The pseudofinite monadic second order theory of finite words.

Deacon Linkhorn

February 17, 2022

Contents

1 Introduction. ................................................................. 1

2 Preliminaries. ............................................................... 2
  2.1 Extended Stone duality. .............................................. 2
  2.2 Monoids. ....................................................................... 3
  2.3 Uppercase/lowercase notation and comprehension schema. .... 4

3 Axiomatising the pseudofinite monadic second order theory of Σ-words. 5

4 The free profinite monoid generated by Σ. 10

References 14

1 Introduction.

In this paper we analyse the pseudofinite monadic second order theory of words over a fixed alphabet Σ. That is we consider the shared monadic second order theory of words \( w : \alpha_w \to \Sigma \) for which \( \alpha_w \) is finite. In particular we give an axiomatisation \( T_{M(\Sigma^*)} \) (definition 3.3) of the shared monadic second order theory of such words, working in a one-sorted first order framework. For each finite word \( w : \alpha_w \to \Sigma \) we work with a first order structure \( M(w) \) (definition 3.2) expanding the Boolean algebra of subsets of \( \alpha_w \). The analysis hinges on the fact that concatenation of finite words interacts nicely with monadic second order logic, in particular when working in the signature \( L^\otimes \) (definition 3.2). More precisely, we heavily exploit the fact that for each \( k \in \mathbb{N} \) equivalence of (monadic second order versions of) words with respect to \( L^\otimes \)-formulas of quantifier depth at most \( k \) is a congruence for concatenation (proposition 3.10). The central result of the paper is theorem 3.15 in which we prove that the theory \( T_{M(\Sigma^*)} \) which we introduced is indeed an axiomatisation for the pseudofinite monadic second order theory of Σ-words.

As an application, this property of concatenation is then exploited to present an alternative proof of a theorem (theorem 4.3) connecting recognisable languages and finitely generated free profinite monoids via extended Stone duality, due to Gehrke, Grigorieff, and Pin. Our proof makes no appeals to heavy machinery from duality theory.

Related literature: A similar approach, exploiting the well behaved interaction of the concatenation of words and logical properties of words, is used in [GS10] to explore the connection
between star-free languages and free pro-aperiodic monoids. The main difference between the two papers is that in [GS19] it is first order logic on finite words, not monadic second order logic, which aligns with star-free languages and free pro-aperiodic monoids.

In the author’s PhD thesis [Lin21] an axiomatisation is given for the monadic second order theory of finite words over a singleton alphabet, i.e. the monadic second order theory of finite linear orders. The connection between recognisable languages over this alphabet and the free profinite monoid on one generator, which arises as a special case of the theorem of Gehrke, Grigorieff, and Pin (theorem 4.3), was considered in the thesis. There is a fundamental difference between the thesis work and the work presented here. In the thesis theorem 4.3 was used as a foundation, and it was observed that the logical analysis of the monadic second order theory of finite linear orders offered a way of viewing the theorem. Whereas in this paper we show that the logical analysis not only offers a point of view, but allows for a streamlined proof of theorem 4.3.

Our logical notation, wherever it is not explicitly defined or referenced, is borrowed from [Hod93].

2 Preliminaries.

2.1 Extended Stone duality.

The following material is adapted from [Han83] which works more generally with spectral spaces and distributive lattices. We also restrict to the case of binary operations and ternary relations.

Definition 2.1. Let $L$ be a Boolean algebra. A binary operation $\cdot : L^2 \to L$ is called,

1. normal if for each $a \in L$, $a \cdot \bot = \bot = \bot \cdot a$,
2. additive if for each $a, b, c \in L$, $(a \lor b) \cdot c = (a \cdot c) \lor (b \cdot c)$ and $a \cdot (b \lor c) = (a \cdot b) \lor (a \cdot c)$.

We view Boolean algebras with normal additive operations as structures for the signature $L = \{\lor, \land, \bot, \top, c, \cdot\}$. We define BAlgO to be the category whose objects are Boolean algebras equipped with a normal additive binary operation, and whose morphisms are the usual homomorphisms for the signature $L$.

Definition 2.2. Let $X$ be a Boolean space. A relation $R \subseteq X^3$ is called compatible if for each pair $(A, B) \in \text{Clop}(X)^2$ the set,

$$R(A, B, -) := \{z \in X : \exists x \in A \exists y \in B \ R(x, y, z)\},$$

is clopen in $X$. In other words, $R \subseteq X^3$ is compatible if the the function $\mathcal{P}(X)^2 \to \mathcal{P}(X)$ defined by $(A, B) \mapsto R(A, B, -)$ restricts to $\text{Clop}(X)$. Given a ternary relation $R \subseteq X^3$ and $z \in X$ we write $R^{-1}(z)$ as shorthand for the set $\{(x, y) \in X^2 : R(x, y, z)\}$. We define BoolComp to be the category whose objects are Boolean spaces equipped with a compatible ternary relation, and whose morphisms are the usual homomorphisms for the signature $L$.

Theorem 2.3 (Extended Stone duality, [Han83] Theorem 1.14). The categories BAlgO and BoolComp are dual to one another.

Let $F$ be the functor BAlgO $\to$ BoolComp which for objects takes $(L, \cdot)$ to $(\text{PrimF}(L), R)$ where,

1. PrimF$(L)$ is the Boolean space with underlying set comprising prime filters of $L$ and with the topology having an open basis comprising sets $D(a) := \{p \in \text{PrimF}(L) : a \notin p\}$ where $a \in L$,
2. \( R := \{(p, q, r) \in \text{PrimF}(L)^3 : \forall a \in p, \forall b \in q(a \cdot b \in r)\} \),
and for morphisms takes \( f : L_1 \to L_2 \) in \( \text{BAlgO} \) to its preimage map \( f^{-1} : \text{PrimF}(L_2) \to \text{PrimF}(L_1) \).

Let \( G \) be the functor \( \text{BoolComp} \to \text{BAlgO} \) which for objects takes \((X, R)\) to \((\text{Clop}(X), \cdot)\) where,

1. \( \text{Clop}(X) \) is the Boolean algebra of clopen subsets of \( X \),
2. \( \cdot : \text{Clop}(X)^2 \to \text{Clop}(X) \) is the operation given by \( A \cdot B := \{z \in X : \exists x \in A, \exists y \in B \ R(x, y, z)\} \),
and for morphisms takes \( f : X_1 \to X_2 \) in \( \text{BoolComp} \) to its preimage map \( f^{-1} : \text{Clop}(X_2) \to \text{Clop}(X_1) \).

Then \( FG \) is naturally isomorphic to the identity functor on \( \text{BoolComp} \) and \( GF \) is naturally isomorphic to the identity functor on \( \text{BAlgO} \).

### 2.2 Monoids.

Our general setup for working with topological monoids and category-theoretic notions for the most part closely follows [Geh16] (in particular section 4.2), however our notation slightly differs in places.

**Notation 2.4.** Let \( X \) be a set and \( \cdot : X^2 \to X \) a binary operation. We write \( \cdot \) for the binary operation on \( \mathcal{P}(X) \) given by,

\[
A \cdot B := \{a \cdot b : a \in A, b \in B\},
\]

for each \( A, B \in \mathcal{P}(X) \). We call \( \cdot \) the complexification of \( \cdot \) (see [Gol89], where algebras of complexes are discussed, for an explanation of the terminology).

**Definition 2.5.** A congruence on a monoid \((M, \cdot)\) is an equivalence relation \( \rho \) on the underlying set \( M \) such that for any \( a, b, a', b' \in M \), \((a, a'), (b, b') \in \rho \) implies \((a \cdot b, a' \cdot b') \in \rho \). In other words, a congruence is an equivalence relation \( \rho \) for which we get a well-defined binary operation on equivalence classes ‘pointwise’ (i.e. taking \([a]_\rho \cdot [b]_\rho := [a \cdot b]_\rho\)). Note that for each congruence \( \rho \) on \( M \) the set of equivalence classes \( M/\rho \) forms a monoid under this ‘pointwise’ operation.

We write \( \text{Con}(M) \) for the collection of congruences on a monoid \( M \), and \( \text{Con}_e(M) \) for the subcollection of finite index congruences, i.e. those with finitely many equivalence classes. For \( \rho \in \text{Con}(M) \), the homomorphism \( \pi_\rho : M \to M/\rho \) given by \( a \mapsto [a]_\rho \) is the called the canonical quotient map of \( \rho \).

**Definition 2.6.** Let \( X \) be a set and take two equivalence relations \( \rho \) and \( \theta \) on \( X \). We say that \( \rho \) refines \( \theta \), or equivalently that \( \theta \) is a coarsening of \( \rho \), precisely when \( \rho \subseteq \theta \). In other words when \( x \sim_\rho x' \) implies \( x \sim_\theta x' \). If \( \rho \) and \( \theta \) are equivalence relations on a set \( X \), and \( \rho \) refines \( \theta \), then there is a unique function \( p_{\rho, \theta} : X/\rho \to X/\theta \) such that \( p_{\rho, \theta} \circ \pi_\rho = \pi_\theta \). The function \( p_{\rho, \theta} \) is given by \( p_{\rho, \theta}([x]_\rho) := [x]_\theta \), and is called the parent map.

**Remark 2.7.** For each monoid \( M \), viewing \( \text{Con}(M) \) as a partial order under \( \supseteq \), the subcollection \( \text{Con}_e(M) \) is a directed subset. In particular for any \( \rho, \theta \in \text{Con}_e(M) \) the set-theoretic intersection \( \rho \cap \theta \) is an element of \( \text{Con}_e(M) \) with \( \rho \supseteq \rho \cap \theta \) and \( \theta \supseteq \rho \cap \theta \).

**Definition 2.8.** The category \( \text{TopMon} \) of topological monoids has as objects triples \((M, \cdot, \tau)\) comprising a monoid \((M, \cdot)\) and a topology \( \tau \) on \( M \) such that \( \cdot : M^2 \to M \) is continuous (\( M^2 \) is given the product topology). Arrows in \( \text{TopMon} \) are functions \( f : M \to M' \) which are monoid homomorphisms that are moreover continuous with respect to the topologies.
Definition 2.9. A directed system of topological monoids is (for us) a contravariant functor \( F : I \to \text{TopMon} \) where \( I \) is a directed partial order viewed as a category.

Recall that a cone for (a contravariant) functor \( F : C \to D \) comprises an object \( d \in D \) and for each \( c \in C \) an arrow \( f_c : d \to F(c) \), such that for each arrow \( f : c \leftrightarrow c' \in C \) the equality \( F(f) \circ f_c = f_{c'} \) holds.

The inverse limit of a directed system of topological monoids \( F : I \to \text{TopMon} \) is a cone \((M, (\hat{\pi}_i)_{i \in I})\) for \( F \) (see \[Lei14\] Definition 5.1.19) satisfying the universal property that for any other cone \((N, (g_i)_{i \in I})\) for \( F \) there is a unique morphism \( h : N \to M \) such that for each \( i \in I \) we have \( g_i = \hat{\pi}_i \circ h \). The inverse limit of the directed system \( F \), which always exists in the category TopMon, is denoted by \( \lim_{\leftarrow} F \).

Lemma 2.10. Let \( F : I \to \text{TopMon} \) be a directed system of topological monoids. If \( C \subseteq I \) is a cofinal subset then the restriction \( F' \) of \( F \) to \( C \) is a directed system of topological monoids and the two topological monoids \( \lim_{\leftarrow} F \) and \( \lim_{\leftarrow} F' \) are isomorphic.

Proof. This is dealt with (in a more general setting and under slightly different notation and setup) in \[Mac98\], in particular Theorem 1 on page 217.

Definition 2.11. Let \( M \) be a monoid. Let \( F_M : (\text{Con}_\omega(M), \supseteq) \to \text{TopMon} \) be the directed system given by taking \( F_M(\rho) \) to be the finite monoid \( M/\rho \) endowed with the discrete topology, and taking the parent map \( p_{\rho, \theta} : M/\rho \rightarrow M/\theta \) (see definition 2.6) whenever \( \theta \supseteq \rho \). (Note this is a contravariant functor, as \( F_M \) sends an arrow \( \rho \leftrightarrow \theta \) to an arrow \( M/\rho \xrightarrow{p_{\rho, \theta}} M/\theta \).) We write \( \hat{M} \) for the inverse limit \( \lim_{\leftarrow} F_M \). The topological monoid \( \hat{M} \) is called the profinite completion of \( M \).

In the case where \( M \) is the free monoid \( \Sigma^* \) on the set \( \Sigma \), the topological monoid \( \hat{\Sigma}^* \) is called the free profinite monoid on \( \Sigma \). In general a topological monoid \((M, \cdot, \tau)\) is called profinite if there exists a directed system \( F : I \to \text{TopMon} \) such that for each \( i \in I \) the topological monoid \( F(i) \) is finite and carries the discrete topology, such that \((M, \cdot, \tau) = \lim_{\leftarrow} F\).

2.3 Uppercase/lowercase notation and comprehension schema.

We distinguish between upper and lower case variables as a notational shorthand for a certain relativisation of formulas (see \[Hod93\] Section 5.1 pp 202 for a general overview of the relativisation technique) in the setting of atomic Boolean algebras.

Definition 2.12. An atom of a partial order \((P, \leq)\) with a smallest element \( \bot \) is a minimal element of \( P \setminus \{\bot\} \). A Boolean algebra, viewed as a partial order, is called atomic if every element lies above an atom. Note that for any partial order \((P, \leq)\) the atoms form a \( 0 \)-definable subset of \( P \), and moreover this can be done uniformly. By uniformly we mean that there is a single \( \mathcal{L} = \{\leq\} \) formula \( \text{At}(x) \) which defines the set of atoms in each partial order \((P, \leq)\). Later, for the purpose of eliminating quantifiers, we will simply include \( \text{At} \) as a unary predicate symbol in the signatures we work with.

Remark 2.13. For a Boolean algebra \( \mathcal{B} \), being atomic is equivalent to the following holding,

\[
\forall x \forall y ( (\forall z (\text{At}(z) \rightarrow (z \leq x \leftrightarrow z \leq y))) \rightarrow x = y). \tag{f}
\]

This says, if two elements sit above the same atoms then they are equal.

Notation 2.14. To make formulae more easily readable when working over theories of expansions of atomic Boolean algebras we make use of some notational conventions. For atoms we reserve lower case letters \( x, y, z, \ldots \) for variables, and lower case letters \( a, b, c, \ldots \) for constants and parameters. For general elements we use upper case letters \( X, Y, Z, \ldots \) for variables and upper
case letters A, B, C, … for general elements and parameters. Formally, the use of a lower case variable is shorthand for relativisation of the formula to the 0-definable set of atoms. Moreover, we write X(x) as shorthand for x ⊆ X.

To illustrate with an example,

\[ \forall X \forall Y (\forall z (X(z) \leftrightarrow Y(z)) \rightarrow X = Y) \]

is the rewriting of using these notational shorthands.

**Definition 2.15.** Let \( \mathcal{L} \) be a signature containing a binary relation symbol \( \subseteq \).

For each \( \mathcal{L} \)-formula \( \phi(x, \vec{Y}) \), we define \( \text{Comp}_\phi \) to be the \( \mathcal{L} \)-sentence,

\[ \forall \vec{Y} \exists Z \forall x (Z(x) \leftrightarrow \phi(x, \vec{Y})) \]

(Recall that \( Z(x) \) is shorthand for \( x \subseteq Z \).)

If \( \mathcal{M} \) is an \( \mathcal{L} \)-structure in which \( \subseteq \) is the ordering of an atomic Boolean algebra, then \( \text{Comp}_\phi \) says that for each tuple of parameters \( \vec{Y} \) there is an element \( Z \) of the Boolean algebra which lies above precisely the atoms \( x \) in the set defined by the formula \( \phi \) with parameters \( \vec{Y} \).

The **comprehension schema** for the signature \( \mathcal{L} \) comprises the sentence \( \text{Comp}_\phi \) for each \( \mathcal{L} \)-formula \( \phi \).

### 3 Axiomatising the pseudofinite monadic second order theory of \( \Sigma \)-words.

In this paper we make use of a one-sorted first-order setup to capture monadic second order structures. Space does not permit a detailed discussion of the approach used or the alternatives available, for which the reader is referred to the authors PhD thesis [Lin21], in particular chapters 1 and 3.

**Definition 3.1.** Let \( \Sigma \) be a finite set. A \( \Sigma \)-word is a function \( w : \alpha_w \rightarrow \Sigma \) where \( \alpha_w \) is a linear order. We write \( \Sigma^* \) for the collection of finite \( \Sigma \)-words, including the unique empty \( \Sigma \)-word \( \varepsilon : \emptyset \rightarrow \Sigma \). For \( \sigma \in \Sigma \) we sometimes identify \( \sigma \) with the \( \Sigma \)-word \( \{ \star \} \rightarrow \Sigma, \star \mapsto \sigma \).

**Definition 3.2.** We define \( \mathcal{L}_{\Sigma}^{\omega} \) to be the signature \( \{ \subseteq, \triangleleft, \text{At}, \bot, (P_{\sigma})_{\sigma \in \Sigma} \} \) in which,

- \( \subseteq \) and \( \triangleleft \) are binary relation symbols,
- At is a unary relation symbol,
- \( \bot \) is a constant symbol,
- \( P_{\sigma} \) is a unary relation symbol for each \( \sigma \in \Sigma \).

For each \( w : \alpha_w \rightarrow \Sigma \) from \( \Sigma^* \) we define \( M(w) \) to be the \( \mathcal{L}_{\Sigma}^{\omega} \)-structure in which,

1. the universe is \( \mathcal{P}(\alpha_w) \),
2. \( \subseteq \) is the usual set-theoretic inclusion,
3. \( \triangleleft \) is \( \{(A, B) \in \mathcal{P}(\alpha)^2 : a < b \text{ for some } a \in A, b \in B\} \), where \( < \) is the ordering of \( \alpha_w \),
4. At is the collection of atoms of \( (\mathcal{P}(\alpha_w), \subseteq) \), i.e. the singleton subsets of \( \alpha \),
5. \( \bot \) is the empty set,
6. for each \( \sigma \in \Sigma \) the unary relation symbol \( P_{\sigma} \) is given the interpretation,

\[ \{ \{a\} : a \in w^{-1}(\sigma) \} \].
We call the $\mathcal{L}_{\Sigma}^\otimes$-structure $M(w)$ the monadic second order version of the word $w$. The pseudofinite monadic second order theory of $\Sigma$-words is defined to be the $\mathcal{L}_{\Sigma}^\otimes$-theory $\bigcap_{w \in \Sigma} \text{Th}(M(w))$. We write $\text{Th}(M(\Sigma^*))$ as a shorthand for this theory.

**Definition 3.3.** We define an $\mathcal{L}_{\Sigma}^\otimes$-theory $T_{M(\Sigma^*)}$ to be the set of sentences expressing,
1. atomic Boolean algebra under $\subseteq$,
2. $\bot$ is the bottom element of the Boolean algebra,
3. the atoms are linearly ordered by $\preceq$,
4. the linear ordering $\preceq$ on the atoms is discrete with endpoints (we denote the smallest and largest atoms by 0 and $0^*$ respectively),
5. each element not equal to $\bot$ lies above a smallest atom with respect to $\preceq$,
6. $\forall X,Y (X \preceq Y \iff \exists x \exists y (X(x) \land Y(y) \land x \preceq y))$,
7. the unary relations ($P_{\sigma})_{\sigma \in \Sigma}$ form a partition\(^1\) of the set of atoms (i.e. each atom is in the extension of precisely one of them),
8. the comprehension schema for $\mathcal{L}_{\Sigma}^\otimes$ (see definition 2.15).

For the remainder of this section, we prove that $T_{M(\Sigma^*)}$ axiomatises the pseudofinite monadic second order theory of linear order. Along the way machinery required in the following section is also set up.

First we establish some lemmas working with a base theory $T_{\text{base}}(\Sigma) \subseteq T_{M(\Sigma)}$.

**Definition 3.4.** We define an $\mathcal{L}_{\Sigma}^\otimes$-theory $T_{\text{base}}(\Sigma)$ to be the set of sentences expressing,
1. atomic Boolean algebra under $\subseteq$,
2. $\bot$ is the bottom element of the Boolean algebra,
3. atoms linearly ordered under $\preceq$,
4. $\forall X,Y (X \preceq Y \iff \exists x \exists y (X(x) \land Y(y) \land x \preceq y))$,
5. the unary relations ($P_{\sigma})_{\sigma \in \Sigma}$ form a partition of the set of atoms.

Note that $T_{\text{base}}(\Sigma)$ is a subset of $T_{M(\Sigma)}$.

**Definition 3.5.** Let $(\alpha, <_\alpha)$ and $(\beta, <_\beta)$ be linear orders. We define $\alpha + \beta$ to be the linear ordering on the disjoint union $\alpha \sqcup \beta$ given by $<_\alpha \sqcup <_\beta \cup (\alpha \times \beta)$. Given $\Sigma$-words $w : \alpha \to \Sigma$ and $v : \beta \to \Sigma$, we define their concatenation to be the $\Sigma$-word $w \cdot v : \alpha + \beta \to \Sigma$ given by,

$$(w \cdot v)(i) := \begin{cases} w(i) & \text{if } i \in \alpha, \\ v(i) & \text{if } i \in \beta. \end{cases}$$

Moreover we define $M(w) + M(v)$ to be the $\mathcal{L}_{\Sigma}^\otimes$-structure $M(w \cdot v)$.

**Definition 3.6.** Given $\mathcal{L}_{\Sigma}^\otimes$-structures $\mathcal{M}$ and $\mathcal{N}$, we define an $\mathcal{L}_{\Sigma}^\otimes$-structure $\mathcal{M} \otimes \mathcal{N}$ as follows,
1. the universe of the structure is $\mathcal{M} \times \mathcal{N}$,
2. $\subseteq$ is defined as it is in the direct product, i.e. for all $(A, B), (C, D) \in \mathcal{M} \otimes \mathcal{N}$ we declare $\mathcal{M} \otimes \mathcal{N} \models (A, B) \subseteq (C, D)$ iff $\mathcal{M} \models A \subseteq C$ and $\mathcal{N} \models B \subseteq D$,
3. $\bot$ is interpreted in $\mathcal{M} \otimes \mathcal{N}$ as $(\bot, \bot)$,

\(^1\)Perhaps unconventionally, we do not require of a partition that it comprises only non-empty sets.
4. At is interpreted in $\mathcal{M} \otimes \mathcal{N}$ as the collection of elements of the form $(A, \bot)$ where $A \in \text{At}_\mathcal{M}$ or $(\bot, B)$ where $B \in \text{At}_\mathcal{N}$.

5. $\mathcal{M} \not\models A = \bot$ and $\mathcal{N} \not\models D = \bot$, or $\mathcal{M} \models A \not\leq C$, or $\mathcal{N} \models B \not\leq D$.

6. for each $\sigma \in \Sigma$, $P_\sigma$ is given by $\mathcal{M} \otimes \mathcal{N} \models P_\sigma((A, B))$ if and only if,

$$\mathcal{M} \models P_\sigma(A) \text{ and } \mathcal{N} \models B = \bot, \text{ or, } \mathcal{M} \models A = \bot \text{ and } \mathcal{N} \models P_\sigma(B).$$

**Remark 3.7.** It is straightforward to check that for any $w, v \in \Sigma^*$ we have $M(w) \otimes M(v) \cong M(w) + M(v)$. The map $M(w) \otimes M(v) \rightarrow M(w) + M(v)$, $(A, B) \mapsto A \cup B$ is an isomorphism. The operation $\otimes$ can therefore be called a generalisation of the operation $+$, from the class $\{M(w) : w \in \Sigma^*\}$ to the class of $\mathcal{L}_\mathcal{E}^\Sigma$-structures. Moreover it is easily checked that the class of models of $T_{\text{base}(\Sigma)}$ is closed under $\otimes$.

**Definition 3.8.** ([Hod93] pp. 103). Given an $\mathcal{L}$-formula $\phi(x)$, the quantifier depth of $\phi$, denoted by $\text{qd}(\phi)$, is defined by induction on the complexity of $\phi$ as follows,

- $\text{qd}(\phi) = 0$ if $\phi$ is quantifier-free,
- $\text{qd}(\phi) = \max\{\text{qd}(\phi_i) : 1 \leq i \leq n\}$ if $\phi$ is a Boolean combination of $\mathcal{L}$-formulae $\phi_1, \ldots, \phi_n$,
- $\text{qd}(\exists \phi) = \text{qd}(\forall \phi) = \text{qd}(\phi) + 1$.

For each $k \in \mathbb{N}$ we can define an equivalence relation $\approx_k$ on the class of $\mathcal{L}$-structures by taking $\mathcal{M} \approx_k \mathcal{N}$ if and only if for each $\mathcal{L}$-sentence $\phi$ of quantifier depth at most $k$, $\mathcal{M} \models \phi$ if and only if $\mathcal{N} \models \phi$.

**Lemma 3.9.** For each $k$, the equivalence relation $\approx_k$ can also be defined in terms of one player having a winning strategy in the so-called untangled Ehrenfeucht-Fraïssé game of length $k$, which we denote by $G_k$, between two $\mathcal{L}$-structures. We make use of this equivalent formulation in the proof of the upcoming proposition [3.10] but space does not permit a full definition of these games. A full account of the games can be found in [Hod93] pp. 102.

**Proposition 3.10.** Let $\mathcal{M}_1, \mathcal{M}_2, \mathcal{N}_1, \mathcal{N}_2 \models T_{\text{base}(\Sigma)}$. Let $k \in \mathbb{N}$. If $\mathcal{M}_1 \approx_k \mathcal{N}_1$ and $\mathcal{M}_2 \approx_k \mathcal{N}_2$ then $\mathcal{M}_1 \otimes \mathcal{M}_2 \approx_k \mathcal{N}_1 \otimes \mathcal{N}_2$.

**Proof.** We outline a winning strategy of $\exists$ for the untangled Ehrenfeucht-Fraïssé game of length $k$ played on the structures $\mathcal{M}_1 \otimes \mathcal{M}_2$ and $\mathcal{N}_1 \otimes \mathcal{N}_2$, i.e. for $G_k(\mathcal{M}_1 \otimes \mathcal{M}_2, \mathcal{N}_1 \otimes \mathcal{N}_2)$. The basic idea is to fuse moves taken from winning strategies for the games $G_k(\mathcal{M}_1, \mathcal{N}_1)$ and $G_k(\mathcal{M}_2, \mathcal{N}_2)$. More explicitly, the strategy of $\exists$ is as follows. We give the case of $\forall$ moves in a move in $\mathcal{M}_1 \otimes \mathcal{M}_2$, the case where $\forall$ moves in $\mathcal{N}_1 \otimes \mathcal{N}_2$ is dealt with symmetrically.

1. Upon receiving the $i$-th move $A_i = (A_{i1}, A_{i2}) \in \mathcal{M}_1 \otimes \mathcal{M}_2$ from $\forall$, take both $A_{i1}$ and $A_{i2}$ and feed them into simulations of the games $G_k[\mathcal{M}_1, \mathcal{N}_1]$ and $G_k[\mathcal{M}_2, \mathcal{N}_2]$ respectively,

2. take the responses $B_{i1} \in \mathcal{N}_1$ and $B_{i2} \in \mathcal{N}_2$ dictated by the winning strategies for those games that we have assumed to exist,

3. form the move $B_i = (B_{i1}, B_{i2}) \in \mathcal{N}_1 \otimes \mathcal{N}_2$ as the response.

Suppose that the strategy just described is deployed, and the resulting play of the game is $((A_1, B_1), \ldots, (A_k, B_k)) \in (\mathcal{M}_1 \otimes \mathcal{M}_2) \times (\mathcal{N}_1 \otimes \mathcal{N}_2))^k$, with $A_1 = (A_{11}, A_{12}) \in \mathcal{M}_1 \times \mathcal{M}_2$ and $B_1 = (B_{11}, B_{12}) \in \mathcal{N}_1 \times \mathcal{N}_2$ for each $i$ such that $1 \leq i \leq k$.
We must check that for each unnested atomic $\mathcal{L}_\omega^\infty$-formula $\phi(\bar{X})$ we have,

$$M_1 \otimes M_2 \models \phi(\bar{A}) \iff N_1 \otimes N_2 \models \phi(\bar{B}).$$

This can be done on a case by case basis. Unnested atomic formulae in $\bar{X}$ are of the form $X = X', X \subseteq X'$, $\text{At}(X)$, $X \not< X'$, or $P_k(X)$, where $X, X' \in X \cup \{\bot\}$.

We deal with the case $X \not< X'$, and leave the remaining cases as an exercise for the reader.

Suppose,

$$M_1 \otimes M_2 \models (A_{i_1}, A_{j_1}) \not< (A_{j_1}, A_{j_2}).$$

By definition of $\otimes$ this is equivalent to at least one of the following conditions holding,

1. $M_1 \not\models A_{i_1} = \bot$ and $M_2 \not\models A_{j_2} = \bot$,
2. $M_1 \models A_{i_1} \not< A_{j_1}$,
3. $M_2 \models A_{i_2} \not< A_{j_2}$.

Now by our assumption that the pairs $(\bar{A}_1, B_1)$ and $(\bar{A}_2, B_2)$ are winning plays for $\exists$ in $G_k[M_1, N_1]$ and $G_k[M_2, N_2]$ respectively, these conditions are equivalent to the following respectively,

1. $M_1 \not\models B_{i_1} = \bot$ and $M_2 \not\models B_{j_2} = \bot$,
2. $M_1 \models B_{i_1} \not< B_{j_1}$,
3. $M_2 \models B_{i_2} \not< B_{j_2}$.

By definition of $\otimes$, the disjunction of these conditions is equivalent to,

$$N_1 \otimes N_2 \models (B_{i_1}, B_{j_2}) \not< (B_{j_1}, B_{j_2}).$$

By symmetry we can therefore conclude,

$$M_1 \otimes M_2 \models (A_{i_1}, A_{j_2}) \not< (A_{j_1}, A_{j_2}) \iff N_1 \otimes N_2 \models (B_{i_1}, B_{j_2}) \not< (B_{j_1}, B_{j_2}).$$

Hence we have outlined a winning strategy for $\exists$ in $G_k[M_1 \otimes M_2, N_1 \otimes N_2]$. \hfill $\square$

**Definition 3.11.** We write $S_k$ for the set of equivalence classes of the equivalence relation $\approx_k$ which contain a structure of the form $M(w)$ where $w \in \Sigma^*$.

Proposition 3.10 shows precisely that $\approx_k$ is a congruence (see definition 2.5) for the operation $\otimes$, we write $\otimes_k$ for the induced operation on the set of equivalence classes of $\approx_k$, restricted to the set $S_k(T_{M(\Sigma^*)})$. Note that for each $k$ the operation $\otimes_k$ is associative and admits the equivalence class containing $M(\varepsilon)$ as an identity element.

Moreover we view $S_k(T_{M(\Sigma^*)})$ as a topological space, by taking as a basis sets of the form,

$$\langle \phi \rangle := \{[\mathcal{M}]_{\approx_k} \in \mathcal{S}_k \mid \mathcal{M} \models \phi\},$$

for each $\mathcal{L}_\omega^\infty$-sentence $\phi$ with quantifier depth at most $k$.

From the Fraïssé-Hintikka theorem (see [Hod93] Theorem 3.3.2 and comments following) it follows that each of the spaces $S_k$ is finite and discrete. Therefore for each $k$, $(S_k, \otimes_k)$ is a finite discrete topological monoid.

**Lemma 3.12.** Let $\mathcal{C}$ be a class of $\mathcal{L}$-structures. For each $\mathcal{L}$-structure $\mathcal{M}$ the following are equivalent,

1. $\mathcal{M} \models \bigcap_{\mathcal{N} \in \mathcal{C}} \text{Th}(\mathcal{N})$,
2. $\text{Th}(\mathcal{M}) \subseteq \bigcup_{\mathcal{N} \in \mathcal{C}} \text{Th}(\mathcal{N})$.
Proof. We write $\text{Th}(\mathcal{F})$ for $\bigcap_{N \in \mathcal{E}} \text{Th}(\mathcal{N})$.

(1) implies (2):
Towards a contradiction suppose that $\mathcal{M} \models \text{Th}(\mathcal{F})$ holds while $\text{Th}(\mathcal{M}) \subseteq \bigcup_{N \in \mathcal{E}} \text{Th}(\mathcal{N})$ fails. Then there exists a sentence $\phi \in \text{Th}(\mathcal{M}) \setminus \bigcup_{N \in \mathcal{E}} \text{Th}(\mathcal{N})$. But now note that for any sentence $\phi$, $\phi \notin \bigcup_{N \in \mathcal{E}} \text{Th}(\mathcal{N})$ implies that $\neg \phi \in \bigcap_{N \in \mathcal{E}} \text{Th}(\mathcal{N}) = \text{Th}(\mathcal{F})$. Therefore we get that $\mathcal{M} \models \phi \land \neg \phi$, giving us a contradiction.

(2) implies (1):
Towards a contradiction suppose that $\text{Th}(\mathcal{M}) \subseteq \bigcup_{N \in \mathcal{E}} \text{Th}(\mathcal{N})$ while $\mathcal{M} \not\models \text{Th}(\mathcal{F})$. Then there exists a sentence $\phi \in \text{Th}(\mathcal{F}) \setminus \text{Th}(\mathcal{M})$. Therefore $\neg \phi \in \text{Th}(\mathcal{M}) \subseteq \bigcup_{N \in \mathcal{E}} \text{Th}(\mathcal{N})$. This implies that for some $\mathcal{N} \in \mathcal{E}$ we have $\mathcal{N} \models \phi \land \neg \phi$, giving us a contradiction. \hfill $\Box$

Definition 3.13. Let $\mathcal{M} \models T_{M(\Sigma^*)}$. For each $a \in \text{At}(\mathcal{M})$, it follows immediately from the comprehension scheme (in particular from Comp$_\sigma$ for $\phi(x,a) : x \gtrsim a \lor x = a$) that the initial interval induced by $a$ is an element of $\mathcal{M}$. We denote this element by $[0,a]$.

Lemma 3.14. Let $\mathcal{M} \models T_{M(\Sigma^*)}$. For each $a \in \text{At}(\mathcal{M})$, the set of elements which are beneath the element $[0,a]$ form the underlying set of a substructure of $\mathcal{M}$. We denote this substructure by $\mathcal{M} \upharpoonright [0,a]$.

By relativisation of quantifiers (see [Hod93] Section 5.1 pp 202), for each $L^\Sigma_w$-formula $\phi(X)$ there is a formula $\phi(\bar{x}) \upharpoonright [0,x]$ where $x$ is a ‘fresh’ variable (i.e. $x \notin X$) such that for each $a \in \text{At}(\mathcal{M})$ and $\bar{A} \in \mathcal{M} \upharpoonright [0,a]$,

$$\mathcal{M} \upharpoonright [0,a] \models \phi(\bar{A}) \iff \mathcal{M} \models \phi(\bar{A}) \upharpoonright [0,a].$$

This generalises in the following way. For each $B \in \mathcal{M}$ the set of elements beneath $B$ form the underlying set of a substructure of $\mathcal{M}$. We denote this substructure by $\mathcal{M} \upharpoonright B$. Relativising quantifiers, for each $L^\Sigma_w$-formula $\phi(X)$ there is a formula $\phi(\bar{x}) \upharpoonright Y$ where $Y \notin X$ such that for each $B \in \mathcal{M}$ and each $\bar{A} \in \mathcal{M} \upharpoonright B$,

$$\mathcal{M} \upharpoonright B \models \phi(\bar{A}) \iff \mathcal{M} \models \phi(\bar{A}) \upharpoonright B.$$

Theorem 3.15. The theory $T_{M(\Sigma^*)}$ is an axiomatisation of the pseudofinite monadic second order theory of $\Sigma$-words, $\bigcap_{\Sigma \in \mathcal{E} \setminus \mathcal{N}} \text{Th}(M(w))$.

Proof. It is clear that for each $w \in \Sigma^*$, $M(w) \models T_{M(\Sigma^*)}$, and hence $T_{M(\Sigma^*)} \subseteq \text{Th}(M(\Sigma^*))$. So we must show that if $\mathcal{M}$ is an $L^\Sigma_w$-structure and $\mathcal{M} \models T_{M(\Sigma^*)}$ then $\mathcal{M} \models \text{Th}(M(\Sigma^*))$. By lemma 3.12, $\mathcal{M} \models \text{Th}(M(\Sigma^*))$ is equivalent to $\text{Th}(\mathcal{M}) \subseteq \bigcup_{w \in \Sigma^*} \text{Th}(M(w))$.

Fix $\mathcal{M} \models T_{M(\Sigma^*)}$. It is enough to prove that for each $k \in \mathbb{N}$ there exists $w_k \in \Sigma^*$ (depending on $\mathcal{M}$) such that $\mathcal{M} \models M(w_k)$. Then $\text{Th}(\mathcal{M}) \subseteq \bigcup_{w \in \Sigma^*} \text{Th}(M(w_k)) \subseteq \bigcup_{w \in \Sigma^*} \text{Th}(M(w))$.

For each $k \in \mathbb{N}$, there is a sentence $\psi_k$ such that for each $L^\Sigma_w$-structure $\mathcal{M}$ the following are equivalent,

- $\mathcal{M} \models \psi_k$,

- $\mathcal{M} \models k \in M(w)$ for some $w \in \Sigma^*$.

Consider the formula $\psi_k \upharpoonright [0,x]$. It is clear that for each $\mathcal{M} \models T_{M(\Sigma^*)}$ we have that $\mathcal{M} \upharpoonright [0,0] \equiv M(\sigma)$ for some $\sigma \in \Sigma$. On the other hand, $\mathcal{M} \upharpoonright [0,0] = \mathcal{M}$.

Clearly $M(\sigma) \models \psi_k$ for each $k \in \mathbb{N}$, hence $\mathcal{M} \models \psi_k \upharpoonright [0,0]$.

Suppose towards a contradiction that $\mathcal{M} \not\models \psi_k$. Equivalently, suppose $\mathcal{M} \not\models \psi_k \upharpoonright [0,0]$. By comprehension there is an element of $\mathcal{M}$ which lies above precisely those $a \in \text{At}(\mathcal{M})$ for which $\mathcal{M} \not\models \psi_k \upharpoonright [0,a]$. This element is not $\bot$, as we noted that it contains $0^*$, therefore it must contain a smallest atom $a \in \text{At}(\mathcal{M})$ (with respect to $\gtrsim$). Now since $\mathcal{M} \models \psi_k \upharpoonright [0,0]$, we must
have \( a \neq 0 \), as such \( a \) has an immediate predecessor \( b \in \text{At}(\mathcal{M}) \). We then get a contradiction using proposition 3.10 and the observation that \( \mathcal{M} \upharpoonright [0, a] \cong \mathcal{M} \upharpoonright [0, b] \otimes M(\sigma) \) where \( \sigma \) is the unique element of \( \Sigma \) such that \( M|_{0, a} = P_{\sigma}(a) \). Since \( \mathcal{M} \upharpoonright [0, b] = \psi_k \) there exists \( w \in \Sigma^* \) such that \( \mathcal{M} \upharpoonright [0, b] \cong M(w) \), and therefore \( \mathcal{M} \upharpoonright [0, a] \cong_k M(w') \) where \( w' \in \Sigma^* \) is given by appending \( \sigma \) to \( w \).

This then implies \( \mathcal{M} \upharpoonright [0, a] \models \psi_k \), giving us a contradiction.

Therefore \( \mathcal{M} \models \psi_k \) as required. \( \square \)

4 The free profinite monoid generated by \( \Sigma \).

**Definition 4.1.** Let \( \Sigma \) be a finite set. A subset \( L \) of \( \Sigma^* \) is called a **recognisable \( \Sigma \)-language** (or a recognisable subset of \( \Sigma^* \)) if there exists a finite semigroup \( S \), a subset \( F \subseteq S \), and a homomorphism \( f : \Sigma^* \to S \) such that \( L = f^{-1}(F) \).

**Remark 4.2.** It is well-known that the collection \( \text{Rec}(\Sigma) \) of recognisable subsets of \( \Sigma^* \) forms a Boolean algebra under the set-theoretic operations of intersection, union, and complement (with the empty set and \( \Sigma^* \) being the bottom and top elements respectively).

Recognisable subsets can also be characterised as the languages accepted by certain finite automata, or as those sets of finite words which are finitely axiomatisable in monadic second order logic (see theorem 4.6).

It is also well known that the collection of recognisable languages is closed under the complexification of concatenation \( \hat{\cdot} : \mathcal{P}(\Sigma^*)^2 \to \mathcal{P}(\Sigma^*) \) (see notation 2.4). It is also straightforward to check that the complexification is both normal and additive (see definition 2.1). It therefore makes sense to talk about the extended Stone dual of \( (\text{Rec}(\Sigma), \hat{\cdot}) \) in the sense of section 2.1.

In [GGP10] the following theorem is stated, but a proof is not given.

**Theorem 4.3 ([GGP10] Theorem 6.1).** The Boolean algebra \( \text{Rec}(\Sigma) \) of recognisable \( \Sigma \)-languages together with the binary operation of concatenation is the extended Stone dual of \( \hat{\Sigma}^* \), the free profinite monoid on \( \Sigma \).

Here the Boolean space underlying \( \hat{\Sigma}^* \) is viewed together with the graph of the monoid operation which it carries, which is a compatible ternary relation. It is said in [GGP10] that a proof would require advanced machinery from duality theory. This theorem can be viewed as a special case of the following considerably more general theorem (by taking \( A \) to be \( \Sigma^* \)).

**Theorem 4.4 ([Geh16] Theorem 4.5).** Let \( A \) be an abstract algebra. The profinite completion \( \hat{A} \) (which is defined in a similar way to the profinite completion of a monoid in definition 2.11) is, up to isomorphism as topological algebras, the extended Stone dual of \( \text{Rec}(A) \), the Boolean algebra of recognisable subsets of \( A \) equipped with residuation operations of recognisable subsets.

Discussion of residuation is outside the scope of this article, and in the cases we are interested in (free profinite monoids on finitely many generators) considering them is unnecessary. A full proof of this theorem is presented at length in [Geh16] alongside other results from duality theory. The purpose of this section is to offer a proof of the more specific theorem from [GGP10], making use of our model-theoretic analysis of \( \text{Th}(\mathcal{M}(\Sigma^*)) \) to present a concise proof which does not rely on advanced machinery from duality theory.

**Definition 4.5.** Let \( T \) be an \( \mathcal{L} \)-theory. The **Lindenbaum-Tarski algebra** of \( T \), which we denote by \( \text{LT}(T) \), is defined to be the Boolean algebra of equivalence classes of \( \mathcal{L} \)-sentences under the equivalence relation \( \sim_T \) given by \( \phi \sim_T \psi \) if and only if,

\[
T \models \phi \leftrightarrow \psi.
\]
Note that this tacitly exploits the fact that $\sim_T$ is a congruence (see definition 2.8) the generalisation of which to the setting of universal algebra is straightforward for the Boolean operations of conjunction, disjunction, and negation. When working with $LT(T)$ we write $\phi$ in place of $[\phi]_{\sim_T}$ to avoid bloated notation.

The following is a concise (but anachronistic) formulation of a theorem of B"uchi on the monadic second order logic of finite words. It conceals the automata theoretic machinery used in the original proofs. It is also worth noting that B"uchi did not work in the signature $\mathcal{L}_2^0$ which we make use of here.

**Theorem 4.6** (B"uchi). Let $\Sigma$ be a finite set. The map $SPEC : LT(T_{M(\Sigma^*)}) \rightarrow Rec(\Sigma)$ given by $\phi \mapsto \{w \in \Sigma^* : M(w) \models \phi\}$ is an isomorphism of Boolean algebras.

In the following proposition we exploit the shorthand for restrictions and relativisations laid out in lemma 3.14.

**Proposition 4.7.** The map $SPEC$ is still an isomorphism if we enrich the two Boolean algebras as follows,

(i) to $LT(T_{M(\Sigma^*)})$ we append the operation $+$ given by,

$$\phi + \psi := \exists X (dwcl(X) \land \phi \land X \land \exists X),$$

where $dwcl(X)$ (downwards closed) is the formula $\forall y, z ((X(y) \land z < y) \rightarrow X(z))$.

(ii) to $Rec(\Sigma)$ we append the operation $\cdot$ given by,

$$L_1 \cdot L_2 := \{w \cdot v : w \in L_1, v \in L_2\}.$$  

**Proof.** All that is required, in light of theorem 4.6, is to show that for any $\phi, \psi \in LT$,

$$SPEC(\phi + \psi) = SPEC(\phi) \cdot SPEC(\psi).$$  

If $w \in SPEC(\phi + \psi)$ this means that $M(w) \models \phi + \psi$, i.e. there is $A \in M(w)$ such that,

$$M(w) \models dwcl(A) \land \phi \land A \land \exists X.$$  

Therefore we get that $M(w) \upharpoonright A \models \phi$ and $M(w) \upharpoonright A^c \models \psi$. Now it is clear, since $M(w) \models dwcl(A)$, that $M(w) \upharpoonright A \cong M(v_1)$ and $M(w) \upharpoonright A^c \cong M(v_2)$ for some $v_1, v_2 \in \Sigma^*$ such that $v_1 \cdot v_2 = w$. As $M(w) \models \phi \upharpoonright A$ is equivalent to $M(v_1) \models \phi$ and $M(w) \models \psi \upharpoonright A^c$ is equivalent to $M(v_2) \models \psi$, we therefore have $w \in SPEC(\phi) \cdot SPEC(\psi)$.

For the converse suppose that $w \in SPEC(\phi) \cdot SPEC(\psi)$. Then there exists $v_1 \in SPEC(\phi)$ and $v_2 \in SPEC(\psi)$ such that $w = v_1 \cdot v_2$. Then it is clear that there exists $A \in M(w)$, namely the initial segment corresponding to $v_1$, such that,

$$M(w) \models dwcl(A) \land \phi \land A \land \exists X.$$  

Hence $w \in SPEC(\phi + \psi)$ as required. 

Proposition 4.7 says that $\langle Rec(\Sigma), \cdot \rangle$ and $\langle LT(T_{M(\Sigma^*)}), + \rangle$ differ only superficially. So towards a proof of theorem 4.3 we can consider the extended Stone dual of $LT(T_{M(\Sigma^*)})$.

**Theorem 4.8.** The topological monoid $\langle S(T_{M(\Sigma^*)}), \otimes \rangle$ is the extended Stone dual of the Boolean algebra with operator $\langle LT(T_{M(\Sigma^*)}), + \rangle$.  

11
Proof. The extended Stone dual of \((LT(T_{M(\Sigma^*)}),+)\) is, by definition (see theorem 2.3), the Boolean space with compatible operation \((S(T_{M(\Sigma^*)}), R_+)\), where \(R_+\) is the ternary relation on \(S(T_{M(\Sigma^*)})\) given by,

\[ R_+(x, y, z) \iff x + y \subseteq z. \]

It is therefore enough to show that for each \(x, y \in S(T_{M(\Sigma^*)})\), \(x \otimes y\) is the unique prime filter on \(LT(T_{M(\Sigma^*)})\) such that, \(x + y \subseteq x \otimes y\). In other words, we will show that \(R_+\) is the graph of \(\otimes\).

We will in fact show that \(R_+\) is a recognisable language. We can enumerate these equivalence classes as \(\eta\), \(\phi\), for each \(\psi\). Let \(\phi \in x\) and \(\psi \in y\) be given. Recall that \(\phi + \psi\) is defined as the sentence,

\[ \exists X (dwcl(X) \land \phi \land X \land \psi \land X^c). \]

Let \(M \models x\) and \(N \models y\), so that \(M \otimes N \models x \otimes y\). Then consider the element \(A = (\top, \bot) \in M \otimes N\). It is immediate that \(A^c = (\bot, \top)\), \(M \otimes N \models A \cong M\), and \(M \otimes N \models A^c \cong N\). Therefore,

\[ M \otimes N \models dwcl(A) \land \phi \land A \land \psi \land A^c. \]

So \(M \otimes N \models \phi + \psi\), and hence \(\phi + \psi \in x \otimes y\) as required.

To conclude we must show that \(x \otimes y \subseteq (x + y)\). We show that each sentence \(\eta \in x \otimes y\) is equivalent over \(T_{M(\Sigma^*)}\) to a finite disjunction of sentences contained in \(x + y\). Take \(\eta \in x \otimes y\). Let \(M \models x\) and \(N \models y\), so that \(M \otimes N \models \eta\). We want to show that there exist \(m \in \mathbb{N}\), \(\phi_1, \ldots, \phi_m \in x\), and \(\psi_1, \ldots, \psi_m \in y\) such that for \(i = 1, \ldots, m\),

\[ T_{M(\Sigma^*)} \models \eta \iff \bigvee_{i=1}^m (\phi_i + \psi_i). \]

Without loss of generality we can assume that \(\eta\) is an unnested sentence of quantifier rank \(k \in \mathbb{N}\). The equivalence relation \(\approx_k\) on the class of \(L^\oplus_{\Sigma}\)-structures admits finitely many equivalence classes, each of which is axiomatised by an \(L^\oplus_{\Sigma}\)-sentence. Taking the finite disjunction of all sentences of the form \(\phi + \psi\), where \(\phi \in x\) and \(\psi \in y\) each axiomatise an equivalence class of \(\approx_k\), which \(\otimes_k\) (see definition 3.11) sends to an equivalence class of \(\approx_k\) containing a model of \(\eta\). This disjunction is equivalent to \(\eta\) over \(T_{M(\text{Fin})}\) so we are done. \(\square\)

Remark 4.9. By taking \(w \approx_k v\) if and only if \(M(w) \approx_k M(v)\), for each \(w, v \in \Sigma^*\), we may think of \(\approx_k\) as an element of \(\text{Con}_w(\Sigma^*)\). In theorem 4.10 and the remainder of the paper we will switch freely between these two views of \(\approx_k\).

Theorem 4.10. The congruences \(\approx_k\) for \(k \in \mathbb{N}\) form a cofinal subset of the poset \((\text{Con}_w(\Sigma^*), \supseteq)\).

Proof. Let \(\rho \in \text{Con}_w(\Sigma^*)\). We must show that there is \(k \in \mathbb{N}\) such that \(w \approx_k w'\) implies \(w \sim_\rho w'\) for each \(w, w' \in \Sigma^*\). Consider the quotient \(\Sigma^*/\rho\). For each \(w \in \Sigma^*/\rho\), we have that the preimage \(\pi^{-1}_\rho([w]_\rho) \subseteq \Sigma^*\) is a recognisable language. We can enumerate these equivalence classes as \(L_1, \ldots, L_n \subseteq \Sigma^*\). Note that these equivalence classes form a partition of \(\Sigma^*\). By theorem 1.6 there are \(L^\oplus_{\Sigma}\)-formulas \(\phi_1, \ldots, \phi_n\) which axiomatise each of these regular languages, meaning for each \(i\) such that \(1 \leq i \leq n\) we have \(L_i = \{ w \in \Sigma^* : M(w) \models \phi_i \}\). Without loss of generality we may assume that each of the sentences \(\phi_1, \ldots, \phi_n\) are unnested. Now let \(k = \max_{1 \leq i \leq n} \text{qd} (\phi_i)\). We claim that \(w \approx_k w'\) implies \(w \sim_\rho w'\). But this is immediate, for if \(w \approx_k w'\) then we have
Theorem 4.11. The topological monoid \((S(T_{M(\Sigma^*)}), \otimes)\) is the inverse limit of the directed system of topological monoids given by restricting \(F_{\Sigma^*} : \text{Con}_c(\Sigma^*) \to \text{TopMon}\) (see definition \[2.11\]) to the cofinal subset \(C\) of \((\text{Con}_c(\Sigma^*), \supseteq)\) comprising the congruences \(\approx_k\) for \(k \in \mathbb{N}\).

Proof. For each \(k\) we have a morphism (a continuous map which respects the binary operation) \(p_k : S(T_{M(\Sigma^*)}) \to S_k\) given by taking a completion of \(T_{M(\Sigma^*)}\) to the \(\approx_k\) class of any of its models. These morphisms commute with the parent maps which make up the directed system (see definition \[2.11\]).

Now suppose \((X, \times)\) is some other topological monoid, and that for each \(k \in \mathbb{N}\) we have a morphism \(q_k : X \to S_k\). Moreover suppose these morphisms \(q_k\) commute with the parent maps. We must show that there exists a unique morphism \(g : X \to S(T_{M(\Sigma^*)})\) such that \(p_k \circ g = q_k\) for each \(k \in \mathbb{N}\). Taking \(x \in X\), we can form \(g(x) := (q_k(x))_{k \in \mathbb{N}}\). Similarly for \(x \in S\) we can form \(\overline{p}(x) := (p_k(x))_{k \in \mathbb{N}}\). By a straightforward application of the compactness theorem, for each \(x \in X\) there exists \(g(x) \in S(T_{M(\Sigma^*)})\) such that \(p_k \circ g(x) = q_k(x)\) for each \(k \in \mathbb{N}\). That \(g(x)\) is unique is clear.

So far we have that \(g\) is a well defined function, but we must check that it is also continuous and respects the binary operations. Let \(x, x' \in X\) and consider \(g(x \times x')\), we want to show that \(g(x \times x') = g(x) \otimes g(x')\). For each \(k \in \mathbb{N}\) we have,

\[
p_k(g(x \times x')) = q_k(x \times x')
\]
\[
= q_k(x) \otimes_k q_k(x'),
\]
\[
= p_k(g(x)) \otimes_k p_k(g(x'))
\]
\[
= p_k(g(x) \otimes g(x')).
\]

But if two elements of \(S(T_{M(\Sigma^*)})\) agree with respect to all of the morphisms \(p_k\), then they must in fact be equal, hence \(g(x \times x') = g(x) \otimes g(x')\). The topology on \(S(T_{M(\Sigma^*)})\) has a basis of clopens comprising sets of the form \(\langle \phi \rangle := \{x : x \models \phi\}\) where \(\phi\) is a sentence. Consider \(g^{-1}(\langle \phi \rangle)\). Taking \(k = \text{qd}(\phi)\), there exists a subset \(F \subseteq S_k\) such that \(\langle \phi \rangle = p_k^{-1}(F)\). Then it follows at once from the fact that \(p_k \circ g = q_k\) that \(g^{-1}(\langle \phi \rangle) = q_k^{-1}(F)\), which is clopen as it is the preimage of a continuous map into a finite discrete space. Hence \(g\) is continuous.

Having verified directly that \((S(T_{M(\Sigma^*)}), \otimes)\) satisfies the universal property for inverse limits the proof is complete.

Theorem 4.12. The topological monoids \((S(T_{M(\Sigma^*)}), \otimes)\) and \(\tilde{\Sigma^*}\) are isomorphic.

Proof. By definition \[2.11\] \(\tilde{\Sigma^*}\) is \(\lim F_{\Sigma^*}\), where \(F_{\Sigma^*} : \text{Con}_c(\Sigma^*) \to \text{TopMon}\) takes a finite index congruence \(\rho\) to the quotient \(\Sigma^*/\rho\). In theorem \[4.10\] it was shown that the congruences of the form \(\approx_k\), for \(k \in \mathbb{N}\), form a cofinal subset \(C\) of the poset \((\text{Con}_c(\Sigma^*), \supseteq)\). Hence \(\tilde{\Sigma^*}\) is \(\lim F'\). Lemma \[2.10\] then tells us that \(\lim F_{\Sigma^*} = \lim F'\) where \(F'\) is the restriction of \(F_{\Sigma^*}\) to \(C\). Finally, theorem \[4.11\] says that \(\lim F' = (S(T_{M(\Sigma^*)}), \otimes)\). Uniqueness of inverse limits up to isomorphism then allows us to conclude that \((S(T_{M(\Sigma^*)}), \otimes)\) and \(\tilde{\Sigma^*}\) are isomorphic as required.

Corollary 4.13. The extended Stone dual of \((\text{Rec}(\Sigma), \uparrow)\) is the topological monoid \(\tilde{\Sigma^*}\).
References

[Geh16] M. Gehrke. “Stone duality, topological algebra, and recognition”. In: *J. Pure Appl. Algebra* 220.7 (2016), pp. 2711–2747.

[GGP10] M. Gehrke, S. Grigorieff, and J.-É. Pin. “A Topological Approach to Recognition”. In: *Automata, Languages and Programming*. Ed. by S. Abramsky et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 151–162.

[Gol89] R. Goldblatt. “Varieties of complex algebras”. In: *Ann. Pure Appl. Logic* 44.3 (1989), pp. 173–242.

[GS19] S. J. van Gool and B. Steinberg. “Pro-aperiodic monoids via saturated models”. In: *Israel J. Math.* 234.1 (2019), pp. 451–498.

[Han83] G. Hansoul. “A duality for Boolean algebras with operators”. In: *Algebra Universalis* 17.1 (1983), pp. 34–49.

[Hod93] W. Hodges. *Model theory*. Vol. 42. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1993, pp. xiv+772.

[Lei14] T. Leinster. *Basic category theory*. Vol. 143. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 2014, pp. viii+183.

[Linkhorn21] D. Linkhorn. “Monadic Second Order Logic and Linear Orders”. (PhD thesis, available at [http://www.deaconlinkhorn.com](http://www.deaconlinkhorn.com) or by email request.) University of Manchester, 2021.

[Mac98] S. Mac Lane. *Categories for the working mathematician*. Second. Vol. 5. Graduate Texts in Mathematics. Springer-Verlag, New York, 1998, pp. xii+314.