Forcing MSO on Infinite Words in Weak MSO

Colin Riba

ENS de Lyon, Université de Lyon, LIP

colin.riba@ens-lyon.fr

http://perso.ens-lyon.fr/colin.riba/

Abstract—We propose a forcing-based interpretation of monadic second-order logic (MSO) on infinite (omega) words in Weak MSO (WMSO). The interpretation is purely syntactic.

We show that a formula with parameters is true in MSO if and only if its interpretation is true in WMSO. We also show that a closed formula is true in MSO if and only if its interpretation is provable under some axioms which hold for WMSO, but without axiomatizing it.

We use model-theoretic arguments. Our approach is inspired from point-free topology: infinite words, seen as topological points, are approximated by filters of bounded segments. We devise forcing conditions such that the corresponding generic filters approximate Ramseyan factorizations of infinite words modulo satisfaction of formulas of a given quantifier depth.

Our interpretation parallels some approaches to Mc-Naughton's Theorem (equivalence between non-deterministic Büchi automata and deterministic Rabin automata) but the obtained formulas do not describe deterministic automata.

I. Introduction

Monadic Second-Order Logic (MSO) on ω-words is known to be decidable since Büchi's work [2]. Its method was to translate formulas to finite state automata running on ω-words. Such automata can in turn be represented by formulas, leading to notions of *automata normal forms* for ω-words.

The automata used in [2], called *Büchi automata*, are usual finite state non-deterministic automata running on infinite words. Their (infinite) accepting runs are those runs which reach a final state infinitely often. In general, Büchi automata can not be determinized.

McNaughton's Theorem (see e.g. [15], [9]) states that non-deterministic Büchi automata are equivalent to *deterministic Rabin automata*. In Rabin automata normal forms, second-order quantifications are only made over bounded predicates. On such formulas, MSO is equivalent to Weak MSO (WMSO) where second-order quantifications are only allowed to range over finite sets. Hence, McNaughton's Theorem provides a translation from MSO to WMSO.

However, such translations lack compositionality on the side of automata, since the operations of projection, union and complementation (which reflect the connectives of MSO) may radically change the structure of automata.

The main aim of this work is to propose a more compositional translation of MSO to WMSO.

In this paper, we devise a forcing interpretation of MSO in WMSO. It is based on the following idea, borrowed from point-free topology: Sequences indexed by a linear order can

be approximated by suitable *filters*. Forcing requires quite special filters, called *generic filters*. We work in the general framework of usual filters since it allows smooth statements and manipulations, in particular when iterating forcing. The good notion of approximation is provided by Ramseyan factorizations, applied to additive colorings of bounded segments by monadic theories.

In contrast with usual approaches (see e.g. [11], [14]), we work with general *Henkin structures*. In Henkin structures, second-order quantification is allowed only on a specified predicate domain, which may not contain all sets of individuals. We work with such general structures mainly because they allow Skolem-Löwenheim arguments, which greatly ease reasoning with iterated forcing.

We also get Henkin completeness. We thus obtain provability results by translating every formula true on ω-words to a formula which holds in any linearly ordered Henkin structure satisfying some arithmetic axioms, and whose predicate domain is only required to contain all definable *bounded* sets. This last requirement is of course satisfied by the full monadic structure of natural numbers, but also by the Henkin structure whose individuals are the natural numbers and whose predicates are the *finite* sets of natural numbers.

Our interpretation parallels some approaches to Mc-Naughton's Theorem. In particular, the way we handle Ramseyan factorizations is inspired from [3] (see also [13]) and also shares some similarities with the algebraic proof of determinization given in [9]. However, the obtained formulas do not by themselves describe deterministic automata.

Organization of the Paper: We begin in Section II by presenting the basic material on MSO and Henkin structures that we will need. We then discuss in Section III results on the additive coloring of segments by MSO theories and the corresponding Ramseyan factorizations. The general setting of filters is then discussed in Section IV. Section V provides an overview of the notions of forcing used in our translation. Section VI then presents the forcing conditions, and the transformation is discussed in Section VII.

II. Monadic Theories of Henkin Structures

This section gathers basic material on Henkin models of second-order logic. After having defined linearly ordered Henkin structures in Section II-A, we present a language for MSO in Section II-B. We recall in Section II-C some basic facts on deduction and completeness for Henkin structures, and discuss in Section II-D relativizations and restrictions.

We finally turn in Section II-E to axiomatics for MSO with bounded comprehension.

A. Linearly Ordered Henkin Structures

1) Linear Orders: A *linear order* is a non-empty set L equipped with a binary relation $<_L$ which is

- **irreflexive:** for all $a \in L$, $a \not<_L a$,
- **transitive:** for all $a, b, c \in L$,
  if $a <_L b$ and $b <_L c$ then $a <_L c$, and
- **total:** for all distinct $a, b \in L$, either $a <_L b$ or $b <_L a$.

A linear order $(L, <_L)$ is *unbounded* if for all $a \in L$, there is $b \in L$ such that $a <_L b$.

2) Henkin Structures: A linearly ordered *Henkin structure* $\mathcal{M}$ is of the form $(\mathcal{M}^\iota, \mathcal{M}^o, <_\mathcal{M})$, 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 linear order on $\mathcal{M}^\iota$.

We are mostly interested in the standard model of ω-words

$$\mathsf{N} \quad := \quad (\mathbb{N}, \mathcal{P}(\mathbb{N}), <)$$

Note that Henkin structures may be not full: $\mathcal{M}^o$ may not contain all $A \subseteq \mathcal{M}^\iota$. For instance, the weak standard model

$$\mathsf{WN} \quad := \quad (\mathbb{N}, \mathcal{P}_{\text{fin}}(\mathbb{N}), <) \ ,$$

whose predicate domain consists of the *finite* subsets of $\mathbb{N}$, is a Henkin structure.

A *p-q-expanded structure* is a structure $\mathcal{M}$ together with $a_1, \ldots, a_p \in \mathcal{M}^\iota$ and $A_1, \ldots, A_q \subseteq \mathcal{M}^\iota$. Note that we do *not* require the A_i's to be predicates of $\mathcal{M}$. Write $\overline{a}$ for a finite sequence of individuals of length $|\overline{a}|$, and similarly for $\overline{A}$.

3) Representation of Words: A linearly ordered Henkin structure $\mathcal{M}$ expanded with sets $A_1, \ldots, A_q$ can be seen as a word on the alphabet $\{0, 1\}^q$, indexed by the linear order $(\mathcal{M}^\iota, <_\mathcal{M})$. Its character at position a is $(b_1, \ldots, b_q)$, where $b_i = 1$ iff $a \in A_i$. In particular, ω-words on $\{0, 1\}^q$ are expanded structures $(\mathsf{N}, , A_1 \ldots A_q)$ with $A_1, \ldots, A_q \subseteq \mathbb{N}$. For instance, if $\text{Odd} \subseteq \mathbb{N}$ is the set of odd numbers, then $(\mathsf{N}, , \text{Odd})$ represents the ω-word

$$(01)^\omega \quad := \quad 01010101 \ldots$$

B. Monadic Quantification

1) MSO for Henkin Structures: We assume given two countable sets $\mathcal{V}^\iota = \{x, y, z, \ldots\}$ and $\mathcal{V}^o = \{X, Y, Z, \ldots\}$ of respectively individual and predicate variables.

The atomic formulas are membership Xx, equality[1] on individuals $x = y$ and comparison $x < y$. Formulas $\phi \in \Lambda$ are build from atomic formulas by the propositional connectives $\neg\phi$ and $\phi \vee \psi$, and the existential quantifiers $\exists x\phi$ and $\exists X\phi$.

The other logical connectives ($\rightarrow$, $\wedge$, $\longleftrightarrow$) and the universal quantifiers ($\forall x$, $\forall X$) are defined as usual from $\neg$, $\vee$ and $\exists$.

For instance, the formula

$$\text{Ub}[X] \quad := \quad \forall x \exists y (x < y \ \wedge \ Xy)$$

expresses that the set X is unbounded.

---

[1] Equality could have been defined as usual (see e.g. [16]), but the presentation is simpler by taking it primitive.

Let us fix enumerations $x_1, \ldots, x_p, \ldots$ and $X_1, \ldots, X_q, \ldots$ of $\mathcal{V}^\iota$ and $\mathcal{V}^o$. We let $\Lambda^{p,q}$ be the set of formulas with free variables among $\{x_1, \ldots, x_p, X_1, \ldots, X_q\}$. $\Lambda^{0,0}$ is the set of *sentences* (or closed formulas).

Consider a *p-q-expanded* structure $(\mathcal{M}, \overline{a}, \overline{A})$. Given a formula $\phi \in \Lambda^{p,q}$, we write

$$(\mathcal{M}, \overline{a}, \overline{A}) \models \phi$$

if ϕ satisfied in $\mathcal{M}$ when interpreting x_i by a_i and X_j by A_j, and when the individual and predicate quantifiers of ϕ range over $\mathcal{M}^\iota$ and $\mathcal{M}^o$ respectively.

For instance, if $\mathcal{M}$ is a linear order and $A \subseteq \mathcal{M}^\iota$, then $(\mathcal{M}, , A)$ satisfies $\text{Ub}[X_1]$ just when A is unbounded. This is in particular the case of $(\mathsf{N}, , \text{Odd})$.

It is often useful to consider formulas with parameters in some fixed linearly ordered Henkin structure $\mathcal{M}$. Such formulas are build as those of Λ, but from more general atomic formulas of the form $\mathsf{A}\,\mathsf{a}$, $\mathsf{a} < \mathsf{b}$ and $\mathsf{a} = \mathsf{b}$ with $\mathsf{a}, \mathsf{b} \in \mathcal{V}^\iota \cup \mathcal{M}^\iota$ and $\mathsf{A} \in \mathcal{V}^o \cup \mathcal{M}^o$. The satisfaction relation for formulas with parameters is defined as for Λ.

2) Weak Monadic Second-Order Logic: Weak MSO (WMSO) is MSO for Henkin structures $\mathcal{M}$ such that $\mathcal{M}^o$ is the set of all finite $A \subseteq \mathcal{M}^\iota$. In particular, WMSO for ω-words is MSO for the weak standard model WN. For instance, the formula $\exists X\,\text{Ub}[X]$ is false in WN.

3) Monadic Theories: The *monadic theory* of a *p-q-* expanded linearly ordered Henkin structure $(\mathcal{M}, \overline{a}, \overline{A})$ is the set of all $\phi \in \Lambda^{p,q}$ satisfied in $(\mathcal{M}, \overline{a}, \overline{A})$.

C. Deduction and Completeness

We recall here some basic facts on deduction and model theory for Henkin structures, namely Henkin completeness and Löwenheim-Skolem Theorem. Both results require non full Henkin structures (see e.g. [16]).

1) Deduction: We consider a complete Natural Deduction system for the language Λ, seen as two-sorted first-order logic with one sort for individuals and one sort for predicates.

The *deduction relation* is written $\Gamma \vdash \phi$, where Γ is a (possibly empty) finite unordered list of formulas, and ϕ is a formula. It is inductively defined by the rules of Table I.

2) Henkin Completeness: Usual Henkin completeness holds for deduction w.r.t. validity in all Henkin structures (see e.g. [16]). In our setting, we have to axiomatize equality and linear orders. The equality axioms are *reflexivity* $\forall x(x = x)$, and the universal closure of the following *Leibniz's axiom scheme*:

$$\text{for all } \phi \in \Lambda, \qquad \forall xy(x = y \ \rightarrow \ \phi[x/z] \ \rightarrow \ \phi[y/z])$$

Let LO consists of the equality axioms together with the universal closures of the formulas corresponding to the linear order axioms of Section II-A1.

Theorem II.1 (Henkin Completeness) *Let Δ be a set of sentences and ϕ be a sentence. Assume that for all linearly ordered 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 $\mathsf{LO}, \Gamma \vdash \phi$.*

TABLE I
DEDUCTION RULES

Rules for Propositional Logic			**Rules for First-Order Logic** (where $\mathcal{X}, \mathcal{Y} \in \mathcal{V}^\iota$ or $\mathcal{X}, \mathcal{Y} \in \mathcal{V}^o$)
		$\dfrac{\Gamma \vdash \phi \qquad \Gamma \vdash \neg\phi}{\Gamma \vdash \psi}$	$\dfrac{\Gamma \vdash \phi[\mathcal{Y}/\mathcal{X}]}{\Gamma \vdash \exists\mathcal{X}\phi}$
$\dfrac{}{\Gamma \vdash \phi \vee \neg\phi}$	$\dfrac{}{\Gamma, \phi \vdash \phi}$		
$\dfrac{\Gamma \vdash \phi}{\Gamma \vdash \phi \vee \psi}$	$\dfrac{\Gamma \vdash \psi}{\Gamma \vdash \phi \vee \psi}$	$\dfrac{\Gamma \vdash \phi \vee \psi \quad \Gamma, \phi \vdash \varphi \quad \Gamma, \psi \vdash \varphi}{\Gamma \vdash \varphi}$	$\dfrac{\Gamma \vdash \exists\mathcal{X}\phi \quad \Gamma, \phi \vdash \psi}{\Gamma \vdash \psi}$ ($\mathcal{X}$ not free in Γ, ψ)

3) Löwenheim-Skolem Theorem: We say that an Henkin structure $\mathcal{M}$ is *countable* if both $\mathcal{M}^\iota$ and $\mathcal{M}^o$ are countable sets. For instance, the weak standard model WN is countable while the full standard model N is not. Countable structures are useful for forcing, and also because they have cofinality at most ω (see Section III-C).

Löwenheim-Skolem Theorem states that we may always assume that structures are countable (see e.g. [16]).

Theorem II.2 (Löwenheim-Skolem) *Let $\mathcal{M}$ be a linearly ordered Henkin structure. There is a countable linearly ordered Henkin structure which has the same monadic theory as $\mathcal{M}$.*

Note the following corollary to Henkin completeness:

Corollary II.3 *Let Δ be a set of sentences and ϕ be a sentence. Assume that for all countable linearly ordered 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 $\mathsf{LO}, \Gamma \vdash \phi$.*

D. Restrictions and Relativizations

We now discuss basic properties of restrictions of structures and relativization of formulas.

1) Restrictions of Structures with Parameters: The *restriction* of a structure $\mathcal{M}$ to some *non-empty* predicate $A \in \mathcal{M}^o$ is the structure $\mathcal{M}{\upharpoonright}A$ 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 $<_{\mathcal{M}{\upharpoonright}A}$ is the restriction of $<_\mathcal{M}$ to A: $<_{\mathcal{M}{\upharpoonright}A} := \; <_\mathcal{M} \cap \, (A \times A)$. It is convenient to write the individual and predicate domains of $\mathcal{M}{\upharpoonright}A$ respectively as $\mathcal{M}^\iota{\upharpoonright}A$ and $\mathcal{M}^o{\upharpoonright}A$.

Consider an expanded structure $(\mathcal{M}, \overline{a}, \overline{A})$ with $\overline{A} = A_1 \ldots A_q$. Let $A \in \mathcal{M}^o$ be non-empty and such that $\overline{a} \in A$. We define the *restriction of $(\mathcal{M}, \overline{a}, \overline{A})$ to A* to be the structure:

$$(\mathcal{M}, \overline{a}, \overline{A}){\upharpoonright}A \quad := \quad (\mathcal{M}{\upharpoonright}A, \overline{a}, (A_1 \cap A) \ldots (A_q \cap A))$$

2) 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 $\phi{\upharpoonright}\varphi[y]$, is defined by induction on ϕ as follows:

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

3) The Transfer Property: We now check that restriction and relativization are equivalent w.r.t. satisfaction.

Proposition II.4 (Transfer) *Let $(\mathcal{M}, \overline{a}, \overline{A})$ be a p-q-expanded structure. Let φ be a formula with parameters in $\mathcal{M}$ and whose free variable are disjoint from $\{x_1, \ldots, x_p, X_1, \ldots, X_q\}$. Given $x_0 \notin \{x_1, \ldots, x_p\}$, let $A \in \mathcal{M}^o$ be non-empty and such that $\overline{a} \in A$ and $(\mathcal{M}, \overline{a}, \overline{A}) \models \forall x \, (Ax \longleftrightarrow \varphi[x/x_0])$. Then, for all $\phi \in \Lambda^{p,q}$ we have*

$$(\mathcal{M}, \overline{a}, \overline{A}){\upharpoonright}A \models \phi \qquad iff \qquad (\mathcal{M}, \overline{a}, \overline{A}) \models \phi{\upharpoonright}\varphi[x_0]$$

4) Segments: A *segment* of a linearly ordered Henkin structure $\mathcal{M}$ is a predicate of one of the two following forms:

$$
\begin{array}{rcll}
[-, b) & := & \{c \in \mathcal{M}^\iota \mid c <_\mathcal{M} b\} & \\
[a, b) & := & \{c \in \mathcal{M}^\iota \mid a \leq_\mathcal{M} c <_\mathcal{M} b\} & (a <_\mathcal{M} b)
\end{array}
$$

Let ${\upharpoonright}[-, y)$ be the formula $\varphi[x_0, y] := (x_0 < y)$. Then, by the Transfer property II.4, if $\overline{a} <_\mathcal{M} b$ we have

$$(\mathcal{M}, \overline{a}, \overline{A}){\upharpoonright}[-, b) \models \phi \qquad iff \qquad (\mathcal{M}, \overline{a}, \overline{A}) \models \phi{\upharpoonright}\varphi[x_0, b]$$

The same can be done for segments of the form ${\upharpoonright}[a, b)$. From now on, we write $\phi{\upharpoonright}[-, y)$ for $\phi{\upharpoonright}\varphi[x_0, y]$, and similarly for $\phi{\upharpoonright}[x, y)$.

E. Linear Orders with Bounded Comprehension

We now present the axiomatics BC-MSO and BC-MSO$^\omega$ used in this paper.

The axioms of BC-MSO are LO together with *unboundedness* $\forall x \exists y (x < y)$ and the universal closure of the following *bounded comprehension scheme*: for all formula ϕ with no free occurrence of y, X:

$$\forall y \, \exists X \, \forall x \, (Xx \longleftrightarrow [x < y \; \wedge \; \phi{\upharpoonright}[-, y)])$$

The axioms of BC-MSO$^\omega$ are BC-MSO augmented with the *predecessor axiom*

$$\forall x(\exists y(y < x) \; \rightarrow \; \exists y[y < x \; \wedge \; \neg\exists z(y < z < x)])$$

and the universal closure of the following *induction axiom scheme*: for all formula ϕ with no free occurrence of y,

$$\forall x(\forall y(y < x \rightarrow \phi[y/x]) \rightarrow \phi) \rightarrow \forall x\phi$$

Note that both N and WN are models of BC-MSO$^\omega$. BC-MSO$^\omega$ will be the target of our translation (see Corollary VII.4), while BC-MSO will be the axiomatics used in

most of the paper. In our context, the distinctive property of BC-MSO$^\omega$ w.r.t. BC-MSO is Doets' Lemma III.3. Also, BC-MSO$^\omega$ with full comprehension completely axiomatizes the monadic theory of N (see [12], [10]).

III. Finite Fusion & Ramseyan Factorizations

In this section, we recall some tools on the composition of monadic theories of segments. We refer to [5], [11], [14]. The main points are the Finite Fusion Lemma III.2, which says that monadic theories of segments define additive colorings, and the existence of Ramseyan factorizations for such colorings (Theorem III.5). We rely on these results for our forcing conditions to be described in Section VI.

A. Finiteness

The starting point is to classify formulas according to their *quantifier depth*, *i.e.* their maximum nesting of quantifiers. The quantifier depth $(\mathrm{qd}(\phi))$ of an atomic formula ϕ is 0, and by induction $\mathrm{qd}(\neg\phi) := \mathrm{qd}(\phi)$, $\mathrm{qd}(\phi \vee \psi) := \max(\mathrm{qd}(\phi), \mathrm{qd}(\psi))$, and $\mathrm{qd}(\exists x\phi) := \mathrm{qd}(\exists X\phi) := \mathrm{qd}(\phi)+1$. We let $\Lambda_n^{p,q}$ be the set of formulas $\phi \in \Lambda^{p,q}$ with $\mathrm{qd}(\phi) \leq n$.

The main interest of this classification is that $\Lambda_n^{p,q}$ is finite modulo logical equivalence. Recall that $\phi, \psi \in \Lambda_n^{p,q}$ are *logically equivalent* if $\vdash \forall \overline{x}\overline{X}(\phi \longleftrightarrow \psi)$ is derivable with the rules of Table I.

We assume given a representative ϕ for each equivalence class of $\Lambda_n^{p,q}$, and write $\mathcal{LA}_n^{p,q}$ for the set of such ϕ. Hence $\mathcal{LA}_n^{p,q}$ represents $\Lambda_n^{p,q}$ quotiented by logical equivalence.

Lemma III.1 (Finiteness) $\mathcal{LA}_n^{p,q}$ *is finite.*

A *characteristic formula of* $\mathcal{LA}_n^{p,q}$ is a boolean combination of *all* the $\phi \in \mathcal{LA}_n^{p,q}$. Given $p, q, n \in \mathbb{N}$ and two p-q-expanded structures $(\mathcal{M}, \overline{a}, \overline{A})$ and $(\mathcal{N}, \overline{b}, \overline{B})$, we let

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

if $(\mathcal{M}, \overline{a}, \overline{A})$ and $(\mathcal{N}, \overline{b}, \overline{B})$ satisfy the same $\phi \in \mathcal{LA}_n^{p,q}$. It follows from Lemma III.1 that each $(\mathcal{M}, \overline{a}, \overline{A})$ has an n-characteristic, *i.e.* a characteristic formula θ such that $(\mathcal{N}, \overline{b}, \overline{B}) \models \theta$ iff $(\mathcal{M}, \overline{a}, \overline{A}) \equiv_n (\mathcal{N}, \overline{b}, \overline{B})$.

B. Finite Fusion of Segments

The Finiteness Lemma III.1 implies the fundamental fact that monadic theories of segments define additive colorings.

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

Lemma III.2 (Finite Fusion) *Let* $(\mathcal{M}, \overline{a}\overline{a}', \overline{A})$ *and* $(\mathcal{N}, \overline{b}\overline{b}', \overline{B})$ *be two expanded models of* BC-MSO. *Let* $n \in \mathbb{N}$ *and* $t_0 <_{\mathcal{M}^{-\infty}} t_1 <_{\mathcal{M}^{-\infty}} t_3$ *and* $u_0 <_{\mathcal{N}^{-\infty}} u_1 <_{\mathcal{N}^{-\infty}} u_3$. *Assume that* $\overline{a} \in \mathcal{M}^\iota{\restriction}[t_0, t_1)$ *and* $\overline{a}' \in \mathcal{M}^\iota{\restriction}[t_1, t_2)$, $\overline{b} \in \mathcal{N}^\iota{\restriction}[u_0, u_1)$ *and* $\overline{b}' \in \mathcal{N}^\iota{\restriction}[u_1, u_2)$.

$$
\begin{array}{llll}
\textit{If} & (\mathcal{M}, \overline{a}, \overline{A}){\restriction}[t_0, t_1) & \equiv_n & (\mathcal{N}, \overline{b}, \overline{B}){\restriction}[u_0, u_1) \\
\textit{and} & (\mathcal{M}, \overline{a}', \overline{A}){\restriction}[t_1, t_2) & \equiv_n & (\mathcal{N}, \overline{b}', \overline{B}){\restriction}[u_1, u_2) \\
\textit{then} & (\mathcal{M}, \overline{a}\overline{a}', \overline{A}){\restriction}[t_0, t_2) & \equiv_n & (\mathcal{N}, \overline{b}\overline{b}', \overline{B}){\restriction}[u_0, u_2).
\end{array}
$$

Consequence for Models of BC-MSO$^\omega$*:* The reduction of MSO to BC-MSO$^\omega$ is possible thanks to the simple but crucial observation that bounded segments of models of BC-MSO$^\omega$ are $\equiv_n$-equivalent to finite linear orders. To our knowledge, this is due to K. Doets [4] for the Π_1^1-case (first-order logic with universal prenex quantification on predicates).

In our context, a *finite linear order* is a structure of the form $\mathsf{N}{\restriction}[m_0, m_1)$ with $m_0 < m_1 \in \mathbb{N}$.

Lemma III.3 (Doets' Lemma) *Let* $\mathcal{M}$ *be a model of* BC-MSO$^\omega$ *and let* $n \in \mathbb{N}$. *For all* $a <_{\mathcal{M}} b$, *there is a finite linear order* $\mathcal{L}$ *such that* $\mathcal{M}{\restriction}[a, b) \equiv_n \mathcal{L}$.

C. Cofinality

Let $\mathcal{M}$ be a model of BC-MSO. A strictly increasing sequence $(\mathfrak{a}_k)_{k \in \mathbb{N}}$ of individuals $\mathfrak{a}_k \in \mathcal{M}^\iota$ is *cofinal* in $\mathcal{M}$ if for all $a \in \mathcal{M}^\iota$ there is $k \in \mathbb{N}$ such that $a <_{\mathcal{M}} \mathfrak{a}_k$.

The following is well-known:

Lemma III.4 *If* $(L, <_L)$ *is an unbounded countable linear order, then there is a cofinal sequence* $(\mathfrak{a}_k)_{k \in \mathbb{N}} \in L$.

We say that $(\mathfrak{b}_k)_{k \in \mathbb{N}}$ is a *subsequence* of $(\mathfrak{a}_k)_{k \in \mathbb{N}}$ if there is a strictly increasing function $f : \mathbb{N} \to \mathbb{N}$ (*i.e.* $n < m$ implies $f(n) < f(m)$) such that $\mathfrak{b}_k = \mathfrak{a}_{f(k)}$ for all $k \in \mathbb{N}$. Note that $(\mathfrak{b}_k)_{k \in \mathbb{N}}$ is cofinal whenever $(\mathfrak{a}_k)_{k \in \mathbb{N}}$ is. We shall thus assume that the first element of a sequence cofinal in $\mathcal{M}$ is not the least element of $\mathcal{M}$ (if it exists).

D. Ramseyan Factorizations

We now discuss Ramseyan factorizations for additive colorings defined by $\equiv_n$-equivalences classes of expanded segments. Ramseyan factorizations are a central tool for MSO on linear orders (see e.g. [11]), and in particular for the complementation of Büchi automata (see e.g. [9], [3], [13]).

Let $[\mathbb{N}]^2$ be the set of pairs $(i, j) \in \mathbb{N}^2$ such that $i < j$.

Consider an expanded model of BC-MSO $(\mathcal{M}, \overline{a}, \overline{A})$, and a sequence $(\mathfrak{a}_k)_{k \in \mathbb{N}}$ cofinal in $\mathcal{M}$ and such that $\overline{a} <_{\mathcal{M}} \mathfrak{a}_0$. Let $n \in \mathbb{N}$. We say that $(\mathfrak{a}_k)_{k \in \mathbb{N}}$ is a *Ramseyan factorization of* $(\mathcal{M}, \overline{a}, \overline{A})$ *modulo* $\equiv_n$ if for all $(i_0, i_1), (j_0, j_1) \in [\mathbb{N}]^2$,

$$(\mathcal{M}, , \overline{A}){\restriction}[\mathfrak{a}_{i_0}, \mathfrak{a}_{i_1}) \equiv_n (\mathcal{M}, , \overline{A}){\restriction}[\mathfrak{a}_{j_0}, \mathfrak{a}_{j_1})$$

Note that Ramseyan factorizations are preserved by taking subsequences.

Theorem III.5 (Ramseyan Factorizations) *Let* $(\mathcal{M}, \overline{a}, \overline{A})$ *be an expanded model of* BC-MSO, *and let* $(\mathfrak{a}_k)_{k \in \mathbb{N}} \in \mathcal{M}^\iota$ *be cofinal and such that* $\overline{a} <_{\mathcal{M}} \mathfrak{a}_0$.

For all $n \in \mathbb{N}$, *there is a subsequence* $(\mathfrak{b}_k)_{k \in \mathbb{N}}$ *of* $(\mathfrak{a}_k)_{k \in \mathbb{N}}$ *such that* $(\mathfrak{b}_k)_{k \in \mathbb{N}}$ *is a Ramseyan factorization of* $(\mathcal{M}, \overline{a}, \overline{A})$ *modulo* $\equiv_n$.

The proof is standard. We recall its main ingredient, the *merging relations* (see e.g. [3], [13], [11], [9]), which is essential to our forcing conditions to be described in Section VI.

Let $(\mathcal{M}, , \overline{A})$ be a q-expanded model of BC-MSO and let $(\mathfrak{a}_k)_{k \in \mathbb{N}} \in \mathcal{M}^\iota$ be cofinal. Given $a, b \in \mathcal{M}^\iota$, let $a \sim_n b$ if

there is $c >_{\mathcal{M}} a, b$ s.t. $(\mathcal{M}, , \overline{A}){\restriction}[a, c) \equiv_n (\mathcal{M}, , \overline{A}){\restriction}[b, c)$

This relation is reflexive, symmetric and transitive.

Moreover, the finiteness of $\mathcal{LA}_n^{0,q}$ (Lemma III.1) implies that $\sim_n$ has only finitely many equivalence classes.

Theorem III.5 easily follows from the following Merging Lemma, which says that Ramseyan factorizations can be extracted from the infinite equivalence classes of $\sim_n$.

Lemma III.6 (Merging) *Let $(\mathcal{M}, , \overline{A})$ be a q-expanded model of* BC-MSO *and let $(\mathfrak{a}_k)_{k \in \mathbb{N}} \in \mathcal{M}^\iota$ be cofinal.*

Let $k_0 \in \mathbb{N}$ and θ be a characteristic formula of $\mathcal{LA}_n^{0,q}$.

Assume that there are infinitely many $t \in \mathbb{N}$ such that $(\mathcal{M}, , \overline{A}){\upharpoonright}[\mathfrak{a}_{k_0}, \mathfrak{a}_t] \models \theta$ and $\mathfrak{a}_{k_0} \sim_n \mathfrak{a}_t$.

Then there is a subsequence $(\mathfrak{c}_k)_{k \in \mathbb{N}}$ of $(\mathfrak{a}_k)_{k \in \mathbb{N}}$ such that $\mathfrak{c}_0 = \mathfrak{a}_{k_0}$ and $(\mathcal{M}, , \overline{A}){\upharpoonright}[\mathfrak{c}_i, \mathfrak{c}_j] \models \theta$ for all $(i, j) \in [\mathbb{N}]^2$.

By combining Theorem III.5 with Lemma III.4, we get:

Corollary III.7 *Let $(\mathcal{M}, \overline{a}, \overline{A})$ be an expanded model of* BC-MSO *such that $\mathcal{M}^\iota$ is countable.*

For all $n \in \mathbb{N}$, there is a Ramseyan factorization of $(\mathcal{M}, \overline{a}, \overline{A})$ modulo $\equiv_n$.

IV. Filters

Our forcing interpretation is based on the following idea, borrowed from point-free topology (see e.g. [7]): Sequences indexed by a linear order can be approximated by suitable *filters*. Forcing requires quite special filters, called *generic filters*. However, working with usual filters allows smooth statements and manipulations. This provides the setting of Sections V to VII.

In this section, we discuss such an approach to points by filters. As expected, the filter structure over WN is isomorphic to N. But it also turns out that filters models built from models of BC-MSO$^\omega$ are MSO-equivalent to N.

A *filter* on a partial order $(\mathcal{C}, \leq_\mathcal{C})$ is a set $\mathcal{F} \subseteq \mathcal{C}$ such that:

- if $P \in \mathcal{F}$ and $P \leq_\mathcal{C} Q$ then $Q \in \mathcal{F}$,
- if $P, Q \in \mathcal{F}$, then there is $R \in \mathcal{F}$ such that $R \leq_\mathcal{C} P, Q$.

Let $\mathcal{M}$ be a model of BC-MSO. A *bounded sequence* of $\mathcal{M}$ is a pair

$$S \quad := \quad \langle A, a \rangle$$

where $A \in \mathcal{M}^o$, $a \in \mathcal{M}^\iota$ and A contains no $b \geq_\mathcal{M} a$. We let $\mathrm{dom}(S)$, the *domain* of S, be the set of all $b <_\mathcal{M} a$, and write $b \in S$ for $b \in A$. We let $\mathsf{S}(\mathcal{M})$ be the set of bounded sequences of $\mathcal{M}$ and write $S \supseteq Q$ if $\mathrm{dom}(Q) \subseteq \mathrm{dom}(S)$ and $S{\upharpoonright}\mathrm{dom}(Q) = Q$. Note that $\mathsf{S}(\mathsf{N}) = \mathsf{S}(\mathsf{WN})$.

A *point* of $\mathcal{M}$ is a filter $\mathcal{F}$ on $(\mathsf{S}(\mathcal{M}), \supseteq)$ which is *total*, *i.e.* for all $b \in \mathcal{M}^\iota$, there is $S \in \mathcal{F}$ such that $b \in \mathrm{dom}(S)$. We let $\mathrm{Pt}(\mathcal{M})$ be the set of points of $\mathcal{M}$.

Note the apparently reversed order in $(\mathsf{S}(\mathcal{M}), \supseteq)$. A bounded sequence $S \in \mathsf{S}(\mathsf{WN})$ generates an open set $[S]$ in the usual product topology for ω-words: $[S]$ is the set of ω-words of which S is a prefix. Then, we have $S \supseteq Q$ iff $[S] \subseteq [Q]$.

Each filter $\mathcal{F}$ on $(\mathsf{S}(\mathcal{M}), \supseteq)$ induces a set of individuals $\mathcal{S}_\mathcal{F}$, defined as the union of the elements of $\mathcal{F}$:

$$\mathcal{S}_\mathcal{F} \quad := \quad \bigcup\{A \mid \langle A, a \rangle \in \mathcal{F}\}$$

The converse holds in the standard model: each ω-word $A \subseteq \mathbb{N}$ is of the form $\mathcal{S}_{\mathcal{F}_A}$ for a unique $\mathcal{F}_A \in \mathrm{Pt}(\mathsf{WN})$ ($\mathcal{F}_A$ is the set of all $\langle A{\upharpoonright}[-, k), k \rangle$ for $k \in \mathbb{N}$).

From now on, we will always denote $\mathcal{S}_\mathcal{F}$ by $\mathcal{F}$. In particular, we write $b \in \mathcal{F}$ for $b \in \mathcal{S}_\mathcal{F}$ and denote expansions $(\mathcal{M}, \overline{a}, \overline{\mathcal{S}_\mathcal{F}})$ by $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$. Note that the second condition on filters ensures that if $b \in \mathcal{F}$, then $b \in S$ for all $S \in \mathcal{F}$ such that $b \in \mathrm{dom}(S)$.

The collection of sets induced by $\mathrm{Pt}(\mathcal{M})$ is also written $\mathrm{Pt}(\mathcal{M})$. With this notation, the points of a given model $\mathcal{M}$ of BC-MSO induce the Henkin structure

$$(\mathcal{M}^\iota, \mathrm{Pt}(\mathcal{M}), <_\mathcal{M})$$

called *filter structure over* $\mathcal{M}$, also denoted $\mathrm{Pt}(\mathcal{M})$. It follows from the remarks above that $\mathrm{Pt}(\mathsf{WN})$ is isomorphic to N.

We now show that $\mathrm{Pt}(\_)$ gives models of BC-MSO from models of BC-MSO. This follows from Lemma IV.1, which holds thanks to the totality condition.

Note that if $\mathcal{M}$ is a model of BC-MSO, then for all $S \in \mathsf{S}(\mathcal{M})$ there is a (unique) point $\mathcal{F} \in \mathrm{Pt}(\mathcal{M})$ *generated by S*, *i.e.* such that $S \in \mathcal{F}$ and $b \in \mathcal{F}$ iff $b \in S$.

Lemma IV.1 *Let $\mathcal{M}$ be a model of* BC-MSO. *Fix $a <_{\mathcal{M}^{-\infty}} b$. For all formula ϕ, if $\phi \in \Lambda^{p,q}$ then for all $\overline{a} \in \mathcal{M}^\iota{\upharpoonright}[a, b)$ of length p and all $\overline{\mathcal{F}} \in \mathrm{Pt}(\mathcal{M})$ of length q we have*

$$(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}){\upharpoonright}[a, b) \models \phi \qquad \text{iff} \qquad (\mathrm{Pt}(\mathcal{M}), \overline{a}, \overline{\mathcal{F}}){\upharpoonright}[a, b) \models \phi$$

In particular, $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$ and $(\mathrm{Pt}(\mathcal{M}), \overline{a}, \overline{\mathcal{F}})$ have the same Ramseyan factorizations. We also deduce that $\mathrm{Pt}(\mathcal{M})$ is a model of BC-MSO whenever $\mathcal{M}$ so is.

Corollary IV.2 *If $\mathcal{M}$ is a model of* BC-MSO *then* $\mathrm{Pt}(\mathcal{M})$ *is a model of* BC-MSO.

Remark IV.3 (Filter Models from BC-MSO$^\omega$) *It is also possible to show that the filter construction yields models MSO-equivalent to* N *from models of* BC-MSO$^\omega$. *This can be proved directly. For countable models, this is a consequence of our forcing transformation (see Corollary VII.5).*

V. Forcing

In this section, we present the (mostly basic) tools and results on forcing that we use for our transformation. We refer to [6] for usual forcing in set theory and to [1] for forcing in second-order arithmetic.

The forcing technique could be presented starting from Remark IV.3 above: We have a model, say $\mathcal{M} \models$ BC-MSO, which is extended to a new model, say $\mathrm{Pt}(\mathcal{M})$, by systematically adding points. In the case of $\mathcal{M} = \mathsf{WN}$ or even $\mathcal{M} \models$ BC-MSO$^\omega$, we obtain with $\mathrm{Pt}(\mathcal{M})$ a model of a given theory, say the MSO-theory of N.

Forcing performs similar extensions of models, say from $\mathcal{M}$ to $\mathcal{M}[\mathsf{G}]$, but with a formula translation from $\mathcal{M}[\mathsf{G}]$ to $\mathcal{M}$ which allows to describe the theory of the extension $\mathcal{M}[\mathsf{G}]$ within the starting (or *ground*) model $\mathcal{M}$. This formula translation allows to describe the theory of $\mathcal{M}[\mathsf{G}]$ without having to actually define the structure $\mathcal{M}[\mathsf{G}]$. Extensions $\mathcal{M}[\mathsf{G}]$ defined by forcing are thus called *generic extensions*.

The version of forcing we use is not the standard one, for reasons related to *naming* and *iterated forcing*, to be discussed in Remark V.6 and Section VII.

We will note in Example V.7 that usual Cohen's generic reals do not permit by themselves to perform a forcing reduction from MSO to WMSO. This motivates the more complex notion of forcing devised in Section VI.

A. Forcing Structures

The formula translation of forcing is performed using a partially ordered set in the ground model $\mathcal{M}$, the *forcing conditions*. We present here a general setting which will be instantiated in Sections VI and VII.

A *forcing structure* is a structure $\mathcal{M}$ of the form

$$(\mathcal{M}^\iota, \mathcal{M}^o, <_\mathcal{M}, \mathtt{C}(\mathcal{M}), \leq_\mathtt{c})$$

where $(\mathcal{M}^\iota, \mathcal{M}^o, <_\mathcal{M})$ is a model of BC-MSO, and $(\mathtt{C}(\mathcal{M}), \leq_\mathtt{c})$ is a partially ordered set of *conditions*. We say that $\mathcal{M}$ is *countable* if $\mathcal{M}^\iota$, $\mathcal{M}^o$ and $\mathtt{C}(\mathcal{M})$ are countable sets. An *expanded forcing structure* is of the form $(\mathcal{M}, \overline{a}, \overline{A})$, where $\overline{a} \in \mathcal{M}^\iota$ and $\overline{A} \subseteq \mathcal{M}^\iota$.

We shall assume given an "erasing map", sending a condition $P \in \mathtt{C}(\mathcal{M})$ to a bounded sequence $P^\bullet \in \mathtt{S}(\mathcal{M})$ such that $P^\bullet \supseteq Q^\bullet$ whenever $P \leq_\mathtt{c} Q$.

For instance, if $\mathcal{M}$ is a model of BC-MSO, then $(\mathcal{M}^\iota, \mathcal{M}^o, <_\mathcal{M}, \mathtt{S}(\mathcal{M}), \supseteq)$ is a forcing structure, that we also write $\mathtt{S}(\mathcal{M})$. In this case, the erasing map $(\_)^\bullet$ is the identity.

B. The Forcing Language

We now define the *forcing language* $\mathtt{C}(\Lambda)$. It will be the target language of the formula translation from the extended model $\mathcal{M}[\mathsf{G}]$ to the ground model $\mathcal{M}$. The *forcing formulas* $\phi \in \mathtt{C}(\Lambda)$ will be interpreted in forcing structures.

We assume given a countable set $P, Q, R, \dots$ of *condition variables*[2]. The atomic formulas of $\mathtt{C}(\Lambda)$ are those of Λ augmented with condition comparison $P \leq_\mathtt{c} Q$ and membership Px. Formulas $\phi \in \mathtt{C}(\Lambda)$ are then formed with the same connectives and quantifiers as those of Λ, plus existential quantifications over conditions $\exists P \phi$.

Satisfiability for forcing formulas in forcing structures $\mathcal{M}$ is defined as for Λ, with condition variables ranging over $\mathtt{C}(\mathcal{M})$, and $\mathcal{M} \models Pa$ iff $a \in P^\bullet$ in the sense of Section IV. We also allow forcing formulas ϕ with parameters in some fixed forcing structure $\mathcal{M}$.

C. The Forcing Translation

As suggested above, we shall define the forcing translation from the extended language $\Lambda[\dot{\mathsf{G}}]$ of $\mathcal{M}[\mathsf{G}]$ to the forcing language, before discussing the actual definition of $\mathcal{M}[\mathsf{G}]$.

Extended formulas $\phi \in \Lambda[\dot{\mathsf{G}}]$ are defined as those of Λ, but by augmenting atomic formulas with $\dot{\mathsf{G}}x$, where $\dot{\mathsf{G}}$ is a constant. We also allow extended formulas with parameters in some fixed model of BC-MSO.

---

[2]Depending on the context, letters P, Q, R range over forcing conditions or forcing variables. This shall rise no confusion.

There are essentially two formulations of forcing considered in the literature, namely *weak* and *strong* forcing (see e.g. [1]). We favor here strong forcing, since it eases model-theoretic reasoning.

We now define the forcing translation. To each extended formula $\phi \in \Lambda[\dot{\mathsf{G}}]$, we associate by induction on ϕ a forcing formula $(P \Vdash \phi) \in \mathtt{C}(\Lambda)$, where P is a free condition variable:

$$
\begin{aligned}
P \Vdash (\dot{\mathsf{G}}x) &:= Px \\
P \Vdash (Xx) &:= Xx \\
P \Vdash (x = y) &:= x = y \\
P \Vdash (x < y) &:= x < y \\
P \Vdash \neg\phi &:= \forall Q(Q \leq_\mathtt{c} P \rightarrow \neg(Q \Vdash \phi)) \\
P \Vdash \phi \vee \psi &:= (P \Vdash \phi) \vee (P \Vdash \psi) \\
P \Vdash \exists x \phi &:= \exists x(P \Vdash \phi) \\
P \Vdash \exists X \phi &:= \exists X(P \Vdash \phi)
\end{aligned}
$$

Note that the predicate quantification of $(\exists X \phi) \in \Lambda[\dot{\mathsf{G}}]$ is interpreted so as to range only over the predicates of the ground model, excluding the constant $\dot{\mathsf{G}}$. This differs from usual forcing (see Remark V.6).

The following is an essential basic property of forcing. The proof is an easy induction on formulas.

Lemma V.1 (Monotonicity) *Let $(\mathcal{M}, \overline{a}, \overline{A})$ be an expanded forcing structure and let $\phi \in \Lambda[\dot{\mathsf{G}}]$. For all $P \in \mathtt{C}(\mathcal{M})$, if $(\mathcal{M}, \overline{a}, \overline{A}) \models P \Vdash \phi$ and $Q \leq_\mathtt{c} P$, then $(\mathcal{M}, \overline{a}, \overline{A}) \models Q \Vdash \phi$.*

D. Generic Filters

Generic extensions $\mathcal{M}[\mathsf{G}]$ are obtained from ground models $\mathcal{M}$ by extending them with special objects called *generic filters*. We need a few definitions.

Given a forcing structure $\mathcal{M}$, a set of conditions $\mathcal{S} \subseteq \mathtt{C}(\mathcal{M})$ is *dense* if for every $P \in \mathtt{C}(\mathcal{M})$ there is $Q \in \mathcal{S}$ such that $Q \leq_\mathtt{c} P$.

Example V.2

(i) *In forcing structures $\mathtt{S}(\mathcal{M})$, given $a \in \mathcal{M}^\iota$, the set of all S such that $a \in \mathrm{dom}(S)$ is dense.*

(ii) *Given an extended forcing structure $(\mathcal{M}, \overline{a}, \overline{A})$ and a formula $\phi \in \Lambda[\dot{\mathsf{G}}]$, by the Monotonicity Lemma V.1 the set of all $P \in \mathtt{C}(\mathcal{M})$ such that $(\mathcal{M}, \overline{a}, \overline{A}) \models P \Vdash (\phi \vee \neg\phi)$ is dense.*

Given an extended forcing structure $(\mathcal{M}, \overline{a}, \overline{A})$, a set of conditions $\mathcal{S}$ is *definable* if there is a forcing formula ϕ with parameters such that for all $P \in \mathtt{C}(\mathcal{M})$, $(\mathcal{M}, \overline{a}, \overline{A}) \models \phi[P]$ if and only if $P \in \mathcal{S}$. The sets of Example V.2 are definable.

Definition V.3 (Generic Filter) *Let $(\mathcal{M}, \overline{a}, \overline{A})$ be an expanded forcing structure. A $\mathtt{C}(\mathcal{M})$-generic filter over $(\mathcal{M}, \overline{a}, \overline{A})$ is a filter G on $(\mathtt{C}(\mathcal{M}), \leq_\mathtt{c})$ such that for every dense $\mathcal{S} \subseteq \mathtt{C}(\mathcal{M})$ definable in $(\mathcal{M}, \overline{a}, \overline{A})$, we have $\mathsf{G} \cap \mathcal{S} \neq \emptyset$.*

The erasing map $(\_)^\bullet$ can be extended to generic filters G on $(\mathtt{C}(\mathcal{M}), \leq_\mathtt{c})$: let $\mathsf{G}^\bullet \subseteq \mathtt{S}(\mathcal{M})$ be the set of all $S \in \mathtt{S}(\mathcal{M})$ such that $P^\bullet \supseteq S$ for some $P \in \mathsf{G}$. The set $\mathsf{G}^\bullet$ is a filter over $(\mathtt{S}(\mathcal{M}), \supseteq)$, but it may be not total.

In the forcing structure $S(\mathcal{M})$, we have $G^\bullet = G$. Moreover, using Example V.2.(i), we get that an $S(\mathcal{M})$-generic filter G is total, hence $G \in \mathrm{Pt}(\mathcal{M})$.

Countable forcing structures have enough generic filters. The proof is standard (see e.g. [6]).

Lemma V.4 (Existence of Generic Filters) *Let $(\mathcal{M}, \overline{a}, \overline{A})$ be an expanded countable forcing structure.*

For every $P \in C(\mathcal{M})$, there is a $C(\mathcal{M})$-generic filter G over $(\mathcal{M}, \overline{a}, \overline{A})$ such that $P \in G$.

Recall that Löwenheim-Skolem Theorem II.2 allows to restrict to countable structures. But a countable $\mathcal{M}$ has uncountably many expansions (in particular, $\mathrm{Pt}(\mathcal{M})$ is in general not countable). Hence, the crucial point in Definition V.3 w.r.t. Lemma V.4, is that generic filters meet every dense set definable using *fixed sets* $S \in \mathcal{P}(\mathcal{M}^\iota) \setminus \mathcal{M}^o$: If $(\mathcal{M}, \overline{a}, \overline{A})$ and $(\mathcal{M}, \overline{b}, \overline{B})$ are expansions such that $\overline{A}$ and $\overline{B}$ contain the same sets $S \in \mathcal{P}(\mathcal{M}^\iota) \setminus \mathcal{M}^o$, then a filter is generic over $(\mathcal{M}, \overline{a}, \overline{A})$ if and only if it is generic over $(\mathcal{M}, \overline{b}, \overline{B})$.

E. Satisfiability in Generic Extensions

We now define a notion of satisfiability for extended formulas and relate it to the forcing interpretation. This will finish our description of generic extensions.

Given a p-q-extended forcing structure $(\mathcal{M}, \overline{a}, \overline{A})$, a $C(\mathcal{M})$-generic filter G, and an extended formula $\phi \in \Lambda[\dot{G}]^{p,q}$, let

$$(\mathcal{M}[G], \overline{a}, \overline{A}) \models \phi$$

if $(\mathcal{M}, \overline{a}, \overline{A}G^\bullet) \models \phi[X_{q+1}/\dot{G}]$ in the sense of Section IV.

Lemma V.5 (Truth Lemma) *Let $(\mathcal{M}, \overline{a}, \overline{A})$ be a p-q-extended forcing structure and G be a $C(\mathcal{M})$-generic filter over $(\mathcal{M}, \overline{a}, \overline{A})$.*

For all extended formula $\phi \in \Lambda[\dot{G}]^{p,q}$, $(\mathcal{M}[G], \overline{a}, \overline{A}) \models \phi$ if and only if there is a $P \in G$ s.t. $(\mathcal{M}, \overline{a}, \overline{A}) \models P \Vdash \phi$.

Proof sketch: By induction on ϕ. The case of $\neg\psi$ follows from Example V.2.(ii). For $\exists X\psi$ (resp. $\exists x\psi$), note that given $A \in \mathcal{M}^o$ (resp. $a \in \mathcal{M}^\iota$), G is generic over $(\mathcal{M}, \overline{a}, \overline{A}A)$ (resp. over $(\mathcal{M}, \overline{a}a, \overline{A})$) iff it is generic over $(\mathcal{M}, \overline{a}, \overline{A})$. ∎

Remark V.6 (Naming) *When pairing or primitive recursive codings are available (it is the case of the usual settings cited above) it is possible to uniformly* name *in the ground model $\mathcal{M}$ the objects of the generic extension $\mathcal{M}[G]$. Second-order quantifications in $\Lambda[\dot{G}]$ are then interpreted by quantifications in $C(\Lambda)$ ranging over the* names *in $\mathcal{M}$ of the sets of $\mathcal{M}[G]$.*

In our setting, such naming is not available. We could have allowed predicate quantifications in $\Lambda[\dot{G}]$ to range over predicates of the ground model and *the generic G, by taking e.g.*

$$P \Vdash \exists X\phi \quad := \quad \exists X(P \Vdash \phi) \ \vee \ (P \Vdash \phi[\dot{G}/X])$$

But this would have lead to heavier definitions, statements and proofs. This choice is the reason for our notation $P \Vdash \phi$, which differs from the usual notation $P \Vdash \phi$.

Example V.7 (Forcing with $S(WN)$ for MSO) *The generics for the structure $S(WN)$ (called* Cohen reals*, see e.g. [6]) are*

not adapted to a forcing reduction from N to WN. *Consider the formula $\mathrm{E}[X] := (\forall x \neg Xx)$ stating that X is empty. This formula is satisfied by the empty set $\emptyset \in \mathcal{P}(\mathbb{N})$, but there is no generic G such that $WN[G] \models \mathrm{E}[\dot{G}]$: Since the set of all non-empty $S \in S(WN)$ is dense and definable, for all generic G we have $n \in G^\bullet$ for some $n \in \mathbb{N}$.*

VI. DESCRIPTION OF THE FORCING CONDITIONS

In this section we describe our forcing conditions and discuss some of their properties. These conditions are then combined in the translation proposed in Section VII.

A. General Idea from Büchi Automata

We have seen in Example V.7 that it is not possible to reduce MSO to WMSO by using only the forcing structure $S(WN)$. Our forcing conditions are build from $S(\mathcal{M})$ (for $\mathcal{M}$ a model of BC-MSO) by adding informations and constraints to the conditions and to their ordering.

The main intuition comes from Büchi automata. Recall that Büchi automata are usual finite state non-deterministic automata running on infinite words. Their (infinite) accepting runs are those who reach a final state infinitely often.

Consider a Büchi automaton $\mathcal{A}$. Our forcing conditions are inspired from the following observation:

- Fix an input ω-word $\overline{A} \subseteq \mathbb{N}$. Let $\mathrm{PAcc}_\mathcal{A}(\overline{A})$ be the set of all finite prefixes of the accepting runs of $\mathcal{A}$ on $\overline{A}$. The set of finite runs of $\mathcal{A}$ on $\overline{A}$ ending in a final state is dense in $\mathrm{PAcc}_\mathcal{A}(\overline{A})$ (for the order $\supseteq$ of Section IV).

The main difficulty is to express "being a prefix of an accepting run" using forcing conditions. We achieve this by following ideas of [3] (see also [13]) on the approach to Ramseyan factorizations described in Section III-D. In particular, we rely on the merging relation to ensure that forcing conditions generate Ramseyan factorizations for a given coloring.

Note that in the above observation, the notion of condition depends both on the automaton $\mathcal{A}$ and on the input ω-word $\overline{A}$. Forcing conditions will be defined by formulas $(C_n^{p,q}, \leq_n^q)$ for each $n, p, q \in \mathbb{N}$, where p, q correspond to the dependence on inputs and n corresponds, *via* the Finiteness Lemma III.1, to the dependence on formulas, classified according to their quantifier depth.

B. Definition of the Forcing Conditions

We now proceed to the definition of the forcing conditions. The main idea is the following. Given an expanded model $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$ of BC-MSO with $\overline{\mathcal{F}} \in \mathrm{Pt}(\mathcal{M})$, a condition P will be a triplet

$$\langle A, a^0, a^1 \rangle \in \mathcal{M}^o \times \mathcal{M}^\iota \times \mathcal{M}^\iota \qquad \text{where} \quad a^0 <_\mathcal{M} a^1$$

and such that for any generic filter G containing P, there is a Ramseyan factorization of $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}G)$ for the theory modulo $\equiv_n$ of $(\mathcal{M}, , \overline{\mathcal{F}}A){\restriction}[a^0, a^1)$.

Forcing conditions are defined using the formulas $C_n^{p,q}$ and $\leq_n^q$ of Table II, where $p, q, n \in \mathbb{N}$. We assume that the individual variables $x, y_0, y_1, z, z_0, z_1, z_2$ are not among $x_1, \ldots, x_p$ and the predicate variables Y, Z not among $X_1, \ldots, X_q$.

We now comment some formulas of Table II.

The formulas $Y{\upharpoonright}[y_0, y_1] \equiv_n^{p,q} Z{\upharpoonright}[z_0, z_1]$ and ${\upharpoonright}[y_0, y_1] \equiv_n^{p,q}$ $Z{\upharpoonright}[z_0, z_1]$ characterize $\equiv_n$-equivalence of bounded segments. For instance, given a p-q-expanded model $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$ and $\langle A, a^0, a^1 \rangle, \langle B, b^0, b^1 \rangle$ such that $\overline{a} <_{\mathcal{M}} a^0, b^0$, we have

$$(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}) \models A{\upharpoonright}[a^0, a^1] \equiv_n^{p,q} B{\upharpoonright}[b^0, b^1]$$

if and only if

$$(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}A){\upharpoonright}[a^0, a^1] \equiv_n (\mathcal{M}, \overline{a}, \overline{\mathcal{F}}B){\upharpoonright}[b^0, b^1]$$

The formula $\mathrm{Id}_n^q[A, a^0, a^1]$ holds in $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$ when the theory of $(\mathcal{M}, , \overline{\mathcal{F}}A){\upharpoonright}[a^0, a^1]$ in $\mathcal{LA}_n^{0,q}$ is idempotent.

In the formula $\mathsf{C}_n^{p,q}$, the quantifier $\exists^\omega z \phi$ stands for $\forall x \exists z (x < z \wedge \phi)$ (x not free in ϕ), and $y_1 \sim_{n+1}^q z$ stands for the merging relation of Section III-D, represented by the formula $\exists x (y_1, z < x \ \wedge \ {\upharpoonright}[y_1, x] \equiv_{n+1}^{0,q} {\upharpoonright}[z, x])$.

C. The Forcing Structures $\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}(\mathcal{M})$

We now discuss the main properties of the forcing structures induced by $(\mathsf{C}_n^{p,q}, \leq_n^q)$, and in particular how they instantiate the framework of Section V.

For the remaining of this section, fix $p, q, n \in \mathbb{N}$ and let $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$ be a p-q-expanded countable model of $\mathsf{BC\text{-}MSO}$ with $\overline{\mathcal{F}} \in \mathrm{Pt}(\mathcal{M})$.

Given $P = \langle A, a^0, a^1 \rangle$ and $Q = \langle B, b^0, b^1 \rangle$, we let

$$P \in \mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}} \quad \text{iff} \quad (\mathcal{M}, \overline{a}, \overline{\mathcal{F}}) \models \mathsf{C}_n^{p,q}[A, a^0, a^1]$$

$$Q \leq_n^{\overline{\mathcal{F}}} P \quad \text{iff} \quad (\mathcal{M}, \overline{a}, \overline{\mathcal{F}}) \models \langle B, b^0, b^1 \rangle \leq_n^q \langle A, a^0, a^1 \rangle$$

$$Q <_n^{\overline{\mathcal{F}}} P \quad \text{iff} \quad (\mathcal{M}, \overline{a}, \overline{\mathcal{F}}) \models \langle B, b^0, b^1 \rangle <_n^q \langle A, a^0, a^1 \rangle$$

We thus obtain

$$\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}(\mathcal{M}) \quad := \quad (\mathcal{M}^\iota, \mathcal{M}^o, <_{\mathcal{M}}, \mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}, \leq_n^{\overline{\mathcal{F}}})$$

The erasing map is defined as $\langle A, a^0, a^1 \rangle^\bullet := \langle A, a^1 \rangle$. The formula Seq of $\mathsf{C}_n^{p,q}$ ensures that $\langle A, a^1 \rangle \in \mathsf{S}(\mathcal{M})$. We write $\mathrm{dom}(P)$ for $\mathrm{dom}(P^\bullet)$. Note that $P \leq_n^{\overline{\mathcal{F}}} Q$ implies $P^\bullet \supseteq Q^\bullet$ thanks to the first two parts of the formula $<_n^q$.

The relation $\leq_n^{\overline{\mathcal{F}}}$ is reflexive. We use the part Id_n^q of the formula $\mathsf{C}_n^{p,q}$ to ensure that it is transitive.

Lemma VI.1 (Transitivity) *Given* $P, Q, R \in \mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}$, *if* $P <_n^{\overline{\mathcal{F}}} Q <_n^{\overline{\mathcal{F}}} R$, *then* $P <_n^{\overline{\mathcal{F}}} R$.

It follows that $\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}(\mathcal{M})$ is a forcing structure in the sense of Section V-A.

The totality of the filters obtained from $\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}$-generics follows from the Merging Lemma III.6, whose premisses are ensured by the last part of the formula $\mathsf{C}_n^{p,q}$.

Lemma VI.2 (Totality) *For all* $P \in \mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}$ *and all* $a \in \mathcal{M}^\iota$, *there is a condition* $Q \leq_n^{\overline{\mathcal{F}}} P$ *such that* $a \in \mathrm{dom}(Q)$.

Let G be a $\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}$-generic filter over $(\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}(\mathcal{M}), \overline{a}, \overline{\mathcal{F}})$. Note that the property expressed by the above Totality Lemma VI.2 (for fixed $a \in \mathcal{M}^\iota$) is definable in $(\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}(\mathcal{M}), \overline{a}, \overline{\mathcal{F}})$. Hence, for all $a \in \mathcal{M}^\iota$, there is $P \in \mathsf{G}$ such that $a \in \mathrm{dom}(P)$.

It follows that the corresponding filter $\mathsf{G}^\bullet$ on $(\mathsf{S}(\mathcal{M}), \supseteq)$ is total, hence $\mathsf{G}^\bullet \in \mathrm{Pt}(\mathcal{M})$.

From now on, we drop the notation $(\_)^\bullet$ whenever possible. In particular, we write $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}\mathsf{G})$ for $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}\mathsf{G}^\bullet)$.

The following Approximation Lemma VI.3 says that points $\mathcal{F} \in \mathrm{Pt}(\mathcal{M})$ can approximated by $\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}$-generic filters modulo Ramseyan factorizations. This is crucial for the correctness of the translation (see Section VII).

Lemma VI.3 (Approximation) *Let* $(\mathfrak{a}_k)_{k \in \mathbb{N}}$ *be a Ramseyan factorization of* $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$ *modulo* $\equiv_{n+1}$.

Let $\mathcal{F} \in \mathrm{Pt}(\mathcal{M})$ *such that* $(\mathfrak{a}_k)_{k \in \mathbb{N}}$ *is also a Ramseyan factorization of* $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}\mathcal{F})$ *modulo* $\equiv_n$.

Then there is a $\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}$-*generic filter* G *and a cofinal sequence* $(\mathfrak{b}_k)_{k \in \mathbb{N}} \in \mathcal{M}^\iota$ *such that*

$$(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}\mathcal{F}){\upharpoonright}[-, \mathfrak{a}_0] \equiv_n (\mathcal{M}, \overline{a}, \overline{\mathcal{F}}\mathsf{G}){\upharpoonright}[-, \mathfrak{b}_0]$$

and for all $(i_0, i_1), (j_0, j_1) \in [\mathbb{N}]^2$,

$$(\mathcal{M}, , \overline{\mathcal{F}}\mathcal{F}){\upharpoonright}[\mathfrak{a}_{i_0}, \mathfrak{a}_{i_1}] \equiv_n (\mathcal{M}, , \overline{\mathcal{F}}\mathsf{G}){\upharpoonright}[\mathfrak{b}_{j_0}, \mathfrak{b}_{j_1}]$$

Lemma VI.3 relies on Lemma V.4 and the following: given a $\mathsf{C}_n^{\overline{a}, \overline{\mathcal{F}}}$-generic filter G, by the Totality Lemma VI.2 we can find a strictly decreasing sequence $\langle A_k, a_k^0, a_k^1 \rangle_{k \in \mathbb{N}} \in \mathsf{G}$ such that $(a_k^1)_{k \in \mathbb{N}}$ is cofinal. The relation $<_n^{\overline{\mathcal{F}}}$ then ensures that for all $(i, j) \in [\mathbb{N}]^2$ we have

$$(\mathcal{M}, , \overline{\mathcal{F}}\mathsf{G}){\upharpoonright}[a_i^1, a_j^1] \equiv_n (\mathcal{M}, , \overline{\mathcal{F}}A_0){\upharpoonright}[a_0^0, a_0^1]$$

VII. COMPOSED ITERATED FORCING

We now present our translation. It maps a formula $\phi \in \Lambda_n^{p,q}$ to a formula $|\phi|_n^{p,q} \in \Lambda^{p,q}$. Its main properties are gathered in Section VII-C.

Roughly speaking, the idea of the translation is to replace in $\phi \in \Lambda_n^{p,q}$ every occurrence of a predicate quantifier $\exists X \psi$ by a new instance of forcing

$$\exists P (\mathsf{C}_{n'}^{p',q'}[P] \wedge P \Vdash |\psi[X_{q'+1}/X]|_{n'}^{p',q'+1}[\dot{\mathsf{G}}/X_{q'+1}])$$

Note that p', q', n' may be different from p, q, n. In particular, the formula $\exists X \psi$ may contain free variables that are bound in ϕ. Hence, the translation of ϕ will involve infinitely many forcing structures $(\mathsf{C}_{n'}^{\overline{a}', \overline{\mathcal{F}}'}, \leq_{n'}^{\overline{\mathcal{F}}'})$ in the sense of Section V.

Of course, only a finite number of distinct formulas $(\mathsf{C}_{n'}^{p',q'}, \leq_{n'}^{q'})_{p',q',n' \in \mathbb{N}}$ will occur in the translation of ϕ. We can thus base the translation on a syntactic formulation of the forcing relations $(\Vdash_n^{p,q})_{p,q,n \in \mathbb{N}}$ induced by $(\mathsf{C}_n^{p,q}, \leq_n^q)_{p,q,n \in \mathbb{N}}$. This formulation is discussed in Section VII-A. We then define the translation in Section VII-B.

Our translation is actually a form of *iterated forcing*, with which the phenomena described above is common. When codings are available, naming allows to get rid of the induced difficulties (see e.g. [6], [1] and Remark V.6). However, in our weak setting naming is not available as such. This motivates our definition of $\Vdash_n^{p,q}$.

TABLE II
FORCING CONDITIONS

$$\text{Seq}[Y,y] \quad := \quad \forall x(Yx \rightarrow x < y) \qquad\qquad Y{\upharpoonright}[-,y) = Z{\upharpoonright}[-,y) \quad := \quad \forall x < y(Yx \longleftrightarrow Zx)$$

$$\text{Gt}^p[y_0,y_1] \quad := \quad (\textstyle\bigwedge_{1 \le i \le p} x_i < y_0) \;\wedge\; (y_0 < y_1)$$

$$Y{\upharpoonright}[y_0,y_1) \equiv_n^{p,q} Z{\upharpoonright}[z_0,z_1) \quad := \quad \textstyle\bigwedge_{\psi \in \mathcal{LA}_n^{p,q+1}} (\psi[Y/X_{q+1}]{\upharpoonright}[y_0,y_1) \longleftrightarrow \psi[Z/X_{q+1}]{\upharpoonright}[z_0,z_1))$$

$${\upharpoonright}[y_0,y_1) \equiv_n^{p,q} {\upharpoonright}[z_0,z_1) \quad := \quad \textstyle\bigwedge_{\psi \in \mathcal{LA}_n^{p,q}} (\psi{\upharpoonright}[y_0,y_1) \longleftrightarrow \psi{\upharpoonright}[z_0,z_1))$$

$$\text{Id}_n^q[Y,y_0,y_1] \quad := \quad \exists z_0 z_1 z_2 \exists Z(z_0 < z_1 < z_2 \;\wedge\; Y{\upharpoonright}[y_0,y_1) \equiv_n^{0,q} Z{\upharpoonright}[z_0,z_1) \equiv_n^{0,q} Z{\upharpoonright}[z_1,z_2) \equiv_n^{0,q} Z{\upharpoonright}[z_0,z_2))$$

$$\text{C}_n^{p,q}[Y,y_0,y_1] \quad := \quad \text{Gt}^p[y_0,y_1] \;\wedge\; \text{Seq}[Y,y_1] \;\wedge\; \text{Id}_n^q[Y,y_0,y_1] \;\wedge\; \exists^\omega z(y_1 \sim_{n+1}^q z \;\wedge\; [y_0,y_1) \equiv_{n+1}^{0,q} [y_1,z))$$

$$\langle Z,z_0,z_1 \rangle <_n^q \langle Y,y_0,y_1 \rangle \quad := \quad y_1 < z_1 \;\wedge\; Y{\upharpoonright}[-,y_1) = Z{\upharpoonright}[-,y_1) \;\wedge\; Z{\upharpoonright}[z_0,z_1) \equiv_n^{0,q} Y{\upharpoonright}[y_0,y_1) \;\wedge\; Z{\upharpoonright}[z_0,z_1) \equiv_n^{0,q} Z{\upharpoonright}[y_1,z_1)$$

$$\langle Z,z_0,z_1 \rangle \le_n^q \langle Y,y_0,y_1 \rangle \quad := \quad \langle Z,z_0,z_1 \rangle <_n^q \langle Y,y_0,y_1 \rangle \;\vee\; (z_0 = y_0 \;\wedge\; z_1 = y_1 \;\wedge\; Z{\upharpoonright}[-,z_1) = Y{\upharpoonright}[-,y_1))$$

A. The Syntactic Forcing Notion $[\!\!\vdash_n^{p,q}$ for $(\text{C}_n^{p,q}, \le_n^q)$

Fix $p,q,n \in \mathbb{N}$. We define the syntactic forcing notion $[\!\!\vdash_n^{p,q}$ for the conditions induced by the formulas $(\text{C}_n^{p,q}, \le_n^q)$.

To each extended formula $\phi \in \Lambda[\dot{\text{G}}]$, we associate by induction on ϕ a formula

$$(\langle Y,y_0,y_1 \rangle \,[\!\!\vdash_n^{p,q} \phi) \in \Lambda$$

where Y, y_0, y_1 are not free in ϕ, Y is not among $X_1, \ldots, X_q$ and y_0, y_1 are not among $x_1, \ldots, x_p$. Note that p, q, n are *not* related to the free variables nor to the quantifier depth of ϕ.

The clauses of $(\langle Y,y_0,y_1 \rangle \,[\!\!\vdash_n^{p,q} \phi)$ are the obvious adaptations of the clauses defining $[\!\!\vdash$ in Section V-C. In particular:

$$\langle Y,y_0,y_1 \rangle \,[\!\!\vdash_n^{p,q} (\dot{\text{G}}x) \quad := \quad (x < y_1) \;\wedge\; Yx$$

Moreover, the formula $\langle Y,y_0,y_1 \rangle \,[\!\!\vdash_n^{p,q} (\phi \vee \psi)$ is

$$(\langle Y,y_0,y_1 \rangle \,[\!\!\vdash_n^{p,q} \phi) \;\vee\; (\langle Y,y_0,y_1 \rangle \,[\!\!\vdash_n^{p,q} \psi)$$

Also, writing P for $\langle Y,y_0,y_1 \rangle$, the formula $P \,[\!\!\vdash_n^{p,q} \neg\phi$ is

$$\forall Z \forall z_0 z_1 [\text{C}_n^{p,q}[Z,z_0,z_1] \;\rightarrow\; \langle Z,z_0,z_1 \rangle \le_n^q P \;\rightarrow\; \neg(\langle Z,z_0,z_1 \rangle \,[\!\!\vdash_n^{p,q} \phi)]$$

with the same freshness conditions on Z, z_0, z_1 as on Y, y_0, y_1 above. Finally, the clauses for $\langle Y,y_0,y_1 \rangle \,[\!\!\vdash_n^{p,q} \exists x\phi$ and $\langle Y,y_0,y_1 \rangle \,[\!\!\vdash_n^{p,q} \exists X\phi$ must be extended with the proviso that x, X are not among $Y, y_0, y_1, x_1, \ldots, x_p, X_1, \ldots X_q$.

Of course, the instances of $[\!\!\vdash_n^{p,q}$ correspond to $[\!\!\vdash$ for suitable forcing structures. Let $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$ be a p-q-expanded model of BC-MSO$^\omega$ and let $P \in \text{C}_n^{\overline{a},\overline{\mathcal{F}}}$. Then, for all $\phi \in \Lambda^{p,q}$ with parameters in $\mathcal{M}$, we have

$$(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}) \models P \,[\!\!\vdash_n^{p,q} \phi \quad \text{iff} \quad (\text{C}_n^{\overline{a},\overline{\mathcal{F}}}(\mathcal{M}), \overline{a}, \overline{\mathcal{F}}) \models P \,[\!\!\vdash \phi$$

B. Definition of the Interpretation $|\_|_n^{p,q}$

The interpretation $|\phi|_n^{p,q}$ of a formula $\phi \in \Lambda_n^{p,q}$ is defined by induction on lexicographically ordered pairs (n, ϕ):

$$|\phi|_n^{p,q} \;:=\; \phi \qquad \text{if } \phi \text{ is atomic}$$
$$|\neg\phi|_n^{p,q} \;:=\; \neg|\phi|_n^{p,q}$$
$$|\phi \vee \psi|_n^{p,q} \;:=\; |\phi|_n^{p,q} \;\vee\; |\psi|_n^{p,q}$$
$$|\exists x\phi|_{n+1}^{p,q} \;:=\; \exists x_{p+1}|\phi[x_{p+1}/x]|_n^{p+1,q}$$
$$|\exists X\phi|_{n+1}^{p,q} \;:=$$
$$\exists P\left(\text{C}_n^{p,q}[P] \;\wedge\; P \,[\!\!\vdash_n^{p,q} |\phi[X_{q+1}/X]|_n^{p,q+1}[\dot{\text{G}}/X_{q+1}]\right)$$

Note that the two clauses for quantifiers update the parameters p, q, n of the translation, following the (maximum) quantifier depths and free variables of formulas.

Note also that the clause for predicate quantification locally involves the extended formula

$$|\phi[X_{q+1}/X]|_n^{p,q+1}[\dot{\text{G}}/X_{q+1}]$$

But $\dot{\text{G}}$ disappears with

$$P \,[\!\!\vdash_n^{p,q} |\phi[X_{q+1}/X]|_n^{p,q+1}[\dot{\text{G}}/X_{q+1}]$$

Finally, note that in the above formula, the condition P "sees" $|\phi[X_{q+1}/X]|_n^{p,q+1}[\dot{\text{G}}/X_{q+1}]$ as an arbitrary extended formula.

C. Correctness and Completeness of the Interpretation

The main results on our interpretation are obtained by combining Ramseyan factorizations (Theorem III.5) with the Infinite Fusion Lemma VII.6 stated in Section VII-D.

Theorem VII.1 *Let $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$ be a p-q-expanded countable model of BC-MSO with $\overline{\mathcal{F}} \in \text{Pt}(\mathcal{M})$. For all $\phi \in \Lambda_n^{p,q}$,*

$$(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}) \models |\phi|_n^{p,q} \qquad \text{iff} \qquad (\text{Pt}(\mathcal{M}), \overline{a}, \overline{\mathcal{F}}) \models \phi$$

Since $(\text{Pt}(\text{WN}), \overline{a}, \overline{A})$ is isomorphic to $(\mathbb{N}, \overline{a}, \overline{A})$, and since WN is a countable model of BC-MSO, we immediately get:

Corollary VII.2 (Reduction of MSO to WMSO) *Let $\overline{a} \in \mathbb{N}$ of length p and $\overline{A} \subseteq \mathbb{N}$ of length q. For all $\phi \in \Lambda_n^{p,q}$ we have*

$$(\mathsf{WN}, \overline{a}, \overline{A}) \models |\phi|_n^{p,q} \qquad iff \qquad (\mathsf{N}, \overline{a}, \overline{A}) \models \phi$$

Reduction of MSO to $\mathsf{BC\text{-}MSO}^\omega$: Let $\mathcal{M}$ be a model of $\mathsf{BC\text{-}MSO}^\omega$ and fix $n \in \mathbb{N}$. By combining Doets' Lemma III.3 with Ramseyan factorizations (Theorem III.5) we obtain cofinal sequences $(\mathfrak{a}_k)_{k \in \mathbb{N}} \in \mathcal{M}^\iota$ and $(\mathfrak{b}_k)_{k \in \mathbb{N}} \in \mathbb{N}$ such that

$$\mathcal{M}{\restriction}[-, \mathfrak{a}_0] \equiv_n \mathsf{N}{\restriction}[-, \mathfrak{b}_0]$$

and for all $(i_0, i_1), (j_0, j_1) \in [\mathbb{N}]^2$,

$$\mathcal{M}{\restriction}[\mathfrak{a}_{i_0}, \mathfrak{a}_{i_1}) \equiv_n \mathsf{N}{\restriction}[\mathfrak{b}_{j_0}, \mathfrak{b}_{j_1})$$

Since N is isomorphic to $\mathrm{Pt}(\mathsf{WN})$, using the Infinite Fusion Lemma VII.6 we get the following:

Theorem VII.3 *Let $\mathcal{M}$ be a countable model of $\mathsf{BC\text{-}MSO}^\omega$. For all sentence ϕ of quantifier depth n, we have*

$$\mathcal{M} \models |\phi|_n^{0,0} \qquad iff \qquad \mathsf{N} \models \phi$$

Thanks to Henkin completeness, Theorem VII.3 implies that our forcing transformation is indeed a reduction from the monadic theory of N to $\mathsf{BC\text{-}MSO}^\omega$:

Corollary VII.4 *For all $\phi \in \Lambda_n^{0,0}$, we have*

$$\mathsf{BC\text{-}MSO}^\omega \vdash |\phi|_n^{0,0} \qquad iff \qquad \mathsf{N} \models \phi$$

By combining Theorem VII.3 with Theorem VII.1, we also obtain the result mentioned in Remark IV.3 that filter structures build from models of $\mathsf{BC\text{-}MSO}^\omega$ have the same monadic theory as N. Note that this result could have been proved directly, without the forcing transformation, thereby relaxing the countability assumption.

Corollary VII.5 *If $\mathcal{M}$ is a countable model of $\mathsf{BC\text{-}MSO}^\omega$, then for all sentence ϕ we have $\mathrm{Pt}(\mathcal{M}) \models \phi$ iff $\mathsf{N} \models \phi$.*

D. Infinite Fusion

The following infinite version of the Finite Fusion Lemma III.2 is the key to the results stated in Section VII-C. It is shown using Theorem III.5 on Ramseyan factorizations and the Approximation Lemma VI.3 on generic filters.

Lemma VII.6 (Infinite Fusion) *Let $(\mathcal{M}, \overline{a}, \overline{\mathcal{F}})$ and $(\mathcal{N}, \overline{b}, \overline{\mathcal{H}})$ be p-q-expanded models of $\mathsf{BC\text{-}MSO}$, with $\mathcal{M}$ countable.*

Assume given cofinal sequences $(\mathfrak{a}_i)_{i \in \mathbb{N}} \in \mathcal{M}$, $(\mathfrak{b}_j)_{j \in \mathbb{N}} \in \mathcal{N}$ such that $\overline{a} <_\mathcal{M} \mathfrak{a}_0$ and $\overline{b} <_\mathcal{N} \mathfrak{b}_0$. Let $n \in \mathbb{N}$ such that

$$(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}){\restriction}[-, \mathfrak{a}_0] \equiv_n (\mathrm{Pt}(\mathcal{N}), \overline{b}, \overline{\mathcal{H}}){\restriction}[-, \mathfrak{b}_0]$$

and for all $(i_0, i_1), (j_0, j_1) \in [\mathbb{N}]^2$,

$$(\mathcal{M}, , \overline{\mathcal{F}}){\restriction}[\mathfrak{a}_{i_0}, \mathfrak{a}_{i_1}) \equiv_n (\mathrm{Pt}(\mathcal{N}), , \overline{\mathcal{H}}){\restriction}[\mathfrak{b}_{j_0}, \mathfrak{b}_{j_1})$$

Then, for all $\phi \in \Lambda_n^{p,q}$, we have

$$(\mathcal{M}, \overline{a}, \overline{\mathcal{F}}) \models |\phi|_n^{p,q} \quad \text{if and only if} \quad (\mathrm{Pt}(\mathcal{N}), \overline{b}, \overline{\mathcal{H}}) \models \phi$$

VIII. Conclusion

We presented a structural and purely syntactic reduction of MSO to WMSO based on forcing. Our proofs are model theoretic, but we hope to find proof transformations on the corresponding axiomatizations, thus allowing to apply the techniques of [8]. Further work will also involve precise comparisons with Shelah's composition method [11] and with (algebraic) determinization (e.g. [9]).

References

[1] J. Avigad, "Formalizing Forcing Arguments in Subsystems of Second-Order Arithmetic," *Annals of Pure and Applied Logic*, vol. 82, no. 2, pp. 165–191, 1996.

[2] J. R. Büchi, "On a Decision Methond in Restricted Second-Order Arithmetic," in *Logic, Methodology and Philosophy of Science*, E. N. et al., Ed. Stanford Univ. Press, 1962, pp. 1–11.

[3] ——, "The Monadic Second Order Theory of ω_1," in *Decidable Theories II : The Monadic Second Order Theory of All Countable Ordinals*, ser. LNM, J. R. Büchi and D. Siefkes, Eds. Springer, 1973, vol. 328, pp. 1–127.

[4] K. Doets, "Monadic Π_1^1-Theories of Π_1^1-Properties," *Notre Dame Journal of Formal Logic*, vol. 30, no. 2, pp. 224–240, 1989.

[5] H.-D. Ebbinghaus and J. Flum, *Finite Model Theory*, 2nd ed., ser. Springer Monographs in Mathematics. Springer, 1999.

[6] T. Jech, *Set Theory: The Third Millennium Edition, revised and expanded*, ser. Springer Monographs in Mathematics. Springer, 2006.

[7] P. Johnstone, *Stone Spaces*, ser. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1986.

[8] A. Miquel, "Forcing as a Program Transformation," in *Proceedings of LICS'11*. IEEE Computer Society, 2011, pp. 197–206.

[9] D. Perrin and J.-É. Pin, *Infinite Words: Automata, Semigroups, Logic and Games*, ser. Pure and Applied Mathematics. Elsevier, 2004.

[10] C. Riba, "A Model Theoretic Proof of Completeness of an Axiomatization of Monadic Second-Order Logic on Infinite Words," in *Proceedings of IFIP-TCS'12*, ser. LNCS, vol. 7604. Springer, 2012, pp. 310–324.

[11] S. Shelah, "The Monadic Theory of Order," *The Annals of Mathematics, Second Series*, vol. 102, no. 3, pp. 379–419, 1975.

[12] D. Siefkes, *Decidable Theories I : Büchi's Monadic Second Order Successor Arithmetic*, ser. LNM. Springer, 1970, vol. 120.

[13] W. Thomas, "Automata on Infinite Objects," in *Handbook of Theoretical Computer Science*, J. van Leeuwen, Ed. Elsevier Science Publishers, 1990, vol. B: Formal Models and Semantics, pp. 133–192.

[14] ——, "Ehrenfeucht Games, the Composition Method, and the Monadic Theory of Ordinal Words," in *Structures in Logic and Computer Science*, vol. 1261. Springer, 1997.

[15] ——, "Languages, Automata, and Logic," in *Handbook of Formal Languages*, G. Rozenberg and A. Salomaa, Eds. Springer, 1997, vol. III, pp. 389–455.

[16] D. van Dalen, *Logic and Structure*, 4th ed., ser. Universitext. Springer, 2004.