{N}$, Idempotent $\equiv^q_n$-Splicing is really an axiom and not an axiom scheme since $\Phi^{0,(q+1}$ is finite by Corollary 3.4. The main difference between the two forms of Splicing is that the predicate U obtained in the idempotent version is correct for *all* segments $[a, b)$ with $(a, b) \in [U]^2$ rather than just for the *consecutive* $a, b \in U$ in the other version.

Thanks to the Finite Sums Lemma 4.3, Idempotent Splicing is easily derived from Splicing in models of MSO^ω.

Proposition 6.4 *Let $\mathcal{M}$ be a model of MSO^ω with correct equality. Then $\mathcal{M}$ satisfies the Idempotent $\equiv^q_n$-Splicing axiom for all $n, q \in \mathbb{N}$.*

Proof. Let $\overline{A} = A_1 \dots A_q \in \mathcal{M}^o$. Assume given $U \in \mathcal{M}^o$ and $\phi \in \Phi^{0,q+1}_n$ such that for all $(a, b) \in [U]^2$ we have $(\mathcal{M}, , \overline{A}) \models \exists X\, \phi[X/X_{q+1}]{\restriction}[a, b)$.

Now, by Splicing (Theorem 6.2), there is a predicate $A \in \mathcal{M}^o$ such that for all consecutive $a, b \in U$ we have $(\mathcal{M}, , \overline{A}A) \models \phi{\restriction}[a, b)$.

Given $u_0 \in U$, we now show that for all $u_1 \in U$ with $u_1 >_\mathcal{M} u_0$ we have $(\mathcal{M}, , \overline{A}A) \models \phi{\restriction}[a, b)$. We apply the axiom of induction (note that the property "$(\mathcal{M}, , \overline{A}A) \models \phi{\restriction}[u_0, u_1)$" is definable by a formula with parameters in $\mathcal{M}$).

So, let $u_1 \in U$ with $u_1 >_\mathcal{M} u_0$ and assume that the property holds for all $w <_\mathcal{M} u_1$. Since $\mathcal{M}$ is a model of MSO^ω, by backward bounded induction (Remark 2.11) let w be the greatest element of U which is $<_\mathcal{M} u_1$. If $w = u_0$, then u_0 and u_1 are consecutive in U and we are done. Otherwise, $w_1 >_\mathcal{M} u_0$ and by induction hypothesis we have $(\mathcal{M}, , \overline{A}A) \models \phi{\restriction}[u_0, w)$. Since w, v_1 are consecutive in U, we have $(\mathcal{M}, , \overline{A}A) \models \phi{\restriction}[w, u_1)$. In order to conclude that $(\mathcal{M}, , \overline{A}A) \models \phi{\restriction}[u_0, u_1)$, we apply the Finite Sums Lemma 4.3, the other structure being provided by condition (ii) of Definition 6.3. □

6.2. Infinite Fusion

As usual with the Ehrenfeucht-Fraïssé method, we will perform an induction on the quantifier depth of formulas. This leads us to consider a version of Lemma 5.6 for structures with parameters.

Lemma 6.5 (Infinite Fusion) *Let $\mathcal{M}$ and $\mathcal{N}$ be linearly-ordered second-order structures with correct equality. Assume that $\mathcal{M}$ and $\mathcal{N}$ satisfy the Idempotent $\equiv^q_n$-Splicing axiom and have Ramseyan $\equiv^q_n$-factorizations for all $n, q \in \mathbb{N}$.*

Let $n \in \mathbb{N}$.

Let $U \in \mathcal{M}^o$ and $V \in \mathcal{N}^o$ be unbounded, and assume that their respective least elements u and v are not the least elements of respectively $\mathcal{M}^\iota$ and $\mathcal{N}^\iota$.

Let $p, q \in \mathbb{N}$, and furthermore
$a_1, \ldots, a_p \in \mathcal{M}^\iota{\restriction}[-, u)$, $A_1, \ldots, A_q \in \mathcal{M}^o$ *and*
$b_1, \ldots, b_p \in \mathcal{N}^\iota{\restriction}[-, v)$, $B_1, \ldots, B_q \in \mathcal{N}^o$.
Assume that

$$(\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-, u) \equiv_n (\mathcal{N}, \overline{b}, \overline{B}){\restriction}[-, v)$$

and that for all $(u_0, u_1) \in [U]^2$ and all $(v_0, v_1) \in [V]^2$ we have

$$(\mathcal{M}, , \overline{A}){\restriction}[u_0, u_1) \equiv_n (\mathcal{N}, , \overline{B}){\restriction}[v_0, v_1)$$

Then

$$(\mathcal{M}, \overline{a}, \overline{A}) \equiv_n (\mathcal{N}, \overline{b}, \overline{B})$$

Note that in the statement of Lemma 6.5, since $\mathcal{M}$ and $\mathcal{N}$ are second-order linearly ordered structures with correct equality, we can always assume that the respective least elements of the unbounded predicates U and V are not the least elements of $\mathcal{M}$ and $\mathcal{N}$.

The rest of this section is devoted to the proof of Lemma 6.5. We reason by (external) induction on $n \in \mathbb{N}$.

Base Case

As for the finite case (Lemma 4.3), we just have to show that $(\mathcal{M}, \overline{a}, \overline{A})$ and $(\mathcal{N}, \overline{b}, \overline{B})$ agree on atomic formulas ϕ with individual variables in $\{x_1, \ldots, x_p\}$ and predicate variables in $\{X_1, \ldots, X_q\}$. We only detail the case of $x_i < x_j$, that of $X_i x_j$ being similar.

Since $a_i, a_j <_\mathcal{M} u$, the formula $(x_i < x_j)$ holds in $(\mathcal{M}, \overline{a}, \overline{A})$ if and only if it holds in $(\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-, u)$. The same holds for $\mathcal{N}$, and we are done since $(\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-, u) \equiv_0 (\mathcal{N}, \overline{b}, \overline{B}){\restriction}[-, v)$.

Inductive Step

We now consider the inductive step: we show the property for $n + 1$ assuming it for n. We use the ($\Longleftarrow$) implication of Lemma 3.7, and consider the different possible moves of Spoiler. We then let Duplicator answer in the corresponding segments according to the ($\Longrightarrow$) implication of Lemma 3.7 in these segments.

Spoiler plays an individual, say $a \in \mathcal{M}^\iota$. Since U is unbounded, there is $u' \in U$ strictly greater than a. Also using the unboundedness of V, Let $v' \in V$ be strictly greater than v. We have

$$(\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-, u) \equiv_{n+1} (\mathcal{N}, \overline{b}, \overline{B}){\restriction}[-, v)$$

Since

$$(\mathcal{M}, , \overline{A}){\restriction}[u, u') \equiv_{n+1} (\mathcal{N}, , \overline{B}){\restriction}[v, v')$$

by the Finite Sums Lemma 4.3, we have

$$(\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-, u') \equiv_{n+1} (\mathcal{N}, \overline{b}, \overline{B}){\restriction}[-, v')$$

Now, by Lemma 3.7 ($\Longrightarrow$) there is some $b \in \mathcal{N}^\iota{\restriction}[-, v')$ such that

$$(\mathcal{M}, \overline{a}a, \overline{A}){\restriction}[-, u') \equiv_{n} (\mathcal{N}, \overline{b}b, \overline{B}){\restriction}[-, v')$$

The predicates $U' := \{s \in U \mid s \geq_\mathcal{M} u'\}$ and $V' := \{t \in V \mid t \geq_\mathcal{N} v'\}$ are both unbounded. Hence, for all $(u_0, u_1) \in [U']^2$, $(v_0, v_1) \in [V']^2$, we have

$$(\mathcal{M}, , \overline{A}){\restriction}[u_0, u_1) \equiv_{n} (\mathcal{N}, , \overline{B}){\restriction}[v_0, v_1)$$

Moreover, since $\mathcal{M}$ and $\mathcal{N}$ are both linearly ordered and with correct equality, u' and v' are the least elements of respectively U' and V'. We can thus conclude by induction hypothesis.

Spoiler plays a predicate, say $A \in \mathcal{M}^o$. Using Corollary 3.4, let $\Phi_n := \Phi_n^{0,q+1}$ be a finite set of n-characteristics which contains an n-characteristic for each 0-$(q+1)$-structure. It follows that for each $(u_0, u_1) \in [U]^2$, there is some $\phi \in \Phi_n$ which is an n-characteristic for $(\mathcal{M}, , \overline{A}A){\restriction}[u_0, u_1)$. Recall that thanks to the Transfer Property (Proposition 4.2), $(\mathcal{M}, , \overline{A}A){\restriction}[u_0, u_1) \models \phi$ implies $(\mathcal{M}, , \overline{A}A) \models \phi{\restriction}[u_0, u_1)$.

Now, since $\mathcal{M}$ has Ramseyan $\equiv_n^{q+1}$-factorizations (Definition 5.3) we get an unbounded predicate $U' \subseteq U$ and an n-characteristic $\phi \in \Phi_n$ such that for all $(u_0, u_1) \in [U']^2$ we have $(\mathcal{M}, , \overline{A}A) \models \phi{\restriction}[u_0, u_1)$, *i.e.* $(\mathcal{M}, , \overline{A}A){\restriction}[u_0, u_1) \models \phi$. Since U' is unbounded and since on the other hand $\mathcal{M}$ is a second-order linearly ordered structure with correct equality, we can assume that U' has a least element u'.

We now claim that for all $(v_0, v_1) \in [V]^2$ we have $(\mathcal{N}, , \overline{B}) \models (\exists X\, \phi){\restriction}[v_0, v_1)$.

- *Proof of the claim.* Fix $(u_0, u_1) \in [U']^2 \subseteq [U]^2$. For all $(v_0, v_1) \in [V]^2$, since by assumption $(\mathcal{N}, , \overline{B}){\restriction}[v_0, v_1) \equiv_{n+1} (\mathcal{M}, , \overline{A}){\restriction}[u_0, u_1)$, by Lemma 3.7 ($\Longrightarrow$) there is some $B^{v_0,v_1} \in \mathcal{N}^o$ such that $(\mathcal{N}, , \overline{B}B^{v_0,v_1}){\restriction}[v_0, v_1) \equiv_n (\mathcal{M}, , \overline{A}A){\restriction}[u_0, u_1)$, hence $(\mathcal{N}, \overline{B}B^{v_0,v_1}){\restriction}[v_0, v_1) \models \phi$. □

Since $\mathcal{N}$ satisfies the Idempotent $\equiv_n^q$-Splicing axiom (Definition 6.3), we obtain a predicate $B \in \mathcal{N}^o$ such that $(\mathcal{N}, , \overline{B}B) \models \phi{\restriction}[v_0, v_1)$ for all $(v_0, v_1) \in [V]^2$. Note that condition (ii) of Definition 6.3 is satisfied with $(\mathcal{M}, , \overline{A}A)$ and any $u'_0 <_\mathcal{M} u'_1 <_\mathcal{M} u'_2$ in the unbounded predicate U'. Moreover, since V is unbounded and since on the other hand $\mathcal{N}$ is a second-order linearly ordered structure with correct equality, we can assume that V has a least element v.

We now build Duplicator's response to $A \in \mathcal{M}^o$, and then conclude by induction hypothesis. Recall that in order to meet the premises of the induction hypothesis, we have to take care of the initial segment $(\mathcal{M},,\overline{A}A)\restriction[-,u')$. By Lemma 3.7 ($\Longrightarrow$), there is some $B' \in \mathcal{N}^o$ such that $(\mathcal{N},,\overline{B}B')\restriction[-,v) \equiv_n (\mathcal{M},,\overline{A}A)\restriction[-,u')$. Since $\mathcal{N}$ is second-order, let $B'' := B'\restriction[-,v) \cup B\restriction[v,-)$. Now, $(\mathcal{N},\overline{b},\overline{B}B'')$ (together with V) satisfies the premise of the induction hypothesis and we are done.

This concludes the proof of Lemma 6.5.

7. Ramsey's Theorem in MSO

In this section, we show that Ramsey's theorem for MSO-definable colorings is provable in MSO.

Theorem 7.1 *Let Φ be a finite set of MSO-formulas. The following is derivable in MSO:*

$$\forall x, y\,(x < y \to \bigvee_{\phi \in \Phi} \phi) \to$$
$$\exists X\,[\forall x \exists y\,(x < y \wedge Xy) \wedge \bigvee_{\phi\in\Phi} \forall x, y\,(Xx \to Xy \to x < y \to \phi)]$$

In general, Ramsey's theorem does not hold for models of MSO. Indeed, if $\mathcal{M}$ is model of MSO, then by Löwenheim-Skolem Theorem (see e.g. [vD04]), we can assume that $\mathcal{M}^\iota$ is countable. It is well-known, then, that there is a partition of $[\mathcal{M}^\iota]^2$ whose infinite homogeneous sets $A \subseteq \mathcal{M}^\iota$ can only have order type $\omega := (\mathbb{N}, <)$ or $(\mathbb{N}, >)$ (see e.g. [Ros82], Proposition 11.3). If $\mathcal{M}$ is model of MSO but not of MSO$^\omega$, then the order type $(\mathcal{M}^\iota, <_\mathcal{M})$ may be some ordinal $> \omega$, so that A can not be unbounded in $\mathcal{M}$.

The partition of $[\mathcal{M}^\iota]^2$ mentioned above involves an ω-indexed cofinal sequence in $\mathcal{M}^\iota$ which is not available as such in the second-order model $\mathcal{M}$. Hence, it seems that Theorem 7.1 is really about structural properties of *MSO-definable colorings*, typically witnessed by the Ehrenfeucht-Fraïssé method and the properties depicted in Sections 4 and 6.

Using Henkin completeness (as formulated in Corollary 2.6), Theorem 7.1 will be obtained from Ramseyan Factorizations in second-order linearly ordered structures with correct equality. This will be done in Section 7.2. Before that, we show in Section 7.1 that models of MSO have Ramseyan factorizations. This will also gives us the occasion to discuss Ramseyan factorizations in MSO$^\omega$.

Before dealing with Ramsey's theorem, let us show the following useful property.

Proposition 7.2 (Infinite Pigeonhole Principle) *Let $\mathcal{M}$ be a linearly ordered second-order structure with correct equality. Assume given predicates $A_1, \ldots, A_k \in \mathcal{M}^o$ and an* *unbounded $U \in \mathcal{M}^o$. If for all $a \in U$ there is $1 \leq i \leq k$ such that $a \in A_i$, then there is an $1 \leq i \leq k$ such that $U \cap A_i$ is unbounded.*

Proof. Assume toward a contradiction that all $U \cap A_i$ are bounded. Hence there are $(a_i \in U \cap A_i)_{1 \leq i \leq k}$ such that $a_i \geq_{\mathcal{M}} b$ for all $b \in U \cap A_i$. Since U is unbounded, there is $u \in U$ such that $u >_{\mathcal{M}} a_1, \ldots, a_k$. Hence we must have $u \in A_i$ for some i, which implies $u \leq_{\mathcal{M}} a_i$, contradicting the antisymmetry of $<_{\mathcal{M}}$. $\square$

7.1. Ramseyan Factorizations in MSO

In this section, we prove the following:

Theorem 7.3 *Let $\mathcal{M}$ be a model of MSO with correct equality. Then $\mathcal{M}$ has Ramseyan $\equiv^q_n$-factorizations for every $n, q \in \mathbb{N}$.*

We adapt to MSO-models the proof of Ramsey's theorem for additive colorings given in [She75]. The main difference is that we apparently can not define the desired homogeneous predicate by induction on an externally given cofinal sequence.

The proof of Theorem 7.3 occupy the next three sections. We first define a candidate unbounded homogeneous predicate H^q_n in Section 7.1.1. This predicate will be homogeneous by construction. That H^q_n is unbounded will be shown in Sections 7.1.2 and 7.1.3.

We make heavy use of the Recursion Theorem 2.10.

7.1.1. The Candidate Unbounded Homogeneous Predicate

Let $\mathcal{M}$ be a model of MSO with correct equality and let $n \in \mathbb{N}$. Furthermore, as in the premises of Definition 5.3, let $q \in \mathbb{N}$, $A_1, \ldots, A_q \in \mathcal{M}^o$ and U be an unbounded predicate. Recall from Corollary 3.4 that for all $(a, b) \in [U]^2$ there is some $\phi \in \Phi^{0,q}_n$ such that $(\mathcal{M}, , \overline{A}) \models \phi{\restriction}[a, b)$. We see $\Phi_n := \Phi^{0,q}_n$ as a coloring of the segments $(\mathcal{M}, , \overline{A}){\restriction}[a, b)$ with $(a, b) \in [U]^2$. The Finite Sums Lemma 4.3 intuitively says that this coloring is additive in the sense of [She75].

We consider the usual "merging" relation. For ω-words, it is an equivalence relation on natural numbers (see e.g. [Tho88, PP04]). It will be here an equivalence relation $\sim_n \subseteq U \times U$. We define it as follows. Let θ_n be the formula:

$$\theta_n[X, x, y] \quad := \quad Xx \wedge Xy \wedge \exists z \left[Xz \wedge (x < z) \wedge (y < z) \wedge \bigvee_{\phi \in \Phi_n} (\phi{\restriction}[x, z) \wedge \phi{\restriction}[y, z)) \right]$$

Let now $a \sim_n b$ if and only if $(\mathcal{M}, , \overline{A}) \models \theta_n[U/X, a/x, b/y]$. The relation $\sim_n$ is clearly reflexive and symmetric. Transitivity follows from the unboundedness of U and the following direct consequence of the Finite Sums Lemma 4.3:

- If $(\mathcal{M}, , \overline{A}){\restriction}[a, c) \equiv_n (\mathcal{M}, , \overline{A}){\restriction}[b, c)$ for some $c >_{\mathcal{M}} a, b$, then $(\mathcal{M}, , \overline{A}){\restriction}[a, d) \equiv_n (\mathcal{M}, , \overline{A}){\restriction}[b, d)$ for all $d >_{\mathcal{M}} c$.

It is well-known that $\sim_n$ has at most $\#\Phi_n$ equivalence classes:

- *Proof.* Given a finite sequence $a_1, \dots, a_k \in \mathcal{M}^\iota$ and $b >_{\mathcal{M}} a_1, \dots, a_k$, if $k > \#\Phi_n$ then we have $(\mathcal{M}, , \overline{A}) \restriction [a_i, b) \equiv_n (\mathcal{M}, , \overline{A}) \restriction [a_j, b)$ for some $i \neq j$. It follows that $a_i \sim_n a_j$, and that $\sim_n$ has antichains of length at most $\#\Phi_n$. $\square$

It follows that $\sim_n$ has an unbounded equivalence class $B \in \mathcal{P}(\mathcal{M}^\iota)$. We now show that B can be taken as a predicate.

Lemma 7.4 *There is an unbounded predicate $B \subseteq U$ such that for all $a, b \in B$ we have $a \sim_n b$.*

Proof. Assume toward a contradiction that all predicates $B \subseteq U$ such that $a \sim_n b$ for all $a, b \in B$ are bounded.

We define a sequence $(B_k, b_k)_{k \in \mathbb{N}}$ where B_k is a predicate $\subseteq U$ and $b_k \in B_k$. The sequence is defined by (external) induction on $k \in \mathbb{N}$. Let b_0 be the least element of U (recall that $\mathcal{M}$ is a model of MSO). If b_k has been defined, using the Recursion Theorem 2.10, let B_k be the unique predicate such that

$$(\mathcal{M}, , \overline{A}) \models \forall x \, (B_k \, x \longleftrightarrow [U \, x \wedge b_k \leq x \wedge \forall z \, (z < x \to B_k \, z \to z \sim_n x)]$$

Note that B_k is homogeneous for $\sim_n$: if $a, b \in B_k$ then $a \sim_n b$. Hence, by assumption, B_k is bounded. Since $\mathcal{M}$ is linearly ordered, with correct equality and satisfies the axiom of induction, let b_{k+1} be the least element of U such that $a <_{\mathcal{M}} b_{k+1}$ for all $a \in B_k$.

Since U is unbounded, all B_k's are non-empty. It follows that $b_i <_{\mathcal{M}} b_j$ for all $i < j$.

Let now $a \in B_i$ for some $i > 0$ and let $j < i$. We claim that $a \sim_n b$ for no $b \in B_j$.

- *Proof of the claim.* Since $j + 1 \leq i$, we have $a \geq_{\mathcal{M}} b_{j+1}$. If $a \sim_n b$ for some $b \in B_j$, then by transitivity of $\sim_n$ we have $a \sim_n b$ for all $b \in B_j$. Since $a \in U$ and $a >_{\mathcal{M}} b_j$, it follows by definition of B_j that $a \in B_j$, contradicting $a \geq_{\mathcal{M}} b_{j+1}$. $\square$

It follows that the sequence $(b_k)_{k \in \mathbb{N}}$ is an infinite antichain for $\sim_n$. But this is not possible since $\sim_n$ has antichains of length at most $\#\Phi_n \in \mathbb{N}$. $\square$

Since $B \in \mathcal{M}^o$ is unbounded, we can assume that it has a least element, say b_0. Using the comprehension scheme, for each $\phi \in \Phi_n$ define the predicate $A_\phi \in \mathcal{M}^o$ as follows:

$$A_\phi \quad := \quad \{a \in B \mid a >_{\mathcal{M}} b_0 \text{ and } (\mathcal{M}, , \overline{A}) \models \phi \restriction [b_0, a)\}$$

Since Φ_n is finite and B is unbounded, by the Infinite Pigeonhole Principle (Proposition 7.2), there is a $\phi \in \Phi_n$ such that A_ϕ is unbounded.

We define our candidate homogeneous predicate H^q_n as follows. Using the Recursion Theorem 2.10, let H^q_n be the unique predicate such that

$$(\mathcal{M}, , \overline{A}) \models \forall x \, (H^q_n x \longleftrightarrow [A_\phi x \wedge \forall z \, (z < x \to H^q_n z \to \phi \restriction [z, x))]) \tag{1}$$

The predicate H^q_n is clearly homogeneous for ϕ. Hence we have proved Theorem 7.3 as soon as we can show that H^q_n is unbounded.

This will be done in the next two sections. First, we deal in Section 7.1.2 with the case of MSO^ω-models. This will give a proof of Theorem 5.4. We give the details because the corresponding Lemma will be useful in Section 7.1.3, where we consider the case of MSO-models.

7.1.2. Ramseyan Factorizations in MSO$^\omega$

In this section, we show that H^q_n has no maximal element, *i.e.* there is no $a \in H^q_n$ such that $b \leq_{\mathcal{M}} a$ for all $b \in H^q_n$. If $\mathcal{M}$ is a model of MSO$^\omega$, since H^q_n is not empty this implies that H^q_n is unbounded thanks to Remark 2.11. We thus get a proof of Theorem 5.4.

The proof that H^q_n has no maximal element follows usual patterns (see e.g. [PP04]).

Lemma 7.5 *The predicate H^q_n defined by (1) has no maximal element.*

Proof. Assume toward a contradiction that H^q_n has a maximal element, say c. Since b_0 and c both belong to B, by definition of $\sim_n$ there is some $d \in U$ such that $d >_{\mathcal{M}} b_0, c$ and $(\mathcal{M}, , \overline{A})\restriction[b_0, d) \equiv_n (\mathcal{M}, , \overline{A})\restriction[c, d)$. By the Finite Sums Lemma 4.3, we can assume that $d \in A_\phi$ since A_ϕ is unbounded.

Since $d \in A_\phi$, we have $(\mathcal{M}, , \overline{A}) \models \phi\restriction[b_0, d)$, and it follows that $(\mathcal{M}, , \overline{A}) \models \phi\restriction[c, d)$. We now claim that $(\mathcal{M}, , \overline{A}) \models \phi\restriction[b, d)$ for all $b \in H^q_n$ with $b <_{\mathcal{M}} d$.

- *Proof of the claim.* Recall that $\mathcal{M}$ has correct equality. Since c is the maximal element of H^q_n, it follows that for all $b \in H^q_n$ different from c we can split $(\mathcal{M}, , \overline{A})\restriction[b, d)$ into

$$(\mathcal{M}, , \overline{A})\restriction[b, c) \qquad \text{and} \qquad (\mathcal{M}, , \overline{A})\restriction[c, d)$$

  Now, since $(\mathcal{M}, , \overline{A})\restriction[b, c) \equiv_n (\mathcal{M}, , \overline{A})\restriction[b_0, c)$ by definition of H^q_n, the Finite Sums Lemma 4.3 implies that $(\mathcal{M}, , \overline{A}) \models \phi\restriction[b, d)$. □

It follows that $d \in H^q_n$, a contradiction since $c <_{\mathcal{M}} d$. This concludes the proof of Lemma 7.5. □

Corollary 7.6 *Models of MSO$^\omega$ have Ramseyan $\equiv^q_n$-factorizations for every $n, q \in \mathbb{N}$.*

7.1.3. Ramseyan Factorizations in MSO

We now show that H^q_n is unbounded when $\mathcal{M}$ is not a model of MSO$^\omega$. The important point is that given $a \in \mathcal{M}^\iota$, the predicate $[-, a) = \{b \mid b <_{\mathcal{M}} a\}$ may not have a maximal element.

Lemma 7.7 *The predicate H^q_n defined by (1) is unbounded.*

Proof. Assume toward a contradiction that H^q_n is not unbounded in $\mathcal{M}$. Since $\mathcal{M}$ is a linear order with correct equality, there is some $a \in \mathcal{M}^\iota$ such that $b \leq_{\mathcal{M}} a$ for all $b \in H^q_n$. Since H^q_n has no maximal element by Lemma 7.5, we moreover have $a \notin H^q_n$, hence $b <_{\mathcal{M}} a$ for all $b \in H^q_n$. Since $\mathcal{M}$ satisfies the axiom of induction, there is a least such a, write it a_0.

We are interested in segments of the form $(\mathcal{M}, , \overline{A})\restriction[b, a_0)$ where $b \in H^q_n \cup \{b_0\}$. We claim that H^q_n is unbounded in such segments: for all $c \in \mathcal{M}^\iota\restriction[b, a_0)$ there is some $d \in H^q_n$ such that $c <_{\mathcal{M}} d <_{\mathcal{M}} a_0$.

- *Proof of the claim.* If $c \in \mathcal{M}^\iota\restriction[-, a_0)$ we have $c <_{\mathcal{M}} a_0$. Since a_0 is the least upper bound of H^q_n, there is some $e \in H^q_n$ such that $c \leq_{\mathcal{M}} e <_{\mathcal{M}} a_0$. Moreover, since H^q_n has no maximal element, it contains some d such that $e <_{\mathcal{M}} d <_{\mathcal{M}} a_0$. □

We now show that

$$(\mathcal{M},,\overline{A})\restriction[b_0,a_0) \equiv_n (\mathcal{M},,\overline{A})\restriction[b,a_0) \qquad \text{for all } b \in H^q_n \tag{2}$$

- *Proof.* Let $b \in H^q_n$. Since H^q_n is unbounded in $[b,a_0)$, there is some $c \in H^q_n$ such that $b <_{\mathcal{M}} c <_{\mathcal{M}} a_0$. Now, since $c \in A_\phi$, we have $(\mathcal{M},,\overline{A}) \models \phi\restriction[b_0,c)$. On the other hand, since $b,c \in H^q_n$, we have $(\mathcal{M},,\overline{A}) \models \phi\restriction[b,c)$. It follows that

  $$(\mathcal{M},,\overline{A})\restriction[b_0,c) \equiv_n (\mathcal{M},,\overline{A})\restriction[b,c)$$

  and we conclude by the Finite Sums Lemma 4.3. □

It is now easy to derive a contradiction. Since A_ϕ is unbounded, let $a \in A_\phi$ be $>_{\mathcal{M}} a_0$. In particular $a \notin H^q_n$. Thanks to Finite Sums Lemma 4.3, for all $b \in H^q_n$ we get from (2)

$$(\mathcal{M},,\overline{A})\restriction[b_0,a) \equiv_n (\mathcal{M},,\overline{A})\restriction[b,a)$$

But this contradicts $a \in A_\phi \setminus H^q_n$ because

(i) $(\mathcal{M},,\overline{A}) \models \phi\restriction[b_0,a)$ since $a \in A_\phi$, and

(ii) there is some $b \in H^q_n$ such that $(\mathcal{M},,\overline{A}) \not\models \phi\restriction[b,a)$ since $a \notin H^q_n$.

This completes the proof of Lemma 7.7. □

This concludes the proof of Theorem 7.3.

7.2. Ramsey's Theorem from Ramseyan Factorizations

We show here that in linearly ordered second-order structures with correct equality, Ramsey's theorem for colorings definable by MSO-formulas follows from the existence of Ramseyan factorizations.

Thanks to Theorem 7.3 and Corollary 2.6, this implies Theorem 7.1.

Theorem 7.8 *Let $\mathcal{M}$ be a second-order linearly ordered structure with correct equality. Assume that $\mathcal{M}$ has Ramseyan $\equiv^q_n$-factorizations for every $n, q \in \mathbb{N}$.*

Let $\phi_1, \dots, \phi_k$ be formulas with parameters in $\mathcal{M}$ and $U \in \mathcal{M}^o$ be an unbounded predicate. Assume that for all $(a,b) \in [U]^2$ we have $\mathcal{M} \models \bigvee_{1 \le i \le k} \phi_i[a/x, b/y]$.

Then there is an unbounded predicate $V \subseteq U$ and an $i \in \{1, \dots, k\}$ such that $\mathcal{M} \models \phi_i[a/x, b/y]$ for all $(a,b) \in [V]^2$.

Proof. By possibly renaming free variables, we can assume that there is a finite partial $\mathcal{M}$-assignment ν such that each ϕ_i is of the form (ψ_i, ν). We can moreover assume that $\mathrm{FV}(\overline{\psi}) \subseteq \{x, y, x_1, \dots, x_p, X_1, \dots, X_q\}$ and that $\mathrm{dom}(\nu) = \{x_1, \dots, x_p, X_1, \dots, X_q\}$ for some $p, q \in \mathbb{N}$. Let $a_i := \nu(x_i)$ and $A_j := \nu(X_j)$ (for $1 \le i \le p$ and $1 \le j \le q$). It follows that for all $i \in \{1, \dots, k\}$ and all $(a,b) \in [U]^2$ we have $\mathcal{M} \models \phi_i[a/x, b/y]$ if and only if $(\mathcal{M}, \overline{a}ab, \overline{A}) \models \psi_i$.

Since U is unbounded and since $\mathcal{M}$ is linearly ordered with correct equality, we can assume that all $a \in U$ are $>_{\mathcal{M}} a_1, \dots, a_p$. Let $n \in \mathbb{N}$ be greater than $\mathrm{qd}(\psi_1), \dots, \mathrm{qd}(\psi_k)$. Thanks to Ramseyan factorizations, we get the following Lemma, from which Theorem 7.8 easily follows.

Lemma 7.9 *There is an unbounded predicate $V \subseteq U$ such that for all $(a,b),(c,d) \in [V]^2$ we have $(\mathcal{M}, \overline{a}ab, \overline{A}) \equiv_n (\mathcal{M}, \overline{a}cd, \overline{A})$.*

We now complete the proof of Theorem 7.8 using Lemma 7.9. Given $(a,b) \in [V]^2$, since $V \subseteq U$, by assumption there is some $i \in \{1,\dots,k\}$ such that $\mathcal{M} \models \phi_i[a/x, b/y]$, hence $(\mathcal{M}, \overline{a}ab, \overline{A}) \models \psi_i$. By Lemma 7.9, since $\mathrm{qd}(\psi_i) \leq n$, this implies that $(\mathcal{M}, \overline{a}cd, \overline{A}) \models \psi_i$ for all $(c,d) \in [V]^2$, and we are done. □

Proof of Lemma 7.9. Given $(a,b) \in [U]^2$, the idea is to split the $(p+2)$-q-structure $(\mathcal{M}, \overline{a}ab, \overline{A})$ into

$$(\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-,a) \qquad (\mathcal{M}, a, \overline{A}){\restriction}[a,b) \qquad (\mathcal{M}, b, \overline{A}){\restriction}[b,-)$$

We then use the Infinite Pigeonhole principle (Proposition 7.2) in the first and last components, and Ramseyan factorizations in the second one.

Let us begin with the first and last components. Recall from Corollary 3.4 that $\equiv_n$-equivalence for $((\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-,a))_{a \in U}$ can be characterized by a finite set Ψ of formulas with parameters $\overline{a}, \overline{A}$ and a single free individual variable. Since $\mathcal{M}$ is second-order, there are predicates $(A_\psi)_{\psi \in \Psi} \in \mathcal{M}^o$ such that for all $a, b \in U$, we have have $a, b \in A_\psi$ if and only if $(\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-,a) \equiv_n (\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-,b)$. The same holds for $((\mathcal{M}, b, \overline{A}){\restriction}[b,-))_{b \in U}$, and by two applications of Proposition 7.2, we get an unbounded predicate $W \subseteq U$ such that for all $a, b \in W$ we have

$$(\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-,a) \equiv_n (\mathcal{M}, \overline{a}, \overline{A}){\restriction}[-,b) \qquad \text{and} \qquad (\mathcal{M}, a, \overline{A}){\restriction}[a,-) \equiv_n (\mathcal{M}, b, \overline{A}){\restriction}[b,-)$$

We now consider the central components $((\mathcal{M}, a, \overline{A}){\restriction}[a,b))_{(a,b) \in [W]^2}$. Let ϕ be a 1-q-formula of q.d. $\leq n$. Then the 0-q-formula $\tilde{\phi} := \exists x \forall y\, (x \leq y \wedge \phi[x/x_1])$ has q.d. $\leq n+3$. Moreover, since $\mathcal{M}$ has correct equality, given $(a,b) \in [W]^2$ we have $(\mathcal{M}, a, \overline{A}){\restriction}[a,b) \models \phi$ if and only if $(\mathcal{M}, , \overline{A}){\restriction}[a,b) \models \tilde{\phi}$. It follows that for all $(a,b),(c,d) \in [W]^2$ we have

$$(\mathcal{M}, a, \overline{A}){\restriction}[a,b) \equiv_n (\mathcal{M}, c, \overline{A}){\restriction}[c,d) \qquad \text{if} \qquad (\mathcal{M}, , \overline{A}){\restriction}[a,b) \equiv_{n+3} (\mathcal{M}, , \overline{A}){\restriction}[c,d)$$

Now, since $\mathcal{M}$ has $\equiv^q_{n+3}$-factorizations, we get an unbounded predicate $V \subseteq W$ such that $(\mathcal{M}, a, \overline{A}){\restriction}[a,b) \equiv_n (\mathcal{M}, c, \overline{A}){\restriction}[c,d)$ for all $(a,b),(c,d) \in [V]^2$.

Thanks to the Finite Sums Lemma 4.3, it follows that for all $(a,b),(c,d) \in [V]^2$ we have

$$(\mathcal{M}, \overline{a}ab, \overline{A}) \equiv_n (\mathcal{M}, \overline{a}cd, \overline{A})$$

□

8. Conclusion

We gave a model-theoretic proof of Siefkes' completeness result for MSO^ω [Sie70]. It is based on Ramsey's Theorem for additive colorings, with constructions reminiscent from algebraic approaches to ω-rational languages [PP04]. Further works will begin by clarifying these relationships. An interesting question is the proof-theoretic analysis of MSO^ω. The algebraic approach to parity conditions [PP04] can be interesting in this perspective. An other direction is the completeness of MSO on infinite trees, and the comparison with Walukiewicz completeness result for the μ-calculus [Wal00].

A. Proof of the Recursion Theorem

In this section, we give a proof of the course-of-values Recursion Theorem 2.10. For models of MSO^ω, it is Theorem 1.b.1 of [Sie70]. We prove it for ordered structures which satisfy the induction axiom.

Let $\mathcal{M}$ be a second-order structure with correct equality. Assume that $\mathcal{M}$ satisfies the induction axiom and moreover that:

$$\mathcal{M} \models \forall x \neg(x < x) \qquad \text{and} \qquad \mathcal{M} \models \forall xyz\,(x < y \to y < z \to x < z)$$

Note that if $a <_{\mathcal{M}} b$ then $b \not<_{\mathcal{M}} a$.

Let $\phi[X < x]$ be a formula with parameters in $\mathcal{M}$ and whose free variables are among $X \in \mathcal{V}_o$ and $x \in \mathcal{V}_\iota$. Assume that ϕ is independent from X for the y such that $\neg(y < x)$:

$$\mathcal{M} \models \forall x \forall XY\,[\forall y < x\,(Xy \longleftrightarrow Yy) \to (\phi[X < x] \longleftrightarrow \phi[Y < x])] \tag{3}$$

We begin by the following Lemma, which will give us the unicity part of Theorem 2.10:

Lemma A.1 *For all $a \in \mathcal{M}^\iota$ and $A, B \in \mathcal{M}^o$ we have*

$$\mathcal{M} \models \forall x \leq a[Ax \longleftrightarrow \phi[A < x]] \to \forall x \leq a[Bx \longleftrightarrow \phi[B < x]] \to \forall x \leq a[Ax \longleftrightarrow Bx]$$

Proof of Lemma A.1. We apply the induction axiom. Let $a \in \mathcal{M}^\iota$ and assume the property for all $b <_{\mathcal{M}} a$.

Let $c \leq_{\mathcal{M}} a$. If $c <_{\mathcal{M}} a$ then using the transitivity of $<_{\mathcal{M}}$ we are done by induction hypothesis. It remains to show that $\mathcal{M} \models Aa \longleftrightarrow Ba$. Using the assumption, it is sufficient to show that $\mathcal{M} \models \phi[A < a] \longleftrightarrow \phi[B < a]$. But by induction hypothesis, we have $A{\restriction}[-, a) = B{\restriction}[-, a)$ and we are done by (3).

This concludes the proof of Lemma A.1. $\square$

Using the comprehension scheme let $R_\phi \in \mathcal{M}^o$ be such that for all $a \in \mathcal{M}^\iota$,

$$\mathcal{M} \models R_\phi\, a \longleftrightarrow \forall X\,[\forall x \leq a(Xx \longleftrightarrow \phi[X < x]) \to Xa]$$

We have to show that R_ϕ is correct for ϕ:

$$\mathcal{M} \models \forall x(R_\phi\, x \longleftrightarrow \phi[R_\phi < x])$$

The last part of Theorem 2.10 will then follow from Lemma A.1.

That R_ϕ is correct for ϕ follows from the following Lemma, which will conclude the proof of Theorem 2.10.

Lemma A.2 *For all $a \in \mathcal{M}^\iota$ we have*

$$\mathcal{M} \models \forall x \leq a(R_\phi\, x \longleftrightarrow \phi[R_\phi < x])$$

Proof of Lemma A.2. We apply the induction axiom. Let $a \in \mathcal{M}^\iota$ and assume the property for all $b <_\mathcal{M} a$.

Let $c \leq_\mathcal{M} a$. If $c <_\mathcal{M} a$ then using the transitivity of $<_\mathcal{M}$ we are done by induction hypothesis. It remains to show that

$$\mathcal{M} \models R_\phi\, a \longleftrightarrow \phi[R_\phi < a]$$

Assume that $\mathcal{M} \models \phi[R_\phi < a]$. We have to show that $a \in R_\phi$. By definition of R_ϕ, we have to show that $\mathcal{M} \models \forall X[\forall x \leq a(Xx \longleftrightarrow \phi[X < x]) \to Xa]$. Let $A \in \mathcal{M}^o$ such that $\mathcal{M} \models \forall x \leq a(Ax \longleftrightarrow \phi[A < x])$. We must show $a \in A$, *i.e.* $\mathcal{M} \models \phi[A < a]$. Since $\mathcal{M} \models \phi[R_\phi < a]$, by (3) we are done if $R_\phi\restriction[-,a) = A\restriction[-,a)$. Let $d <_\mathcal{M} a$. By Lemma A.1 we get $\mathcal{M} \models \forall x \leq d(R_\phi x \longleftrightarrow Ax)$ from the induction hypothesis applied to the $e \leq_\mathcal{M} d$ and the assumption on A. It follows that $R_\phi\restriction[-,a) = A\restriction[-,a)$.

Assume now that $a \in R_\phi$. Using the comprehension scheme, let $A \in \mathcal{M}^o$ be such that for all $b \in \mathcal{M}^\iota$,

$$\mathcal{M} \models Ab \longleftrightarrow [(b < a \wedge R_\phi\, b) \vee (b \doteq a \wedge \phi[R_\phi < b])]$$

By definition of R_ϕ, we are done if $\mathcal{M} \models \forall x \leq a(Ax \longleftrightarrow \phi[A < x])$. Given $b <_\mathcal{M} a$, we have $\mathcal{M} \models Ab \longleftrightarrow R_\phi\, b$ and we conclude by induction hypothesis. For the case of $b = a$, we have to show $\mathcal{M} \models \phi[R_\phi < a] \longleftrightarrow \phi[A < a]$. This follows from (3) since $R_\phi\restriction[-,a) = A\restriction[-,a)$.

This concludes the proof Lemma A.2. $\square$

B. Proof of the Splicing axiom scheme in MSO^ω

In this section, we give a proof of Siefkes' result that models of MSO^ω satisfy the Splicing axiom scheme (Definition 6.1). This is Theorem I.5.b.1 of [Sie70] (see also Theorem 4.1 of [BS73]). We provide a proof here for completeness.

Theorem B.1 (Splicing for MSO^ω) *Let $\mathcal{M}$ be a model of MSO^ω with correct equality. Then $\mathcal{M}$ satisfies the Splicing axiom scheme:*

For all $q \in \mathbb{N}$, all $A_1, \dots, A_q \in \mathcal{M}^o$, all $U \in \mathcal{M}^o$, and all 0-$(q+1)$-formula ϕ, if for all $(a,b) \in [U]^2$ we have $(\mathcal{M},, \overline{A}A^{a,b}) \models \phi\restriction[a,b)$ for some $A^{a,b} \in \mathcal{M}^o$, then there is a predicate $A \in \mathcal{M}^o$ such that for all a, b consecutive *in U we have $(\mathcal{M},, \overline{A}A) \models \phi\restriction[a,b)$.*

Proof. Let $\mathcal{M}$ be a model of MSO^ω and $U \in \mathcal{M}$. Consider a formula $\phi[X]$ with parameters in $\mathcal{M}$ and with only one free variable $X \in \mathcal{V}_o$. Assume that for all $(a,b) \in [U]^2$, there is some $A \in \mathcal{M}^o$ such that $\mathcal{M} \models \phi[A]\restriction[a,b)$. We want to define an $A \in \mathcal{M}^o$ such that $\mathcal{M} \models \phi[A]\restriction[a,b)$ for all consecutive $(a,b) \in [U]^2$.

In order to define such an $A \in \mathcal{M}^o$, we have to select for each segment $[a,b)$ (with a, b consecutive in U) a *unique* boolean $[a,b)$-sequence $A\restriction[a,b)$. This boolean $[a,b)$-sequence has to be defined by a formula $\tilde{\phi}[X,a,b]$ which depends uniformly on a, b. For this, we build a formula $\tilde{\phi}[X,x,y]$ which *uniformizes*[3] $\phi[X]$ on segments $[a,b)$:

[3]in the sense of [Sim10].

(a) $\mathcal{M} \models \exists X\, (\phi[X]{\restriction}[a,b)) \to \exists X\, \tilde{\phi}[X,a,b]$,

(b) $\mathcal{M} \models \forall X\, (\tilde{\phi}[X,a,b] \to \phi[X]{\restriction}[a,b))$,

(c) $\mathcal{M} \models \forall XY\, (\tilde{\phi}[X,a,b] \to \tilde{\phi}[Y,a,b] \to \forall x\, (a \le x < b \to [Xx \longleftrightarrow Yx]))$

Given $\tilde{\phi}$ satisfying (a), (b) and (c), define $\tilde{A} \in \mathcal{M}^o$ as follows:

$$\begin{array}{rcl} \tilde{A} & := & \{c \mid a \le_{\mathcal{M}} c <_{\mathcal{M}} b \text{ and } c \in B \\ & & \qquad \text{for some } B \in \mathcal{M}^o \text{ and consecutive } a,b \in U \text{ such that } \mathcal{M} \models \tilde{\phi}[B,a,b]\} \end{array}$$

We show that $\tilde{A}$ is correct. Let a,b be consecutive in U. Since $\mathcal{M} \models \exists X\, \phi[X]{\restriction}[a,b)$, by (a) above there is $B \in \mathcal{M}^o$ such that $\mathcal{M} \models \tilde{\phi}[B,a,b]$. This B is moreover unique on the segment $[a,b)$ by (c). It follows that $B{\restriction}[a,b) = \tilde{A}{\restriction}[a,b)$ since a,b are consecutive in U. We then get $\mathcal{M} \models \phi[\tilde{A}]{\restriction}[a,b)$ by (b).

Let us now build $\tilde{\phi}[X,x,y]$ such that (a), (b) and (c) hold. The idea (also used in [BS73]) is to lexicographically order the boolean $[a,b)$-sequences $A \in \mathcal{M}^o{\restriction}[a,b)$. We can then define $\tilde{\phi}[X,a,b]$ so that it holds on $A \in \mathcal{M}^o$ if $A{\restriction}[a,b)$ is minimal such that $\mathcal{M} \models \phi[A]{\restriction}[a,b)$. Let

$$\begin{array}{rcl} \tilde{\phi}[X,a,b] & := & \phi[X]{\restriction}[a,b) \wedge \\ & & \forall Y\, (\phi[Y]{\restriction}[a,b) \to \exists x[a \le x < b \wedge \neg(Xx \longleftrightarrow Yx)] \to \\ & & \quad \exists y[\forall x < y(Xx \longleftrightarrow Yx) \wedge \neg Xy \wedge Yy]) \end{array}$$

Property (b) is obvious if $(a,b) \in [U]^2$. For (c), given $(a,b) \in [U]^2$ let $A,B \in \mathcal{M}^o$ such that $\mathcal{M} \models \phi[A,a,b] \wedge \tilde{\phi}[B,a,b]$. If $A{\restriction}[a,b) \neq B{\restriction}[a,b)$, then by the induction axiom there is a least $c \in [a,b)$ such that $\mathcal{M} \models \neg(Ac \longleftrightarrow Bc)$. Then we have say $c \notin A$ and $c \in B$, contradicting $\mathcal{M} \models \tilde{\phi}[B,a,b)$.

To show (a) we proceed by course-of-values recursion. Fix $(a,b) \in [U]^2$. By Theorem 2.10, let A be the unique predicate such that

$$\mathcal{M} \models \forall x\, (Ax \longleftrightarrow \forall Z\, [\forall z\, (a \le z < x \to [Zz \longleftrightarrow Az]) \to \phi[Z]{\restriction}[a,b) \to Zx]) \tag{4}$$

Lemma B.2 *If $\mathcal{M} \models \exists X\, (\phi[X]{\restriction}[a,b))$ then*

$$\mathcal{M} \models \forall x \ge a\, \exists Z\, [\forall z\, (a \le z < x \to [Zz \longleftrightarrow Az]) \wedge \phi[Z]{\restriction}[a,b)]$$

Proof of Lemma B.2. We use the induction axiom (on $x \ge a$). Let $c \ge_{\mathcal{M}} a$ such that the property holds for all $d \in [a,c)$.

If $c = a$ (recall from Definition 2.7 that $\mathcal{M}$ has correct equality), take for Z some $B \in \mathcal{M}^o$ such that $\mathcal{M} \models \phi[B]{\restriction}[a,b)$. Such B exists by assumption, and we have $B{\restriction}[a,a) = A{\restriction}[a,a)$ since $[a,a) = \emptyset$.

Otherwise, let $d \ge_{\mathcal{M}} a$ be the predecessor of c.

- If $c \notin A$, then by (4) there is a $B \in \mathcal{M}^o$ such that $\mathcal{M} \models \phi[B]{\restriction}[a,b)$, $B{\restriction}[a,d) = A{\restriction}[a,d)$ and $c \notin B$. We thus have $B{\restriction}[a,c) = A{\restriction}[a,c)$ and we are done.

- Otherwise, by induction hypothesis, let $B \in \mathcal{M}^o$ be such that $\mathcal{M} \models \phi[B]{\restriction}[a,b)$ and $B{\restriction}[a,d) = A{\restriction}[a,d)$. Since $c \in A$ we have $c \in B$ by (4). Hence $B{\restriction}[a,c) = A{\restriction}[a,c)$ and we are done.

This concludes the proof of Lemma B.2. $\square$

Assume that $\mathcal{M} \models \phi[B]{\restriction}[a,b)$ for some $B \in \mathcal{M}^o$. The above Lemma implies that there is $C \in \mathcal{M}^o$ such that $\mathcal{M} \models \forall z\,(a \leq z < b \to [Cz \longleftrightarrow Az]) \wedge \phi[C]{\restriction}[a,b)$, hence that $\mathcal{M} \models \phi[A]{\restriction}[a,b)$ since X is bounded in $\phi[X]{\restriction}[a,b)$.

Let now be $D \in \mathcal{M}^o$ such that $\mathcal{M} \models \phi[D]{\restriction}[a,b)$ and $A{\restriction}[a,b) \neq D{\restriction}[a,b)$. By the minimum principle, there is a least $c \in [a,b)$ such that $\mathcal{M} \models \neg(Ac \longleftrightarrow Dc)$. We must show that $c \notin A$ and $c \in D$. Assume toward a contradiction that $c \in A$. Then by (4) we have

$$\mathcal{M} \models \forall z\,(a \leq z < c \to [Dz \longleftrightarrow Az]) \to \phi[D]{\restriction}[a,b) \to Dc$$

and we deduce $c \in D$ by assumption on c and D. But this contradicts the definition of c. Hence we deduce that $c \in A$, and thus $c \in D$. It follows that $\mathcal{M} \models \tilde{\phi}[A,a,b]$.

This concludes the proof of Theorem B.1. $\square$

References

[BBJ07] G. S. Boolos, J. P. Burgess, and R. C. Jeffrey. *Computability and Logic*. Cambridge University Press, fifth edition, 2007. 10

[BS73] J. R. Büchi and D. Siefkes. Axiomatization of the Monadic Second Order Theory of ω_1. In J. R. Büchi and D. Siefkes, editors, *Decidable Theories II : The Monadic Second Order Theory of All Countable Ordinals*, volume 328 of *LNM*, pages 129–217. Springer, 1973. 3, 18, 21, 22, 32, 33

[Büc62] J. R. Büchi. On a Decision Methond in Restricted Second-Order Arithmetic. In E. Nagel et al., editor, *Logic, Methodology and Philosophy of Science (Proc. 1960 Intern. Congr.)*, pages 1–11. Stanford Univ. Press, 1962. 3, 8

[Doe89] K. Doets. Monadic Π^1_1-Theories of Π^1_1-Properties. *Notre Dame Journal of Formal Logic*, 30(2):224–240, 1989. 3, 19

[EF99] H.-D. Ebbinghaus and J. Flum. *Finite Model Theory*. Springer Monographs in Mathematics. Springer, second edition, 1999. 10, 11, 12

[GS83] Y. Gurevich and S. Shelah. Interpreting Second-Order Logic in the Monadic Theory of Order. *Journal of Symbolic Logic*, 48(3):816–828, 1983. 10

[GtC09] A. Gheerbrant and B. ten Cate. Complete Axiomatizations of MSO, FO(TC1) and FO(LFP1) on Finite Trees. In *LFCS*, volume 5407 of *LNCS*, pages 180–196. Springer, 2009. 3, 8, 9, 12, 13, 17

[GTW02] E. Grädel, W. Thomas, and T. Wilke, editors. *Automata, Logics, and Infinite Games: A Guide to Current Research*, volume 2500 of *Lecture Notes in Computer Science*. Springer, 2002. 3

[Gur85] Y. Gurevich. Monadic Second-Order Theories. In J. Barwise and S. Feferman, editors, *Model-Theoretical Logics*, Perspective in Mathematical Logic, pages 479–506. Springer, 1985. 3, 9, 13

[PP04] D. Perrin and J.-É. Pin. *Infinite Words: Automata, Semigroups, Logic and Games*. Pure and Applied Mathematics. Elsevier, 2004. 3, 19, 20, 26, 28, 30

[Ros82] J. G. Rosenstein. *Linear Orderings*. Pure and Applied Mathematics. Academic Press, 1982. 12, 25

[Sha91] S. Shapiro. *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford University Press, 1991. 4, 6, 7, 10

[She75] S. Shelah. The Monadic Theory of Order. *The Annals of Mathematics, Second Series*, 102(3):379–419, 1975. 3, 10, 13, 17, 19, 26

[Sie70] D. Siefkes. *Decidable Theories I : Büchi's Monadic Second Order Successor Arithmetic*, volume 120 of *LNM*. Springer, 1970. 3, 8, 10, 18, 19, 20, 21, 22, 30, 31, 32

[Sim10] S.G. Simpson. *Subsystems of Second Order Arithmetic*. Perspectives in Logic. Cambridge University Press, 2nd edition, 2010. 32

[Tho88] W. Thomas. Automata on Infinite Objects. Technical Report AIB-88-17, RWTH Aachen, October 1988. 26

[Tho96] W. Thomas. Languages, Automata, and Logic. Technical Report 9607, Institut für Informatik und Praktische Mathematik, Christian-Albrechts-Universität zu Kiel, Germany, May 1996. 3

[vD04] D. van Dalen. *Logic and Structure*. Universitext. Springer, fourth edition, 2004. 10, 25

[Wal00] I. Walukiewicz. Completeness of Kozen's Axiomatisation of the Propositional μ-Calculus. *Information and Computation*, 157(1-2):142–182, 2000. 30