A MODEL-THEORETIC CHARACTERIZATION OF
MONADIC SECOND ORDER LOGIC ON INFINITE WORDS

SILVIO GHILARDI AND SAMUEL J. VAN GOOL

Abstract. Monadic second order logic and linear temporal logic are two logical formalisms that can be used to describe classes of infinite words, i.e., first-order models based on the natural numbers with order, successor, and finitely many unary predicate symbols.

Monadic second order logic over infinite words (S1S) can alternatively be described as a first-order logic interpreted in $\mathcal{P}(\omega)$, the power set Boolean algebra of the natural numbers, equipped with modal operators for 'initial', 'next' and 'future' states. We prove that the first-order theory of this structure is the model companion of a class of algebras corresponding to a version of linear temporal logic (LTL) without until.

The proof makes crucial use of two classical, non-trivial results from the literature, namely the completeness of LTL with respect to the natural numbers, and the correspondence between S1S-formulas and Büchi automata.

§1. Introduction. Monadic second order logic over the natural numbers with successor operation is a rather expressive, but still decidable formalism. The decision result, originally due to Büchi [2], makes use of a conversion between logic and automata. The key idea is to view interpretations of unary predicates over natural numbers as infinite words over a suitable alphabet: one can associate an automaton with a formula, and vice versa, in such a way that, roughly speaking, the automaton accepts exactly those infinite words that, viewed as interpretations of second-order variables, satisfy the formula. Converting a formula φ into an automaton A_φ and then the automaton A_φ again into a formula, one does not get back the same formula φ, but a formula φ' which is equivalent to φ in the intended model of the natural numbers. Loosely speaking, one may view φ' as a kind of 'normalization' of φ. Morally, φ' is an existential formula; although this is not formally true, the 'existential nature' of φ' is rather evident. Nevertheless, in order to make φ' into an actual existential formula, an enlargement of the language is needed.

In this paper we make the above observations precise, which allows us to fit monadic second order logic into the framework of modern model-theoretic algebra (Robinson's legacy [14]), using the concepts of model completeness and model companions. We briefly recall these concepts now; precise definitions are

---

given in Section 2. Theories T^* with the property that every formula is equivalent modulo T^* to an existential (equivalently: to a universal) formula are called *model complete*; model complete theories T^* can also be characterized semantically by requiring that all embeddings between models of T^* are elementary. Algebraically closed fields and real closed fields are classical examples of model complete theories (see, e.g., [16] for many further examples). Moreover, as the previous paragraph indicates, monadic second order logic, viewed as the first-order theory of $\mathcal{P}(\omega)$, is close to being model complete; we will see later in this paper (Section 4) that it actually becomes model complete in an enriched language. If T is the set of universal consequences of a model complete theory T^*, then T^* axiomatizes the class of structures which, as models of T, are *existentially closed*: for the case of algebraically closed fields, T is the theory of integral domains, for the case of real closed fields, T is the theory of formally real fields, etc. In this situation, T^* is called a *model companion* of T.

In this paper, we exhibit a natural enrichment of the language of monadic second order logic, viewed as the first-order theory of $\mathcal{P}(\omega)$, so that the theory becomes model complete. Moreover, we axiomatize the underlying universal fragment, i.e., the universal theory T of which monadic second order logic over natural numbers is the model companion. Our idea is to use *modal and temporal logic* to attain this goal. More precisely: we will use a variant of the modal logic LTL ('linear temporal logic' [12]) with an actual instant of time (i.e. with an atom for zero) and without until; i.e., we have a basic modal logic with a reachability operator and one additional constant.[1] We introduce a universal class of algebras algebraizing this logic and we show that this class fits our purposes. Indeed, this class is axiomatized by a theory T admitting a model companion T^* which is exactly monadic second order logic over the natural numbers (Theorem 2.2). In order to establish this result, we shall make use, besides automata techniques, of standard modal logic machinery, in particular, duality and filtrations [3, 9, 10].

The paper is organized in the following way: in Section 2 we state our result, in the (rather technical) Section 3 we give the part of the proof requiring modal logic techniques (completeness of the axiomatization) and in Section 4 the part of the proof requiring automata theory ingredients (model completeness of the theory in the enriched language). Section 5 concludes and indicates directions for future work.

§2. **Definitions and statement of the main theorem.** In the following definition, we introduce the relevant class of algebras.

DEFINITION 2.1. A $(\Diamond, \mathbf{X}, I)$-*algebra* is a tuple $(A, \vee, \neg, \bot, \Diamond, \mathbf{X}, I)$, where $(A, \vee, \neg, \bot)$ is a Boolean algebra, $\Diamond$ is a unary normal modal operator on A (i.e., a $(\vee, \bot)$-semilattice endomorphism of A), $\mathbf{X}$ is a Boolean endomorphism of A, I is an element of $A \setminus \{\bot\}$, and, for any $a \in A$, the following conditions hold:

---

[1] The reason for leaving out the Until operator is that we opted for as small a language as possible; Until is not needed for our purpose of expressing the Büchi acceptance condition. In fact, any operator definable in monadic second order logic over the natural numbers could be added to our T (together with its axiomatization) without affecting our results.

(i) $\Diamond a = a \vee \mathbf{X}\Diamond a$,   (iii) if $a \neq \bot$ then $I \leq \Diamond a$.

(ii) if $\mathbf{X}a \leq a$ then $\Diamond a \leq a$,   (iv) $\mathbf{X}I = \bot$.

The class of $(\Diamond, \mathbf{X}, I)$-algebras (which is a *universal* class) algebraizes a version of linear temporal logic without the until connective and with a constant I for 'initial element'. The structure of $(\Diamond, \mathbf{X}, I)$-algebras will be studied in detail in Section 3.

An important example of a $(\Diamond, \mathbf{X}, I)$-algebra is the *power set algebra of the natural numbers*, $\mathcal{P}(\omega)$, equipped with the usual Boolean operations, and the operations $\Diamond S := \{n \in \omega \mid n \leq s \text{ for some } s \in S\}$, $\mathbf{X}S := \{n \in \omega \mid n+1 \in S\}$, and $I := \{0\}$. We will see in Section 4 that first-order formulas in the signature of $(\Diamond, \mathbf{X}, I)$-algebras, interpreted in $\mathcal{P}(\omega)$, are interdefinable with formulas in the system S1S, monadic second order logic over the natural numbers with order and successor relations.

If T and T^* are first-order theories in the same signature, recall that T^* is called a *model companion* of T if (i) the theories T and T^* have the same universal consequences (i.e., T^* is a *companion* or *co-theory* of T) and (ii) any first-order formula is equivalent over T^* to an existential formula (i.e., T^* is *model complete*). The model companion of T is unique if it exists, and in this case it is the theory of the existentially closed T-models [16].

Our main theorem is the following.

THEOREM 2.2. *The first-order theory of the $(\Diamond, \mathbf{X}, I)$-algebra $\mathcal{P}(\omega)$ is the model companion of the first-order theory of $(\Diamond, \mathbf{X}, I)$-algebras.*

In the rest of this paper, we denote by T the first-order theory of $(\Diamond, \mathbf{X}, I)$-algebras, and by T^* the first-order theory of the $(\Diamond, \mathbf{X}, I)$-algebra $\mathcal{P}(\omega)$. In Section 3, we prove that T^* is a companion of T. In Section 4, we prove that T^* is model complete. Together, these two sections prove Theorem 2.2.

§3. T^* **is a companion of** T. In this section, we prove that T^* is a companion of T, i.e., that T and T^* have the same universal consequences.

Notation. Throughout this section, we denote by $\mathcal{L} = \{\vee, \neg, \bot, \Diamond, \mathbf{X}, I\}$ the first-order signature of $(\Diamond, \mathbf{X}, I)$-algebras. We employ the usual abbreviations $a \wedge b := \neg(\neg a \vee \neg b)$, $\Box a := \neg\Diamond\neg a$, $a \to b := \neg a \vee b$, and $a \leftrightarrow b := (a \to b) \wedge (b \to a)$.

We begin with a useful observation: in the theory T, we can reduce quantifier-free $\mathcal{L}$-formulas to single equations.

LEMMA 3.1. *For any quantifier-free $\mathcal{L}$-formula φ, there exists an $\mathcal{L}$-term t such that φ is equivalent to $t = \top$ in T.*

PROOF. Observe that, for any element w in a $(\Diamond, \mathbf{X}, I)$-algebra, we have

$$w \neq \top \text{ if, and only if, } I \leq \Diamond\neg w \text{ if, and only if, } I \to \Diamond\neg w = \top. \qquad (*)$$

Let φ be a quantifier-free formula. We may assume that φ is a disjunction of formulas $\psi_1, \ldots, \psi_n$, where each ψ_j is a conjunction of literals, i.e., has the form

$$r_1^j \neq s_1^j \text{ and } \ldots \text{ and } r_k^j \neq s_k^j \text{ and } u_1^j = v_1^j \text{ and } \ldots \text{ and } u_\ell^j = v_\ell^j,$$

where the r_i^j, s_i^j, u_i^j, v_i^j are $\mathcal{L}$-terms. Define the $\mathcal{L}$-terms

$$t_j := \bigwedge_{i=1}^{k} (I \to \Diamond \neg (r_i^j \leftrightarrow s_i^j)) \wedge \bigwedge_{i=1}^{\ell} (u_i^j \leftrightarrow v_i^j), \quad t := I \to \bigvee_{j=1}^{n} \Box t_j.$$

Using (*) and basic facts about Boolean algebras, we obtain that ψ_j is equivalent in T to $t_j = \top$, so φ is equivalent in T to the disjunction ($t_1 = \top$ or ... or $t_n = \top$). This disjunction, in turn, is equivalent in T to $t = \top$, as can be proved easily using (*) again, together with the axioms $\Box\top = \top$ and $I \neq \bot$.   ⊣

In order to prove that T^* is a companion of T, since $T \subseteq T^*$, it suffices to prove that any universal formula that is valid in $\mathcal{P}(\omega)$ is valid in any $(\Diamond, \mathbf{X}, I)$-algebra. Therefore, by Lemma 3.1, it is enough to prove the following theorem.

THEOREM 3.2. *If t is an $\mathcal{L}$-term and $\mathcal{P}(\omega) \models t = \top$, then, for any $(\Diamond, \mathbf{X}, I)$-algebra A, $A \models t = \top$.*

In the rest of this section, we prove Theorem 3.2. The proof we give here is an adaptation to our setting of the standard completeness theorem for LTL with respect to ω; in particular, it is directly inspired by the proof given in [10, Ch. 9]. Our proof of Theorem 3.2 is structured as follows. We will first show, in Theorem 3.4, that $(\Diamond, \mathbf{X}, I)$-algebras can be represented as dual algebras of $(\Diamond, \mathbf{X}, I)$-*spaces* (Definition 3.3), by means of an adaptation of the standard (Stone-)Jónsson-Tarski representation theorem. We then combine this theorem with an adaptation of a filtration argument for LTL (Lemma 3.8) to prove Theorem 3.2.

Below, we use the following notational conventions for a preorder (= set endowed with a reflexive and transitive relation) $(X, \leq)$. We define $x < y$ as ($x \leq y$ and $x \neq y$); for $Y \subseteq X$, we write $\uparrow Y := \{x \in X \mid y \leq x \text{ for some } y \in Y\}$, and similarly $\downarrow Y$; finally, for $x \in X$, we write $\uparrow x$ and $\downarrow x$ as shorthand for $\uparrow\{x\}$ and $\downarrow\{x\}$, respectively.

We will now formulate the duality between algebras and general frames (viewed as Stone spaces) that we need here. Our exposition will be brief, as we are only using standard modal logic machinery. We refer to, e.g., [1, Ch. 5], for more details on duality for modal algebras. To motivate the definition of $(\Diamond, \mathbf{X}, I)$-spaces, observe that the $\Diamond$-reduct of a $(\Diamond, \mathbf{X}, I)$-algebra is an S4-algebra. Through Jónsson-Tarski representation, S4-algebras correspond to Boolean spaces equipped with a *topological preorder*, i.e., a preorder such that $\uparrow x$ is closed for any point x and $\downarrow K$ is clopen for any clopen set K (see, e.g., [5, Rem. 1 in Sec. 2.6]). The additional structure and properties of $(\Diamond, \mathbf{X}, I)$-algebras now correspond to additional structure and properties of these preordered Boolean spaces, as follows.

DEFINITION 3.3. We define a $(\Diamond, \mathbf{X}, I)$-*space*[2] to be a tuple $(X, \leq, f, x_0)$, where X is a Boolean topological space, $\leq$ is a topological preorder on X, $f : X \to X$ is a continuous function, $x_0 \in X$ is a point such that $\{x_0\}$ is clopen, and, for any $x, y \in X$ and clopen $K \subseteq X$:

---

[2]Note that our definition of $(\Diamond, \mathbf{X}, I)$-spaces makes crucial use of the second-order structure (topology) on the underlying Kripke frames. This is necessarily so: the class of $(\Diamond, \mathbf{X}, I)$-algebras is not canonical, so it can not be dual to an elementary class of Kripke frames, by a theorem of Fine [4].

(i) $x \leq f(x)$, and if $x < y$ then $f(x) \leq y$,

(ii) if $f(K) \subseteq K$ then $\uparrow\!K \subseteq K$,

(iii) $x_0 \leq x$,

(iv) $f(x) \neq x_0$.

The *dual algebra* of a $(\Diamond, \mathbf{X}, I)$-space $(X, \leq, f, x_0)$ is defined to be the tuple $(A, \Diamond, \mathbf{X}, I)$, where A is the Boolean algebra of clopen subsets of X, and for any $K \in A$, $\Diamond K := {\downarrow}K$, $\mathbf{X}K := f^{-1}(K)$, and $I := \{x_0\}$.

We now prove our representation theorem for $(\Diamond, \mathbf{X}, I)$-algebras.

THEOREM 3.4. *The class of $(\Diamond, \mathbf{X}, I)$-algebras coincides with the class of algebras that are isomorphic to the dual algebra of a $(\Diamond, \mathbf{X}, I)$-space.*

PROOF. Note that, *in any $(\Diamond, \mathbf{X}, I)$-algebra, I is an atom*: let $\bot < a \leq I$ be arbitrary. Then $I \leq \Diamond a$ by Def. 2.1(iii). Also, $\Diamond I \leq I$ by Def. 2.1(ii) and (iv). Using these facts, Def. 2.1(i) and the monotonicity of $\mathbf{X}$ and $\Diamond$, we obtain that $I \leq a \vee \mathbf{X}\Diamond a \leq a \vee \mathbf{X}\Diamond I \leq a \vee \mathbf{X}I$, i.e., $I \leq a$ by Def. 2.1(iv), as required.

Now, by the Jónsson-Tarski representation theorem (see, e.g., [1, Section 5.3]), the class of algebras $(A, \Diamond, \mathbf{X}, I)$ where $(A, \Diamond)$ is an S4-algebra, $\mathbf{X}$ is an endomorphism, and I is an atom of A, coincides with the class of algebras that are isomorphic to the dual algebra of a tuple of the form $(X, \leq, f, x_0)$, where X is a Boolean space, $\leq$ is a topological preorder on X, f is a continuous function on X and $x_0 \in X$ is such that $\{x_0\}$ is clopen.

It remains to prove that $(X, f, \leq, x_0)$ validates (i)-(iv) in the definition of $(\Diamond, \mathbf{X}, I)$-space if, and only if, its dual algebra A is a $(\Diamond, \mathbf{X}, I)$-algebra. This follows from the following claim.

Claim. The following equivalences hold.

(a) $x \leq f(x)$ for all $x \in X$ $\iff$ $\mathbf{X}\Diamond a \leq \Diamond a$ for all $a \in A$.

(b) if $x < y$ then $f(x) \leq y$ for all $x, y \in X$ $\iff$ $\Diamond a \leq a \vee \mathbf{X}\Diamond a$ for all $a \in A$.

(c) $(X, f, \leq)$ validates (ii) in Definition 3.3 $\iff$ A validates (ii) in Definiton 2.1.

(d) $x_0 \leq x$ for all $x \in X$ $\iff$ $I \leq \Diamond a$ for all $a \in A \setminus \{\bot\}$.

(e) $f(x) \neq x_0$ for all $x \in X$ $\iff$ $\mathbf{X}I = \bot$.

Proof of Claim. (a) ($\Rightarrow$) Let $a \in A$ be arbitrary, and suppose that $x \in \mathbf{X}\Diamond a$. Then $f(x) \in \Diamond a$, and $x \leq f(x)$, so $x \in \Diamond a$.

($\Leftarrow$) Suppose that $x \not\leq f(x)$ for some $x \in X$. Since $\uparrow\!x$ is closed, there exists a clopen set $a \in A$ such that $f(x) \in a$ and $\uparrow\!x \cap a = \emptyset$. In particular, $f(x) \in \Diamond a$, so $x \in \mathbf{X}\Diamond a$, but $x \notin \Diamond a$.

(b) ($\Rightarrow$) Let $a \in A$ be arbitrary, and suppose that $x \in \Diamond a$ but $x \notin a$. Since $x \in \Diamond a$, pick $y \geq x$ such that $y \in a$. Since $x \notin a$, we have $x \neq y$, so $f(x) \leq y$. Hence, $x \in \mathbf{X}\Diamond a$.

($\Leftarrow$) Suppose that there exist $x, y \in X$ such that $x < y$, but $f(x) \not\leq y$. Since $\uparrow\!f(x)$ is closed, pick $a_1 \in A$ such that $y \in a_1$ and $\uparrow\!f(x) \cap a_1 = \emptyset$. Since $x \neq y$, pick $a_2 \in A$ such that $y \in a_2$ and $x \notin a_2$. Let $a := a_1 \cap a_2$. Note that $x \in \Diamond a$, since $x \leq y$ and $y \in a$. However, we have $x \notin a$, and $x \notin \mathbf{X}\Diamond a$, contrary to the assumption that $\Diamond a \leq a \vee \mathbf{X}\Diamond a$ for all $a \in A$.

(c) Note that $f(K) \subseteq K$ if, and only if, $f^{-1}(K^c) \subseteq K^c$, and that $\uparrow\!K \subseteq K$ if, and only if, ${\downarrow}(K^c) \subseteq K^c$. The stated equivalence now follows from the definitions of $\mathbf{X}$ and $\Diamond$.

(d) ($\Rightarrow$) If $a \in A \setminus \{\bot\}$, then there is $x \in a$. Since $x_0 \leq x$, it follows that $I = \{x_0\} \leq \Diamond a$.
($\Leftarrow$) If $x_0 \not\leq x$, pick $a \in A$ such that $x \in a$ and $\uparrow x_0 \cap a = \emptyset$. Then $a \neq \bot$, but $x_0 \notin \Diamond a$, so $I \not\leq \Diamond a$.

(e) Clear from the definitions. $\dashv$

We may use the representation theorem, Theorem 3.4, to prove the following proposition, which will be used in the proof of Lemma 3.8.

PROPOSITION 3.5. *Any $(\Diamond, \mathbf{X}, I)$-algebra validates the equations*

(Con) $\square(\square a \to b) \vee \square(\square b \to a) = \top$,
(Dum) $\square(\square(a \to \square a) \to a) \wedge \Diamond \square a \leq a$.

Moreover, the preorder on any $(\Diamond, \mathbf{X}, I)$-space is linear.

PROOF. Let A be a $(\Diamond, \mathbf{X}, I)$-algebra dual to a $(\Diamond, \mathbf{X}, I)$-space $(X, \leq, f, x_0)$.

(Con) Let $a, b \in A$ be arbitrary, and write $K := \Diamond \neg (\square a \to b) \wedge \Diamond \neg (\square b \to a)$, the complement of $\square(\square a \to b) \vee \square(\square b \to a)$. We need to show that $K = \emptyset$. We prove first that $f(K) \subseteq K$. If $x \in K$, pick $y \in \neg(\square a \to b) = \square a \wedge \neg b$ and $z \in \neg(\square b \to a) = \square b \wedge \neg a$ with $y, z \geq x$. Then $y \neq x$, since $y \in \square a$ but $z \notin a$, while $x \leq z$. So $f(x) \leq y$. Similarly, $f(x) \leq z$, so $f(x) \in K$. By Def. 3.3(ii), we obtain that $\uparrow K \subseteq K$. Now, if we would also have that $K \neq \emptyset$, then there would exist $x \in K$ and $y \geq x$ such that $y \in \square a \wedge \neg b$. But then $y \in \square(\square b \to a)$, so $y \notin K$, contradicting that $\uparrow K \subseteq K$. Therefore, we must have $K = \emptyset$, as required.

(Dum) Let $a \in A$ be arbitrary. We will prove $\square(\square(a \to \square a) \to a) \wedge \Diamond \square a \leq \square a$, from which (Dum) follows since $\square a \leq a$. The method is the same as before. Let us write $K := \square(\square(a \to \square a) \to a) \wedge \Diamond \square a \wedge \neg \square a$. We need to show that $K = \emptyset$. We prove first that $f(K) \subseteq K$. Let $x \in K$ be arbitrary. Pick $y \geq x$ such that $y \notin a$, and pick $z \geq x$ such that $z \in \square a$. Then $z \neq x$, so $f(x) \leq z$, and $f(x) \in \Diamond \square a$. Also, since $\uparrow(\square b) \subseteq \square b$ holds for any b, and $x \leq f(x)$, we have $f(x) \in \square(\square(a \to \square a) \to a)$. It remains to prove that $f(x) \in \neg \square a$. If $y \neq x$, then $f(x) \leq y$, and we are done. Otherwise, we have that $y = x$, so $x \notin a$. Since $x \in \square(a \to \square a) \to a$, pick $w \geq x$ such that $w \in a$ and $w \in \neg \square a$. Then $w \neq x$, so $f(x) \leq w$, so $f(x) \in \neg \square a$, as required. Again, by Def. 3.3(ii), we obtain $\uparrow K \subseteq K$, and if K were non-empty, we would have $x \in K$ and $z \geq x$ such that $z \in \square a$, but then $z \notin K$, contradiction. So $K = \emptyset$, as required.

Finally, since (Con) holds in A, the preorder $\leq$ is connected, i.e., for any $z, x, y \in X$, if $z \leq x$ and $z \leq y$, then $x \leq y$ or $y \leq x$ (see, e.g., [1, Ex. 4.3.3]). Since the $(\Diamond, \mathbf{X}, I)$-space $(X, \leq, f, x_0)$ has a minimum element, x_0, it follows that the preorder is linear. $\dashv$

We now turn to the filtration argument. The following syntactic lemma, which allows us to rewrite terms into negation normal form, will be useful. We call an $\mathcal{L}$-term a *literal* if it is either a variable, a constant ($\bot$ or I), a negated variable or a negated constant. An $\mathcal{L}$-term is in *negation normal form* (NNF) if it is built up from literals by repeated applications of $\vee, \wedge, \Diamond, \square$ and $\mathbf{X}$.

LEMMA 3.6. *For any $\mathcal{L}$-term t, there exists an $\mathcal{L}$-term t' in negative normal form such that $T \vdash t = t'$.*

PROOF. By induction on the complexity of the $\mathcal{L}$-term t, we can push any negation inwards, making use of the following T-provable equalities: $\neg(s_1 \vee s_2) = \neg s_1 \wedge \neg s_2$, $\neg(s_1 \wedge s_2) = \neg s_1 \vee \neg s_2$, $\neg \Diamond s = \Box \neg s$, $\neg \Box s = \Diamond \neg s$ and $\neg \mathbf{X} s = \mathbf{X} \neg s$. ⊣

DEFINITION 3.7. A finite set Γ of NNF $\mathcal{L}$-terms is *filterable* if

(i) Γ contains I,
(ii) Γ is closed under subterms,
(iii) whenever Γ contains $\Diamond s$ for some term s, Γ also contains $\mathbf{X} \Diamond s$,
(iv) whenever Γ contains $\Box s$ for some term s, Γ also contains $\mathbf{X} \Box s$.

LEMMA 3.8. *Let Γ be a filterable set, and denote by $v_1, \ldots, v_N$ the variables occurring in Γ. For any $(\Diamond, \mathbf{X}, I)$-space $(X, f, \leq_X, x_0)$ with dual algebra A and $\bar{a} \in A^N$, there exists $\bar{p} \in \mathcal{P}(\omega)^N$ with the following property: for any $x \in X$, there is $n_x \in \omega$ such that for all $s \in \Gamma$,*

$$\text{if } x \in s^A(\bar{a}) \text{ then } n_x \in s^{\mathcal{P}(\omega)}(\bar{p}). \tag{$\star$}$$

PROOF. Let $(X, f, \leq_X, x_0)$ be a $(\Diamond, \mathbf{X}, I)$-space and let $\bar{a} \in A^N$. Throughout this proof, if $x \in X$ and s is a term, we will write "$x \in s$" as shorthand for "$x \in s^A(\bar{a})$".

Let $\sim$ be the equivalence relation on X defined by

$$x \sim x' \iff \text{for all } s \in \Gamma : x \in s \text{ if, and only if } x' \in s.$$

Since Γ is finite, $X/\sim$ is finite. We write $Y := X/\sim$ and $q : X \twoheadrightarrow Y$ for the quotient map. Note that q is continuous with respect to the discrete topology on Y, since each equivalence class $[x]_\sim$ is clopen: it can be described by the formula $\bigwedge_{x \in \gamma \in \Gamma} \gamma \wedge \bigwedge_{x \notin \gamma \in \Gamma} \neg \gamma$.

We will define three relations, F, $\leq_Y$ and $\preceq$, on the quotient Y. First, define the relation F by:

$$yFy' \iff \text{there exists } x \in X \text{ such that } q(x) = y \text{ and } q(f(x)) = y',$$

that is, F is the smallest filtration of the (functional) relation f on X, cf., e.g., [1, Lem. 2.40]. Let $\leq_Y$ be the reflexive and transitive closure of F, i.e.,

$$y \leq_Y y' \iff \text{there exist } m \geq 0 \text{ and } z_0, \ldots, z_m \in Y \text{ such that}$$
$$z_0 = y, z_m = y', \text{ and } z_i F z_{i+1} \text{ for all } 0 \leq i < m.$$

Finally, we define a relation $\preceq$ on Y by

$$y \preceq y' \iff \text{for all } x \in X, \text{ if } q(x) = y, \text{ then there exists}$$
$$x' \in X \text{ such that } x \leq x' \text{ and } q(x) = y'.$$

In the following claim, we collect several properties of the three relations defined above. In particular, (i) & (ii) show that F is a filtration of f, (iii) & (iv) show that $\leq_Y$ is a filtration of $\leq_X$, (v)–(viii) provide detailed properties of the relation $\preceq$ that we need in our construction, and (ix) & (x) show that the properties of the minimum element x_0 are preserved in $(Y, \preceq)$.

Claim 1. For any $x, x' \in X$, $y, y' \in Y$, the following hold.

(i) $q(x)Fq(f(x))$.
(ii) whenever $\mathbf{X} t \in \Gamma$, if $q(x)Fq(x')$ then $x \in \mathbf{X} t$ if, and only if, $x' \in t$.

(iii) if $x \leq_X x'$, then $q(x) \leq_Y q(x')$.

(iv) whenever $\Box t \in \Gamma$, if $q(x) \leq_Y q(x')$ and $x \in \Box t$, then $x' \in t$.

 (v) $\preceq$ is a linear preorder.

(vi) if $y \preceq y'$, then $y \leq_Y y'$.

(vii) whenever $\Diamond t \in \Gamma$, if $x \in \Diamond t$ then there exists w such that $q(x) \preceq q(w)$ and $w \in t$.

(viii) if $y \preceq y'$ and $y' \preceq y$, then either $y = y'$, or $z \preceq y$ for all $z \in Y$.

 (ix) $q(x) = q(x_0)$ if, and only if, $x = x_0$.

 (x) $q(x_0) \preceq y$ and not $yFq(x_0)$.

PROOF OF CLAIM 1. (i) and (ii) are standard, cf. e.g. [1, Lem. 2.40].

(iii) Note that it suffices to prove that $K_y := \{x' \in X \mid y \leq_Y q(x')\}$ is an up-set in $\leq_X$ for any $y \in Y$. Notice that K_y is clopen, since q is continuous and Y is finite. To show that K_y is an up-set, by Def 3.3(ii), it suffices to prove that $f(K_y) \subseteq K_y$. Indeed, if $x' \in K_y$, then $y \leq_Y q(x')Fq(f(x'))$, and $\leq_Y$ is a transitive relation containing F, so $q(f(x')) \in K_y$.

(iv) Suppose that $\Box t \in \Gamma$ and $x \in \Box t$. By induction on m, using that $\mathbf{X}\Box t \in \Gamma$ since Γ is filterable, one can prove that if there exists an F-path of length m from $q(x)$ to $q(x')$, then $x' \in \Box t$. Now, if $q(x) \leq_Y q(x')$, then by definition of $\leq_Y$ there exists an F-path from $q(x)$ to $q(x')$, so $x' \in \Box t$. In particular, $x' \in t$.

(v) Reflexivity and transitivity are straight-forward to prove. For linearity, suppose that $y \npreceq y'$. Pick $x \in X$ such that $q(x) = y$ and for any x' such that $q(x') = y'$, we have $x \nleq_X x'$. Since $\leq_X$ is linear (Prop. 3.5), we have $x' \leq_X x$ for any x' such that $q(x') = y'$, so $y' \preceq y$.

(vi) Suppose that $y \preceq y'$. Pick $x \in X$ such that $q(x) = y$. By definition of $\preceq$, pick $x' \geq_X x$ such that $q(x') = y'$. By (iii), we have $y = q(x) \leq_Y q(x') = y'$.

(vii) Suppose that $\Diamond t \in \Gamma$ and $x \in \Diamond t$. Let $T := \{y \in Y \mid q(x) \npreceq y\}$. For each $y \in T$, we have $q(x) \npreceq y$, so pick $x_y \in X$ such that $q(x_y) = q(x)$ and, for any w such that $q(w) = y$, we have $x_y \nleq w$. Since $\leq_X$ is linear (Prop. 3.5), choose an enumeration $y_0, \dots, y_m$ of the elements of T such that $x_{y_0} \leq_X \cdots \leq_X x_{y_m}$. Since $x \in \Diamond t \in \Gamma$ and $x_{y_m} \sim x$, we have $x_{y_m} \in \Diamond t$. Pick $w \geq_X x_{y_m}$ such that $w \in t$. Note that $q(w) \notin T$: otherwise, we would have $q(w) = y_j$ for some $y_j \in T$, and $x_{y_j} \leq_X x_{y_m} \leq_X w$, contradicting the choice of x_{y_j}. Therefore, since $\preceq$ is linear by (v), we must have $q(x) \preceq q(w)$, as required.
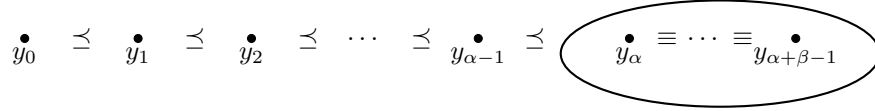
(viii) Suppose that $y \preceq y'$, $y \neq y'$ and that there exists $z \in Y$ such that $z \npreceq y$. We prove that $y' \npreceq y$. Let $a := \{v \in X \mid q(v) \neq y\}$, a clopen subset of X since q is continuous. Pick an element $x \in X$ such that $q(x) = y$, i.e., $x \notin a$. We note first that $x \in \Diamond \Box a$: since $z \npreceq y$, pick $w \in X$ such that $q(w) = z$ and, whenever $w \leq_X v$, $q(v) \neq y$. Now $w \in \Box a$, so $w \nleq_X x$, and hence $x \leq_X w$ since $\leq_X$ is linear. Now, since A verifies (Dum) (Prop. 3.5), $x \notin a$, and $x \in \Diamond \Box a$, we obtain $x \notin \Box(\Box(a \to \Box a) \to a)$. Pick $x_1 \geq_X x$ such that $x_1 \in \Box(a \to \Box a)$ and $x_1 \notin a$. By definition of a, we have $q(x_1) = y \preceq y'$, so we may pick $x_2 \geq_X x_1$ such that $q(x_2) = y'$. Since $y \neq y'$, we have $x_2 \in a$. Thus, since $x_1 \in \Box(a \to \Box a)$, and $x_1 \leq_X x_2$, we obtain $x_2 \in \Box a$. In particular, there is no $v \geq_X x_2$ such that $q(v) = y$, so $y' \npreceq y$, as required.

(ix) If $q(x) = q(x_0)$, then since $x_0 \in I$ and $I \in \Gamma$, we have $x \in I$. Therefore, $x = x_0$.

(x) By (ix) and the fact that $x_0 \leq_X x$ for any $x \in X$, it is clear that $q(x_0) \preceq y$ for any $y \in Y$. If we would have $yFq(x_0)$, then there would exist $x \in X$ such that $q(x) = y$ and $q(f(x)) = q(x_0)$. However, by (ix) we would get $f(x) = x_0$, contradicting Def. 3.3(iv). $\dashv$

We will write $\equiv$ for the equivalence relation induced by $\preceq$, i.e., $y \equiv y'$ if, and only if, $y \preceq y'$ and $y' \preceq y$. By Claim 1(viii), the preorder $(Y, \preceq)$ is a "balloon", i.e., there exist $\alpha \geq 0$, $\beta \geq 1$ and an enumeration $y_0, \ldots, y_\alpha, \ldots, y_{\alpha+\beta-1}$ of the elements of Y such that, for $0 \leq i \leq \alpha - 1$, $y_i \preceq y_{i+1}$ and $y_{i+1} \not\preceq y_i$, and for $0 \leq j, j' \leq \beta - 1$, $y_{\alpha+j} \equiv y_{\alpha+j'}$. In a picture:

$$\underset{y_0}{\bullet} \quad \preceq \quad \underset{y_1}{\bullet} \quad \preceq \quad \underset{y_2}{\bullet} \quad \preceq \quad \cdots \quad \preceq \quad \underset{y_{\alpha-1}}{\bullet} \quad \preceq \quad \overbrace{\underset{y_\alpha}{\bullet} \quad \equiv \cdots \equiv \quad \underset{y_{\alpha+\beta-1}}{\bullet}}$$

As a convenient notation, we extend this enumeration of the elements of Y to an infinite sequence, by defining $y_{\alpha+\gamma} := y_{\alpha+(\gamma \bmod \beta)}$ for any $\gamma \geq \beta$.

We now define a function $\sigma : \omega \to Y$ and a strictly increasing sequence of natural numbers $(\ell_i)_{i \in \omega}$ with the following properties:

(i) $\sigma(n)F\sigma(n+1)$ for all n,
(ii) $\sigma(\ell_i) = y_i$ for all i.

Let $\sigma(0) := y_0$ and $\ell_0 := 0$. Assume that ℓ_i and $\sigma(n)$ have been defined correctly for all $n \leq \ell_i$. Since $y_i \preceq y_{i+1}$, by Claim 1(vi), we have $y_i \leq_Y y_{i+1}$, so we may pick $m > 0$, $\sigma(\ell_i + 1), \ldots, \sigma(\ell_i + m) \in Y$ with $\sigma(\ell_i + m) = y_{i+1}$ and $\sigma(\ell_i + j)F\sigma(\ell_i + j + 1)$ for all $0 \leq j < m$. Define $\ell_{i+1} := \ell_i + m$. Note that we may indeed arrange the choice so that $m > 0$: if $y_i \neq y_{i+1}$, this is automatic, and if $y_i = y_{i+1}$ then $\beta = 1$ and $i \geq \alpha$; in this case first choose any $\sigma(\ell_i + 1) \in Y$ so that $y_i F\sigma(\ell_i + 1)$, and then choose an F-path from $\sigma(\ell_i + 1)$ to $y_i = y_{i+1}$, which can be done since $\sigma(\ell_i + 1) \preceq y_i$ so $\sigma(\ell_i + 1) \leq_Y y_i$ by Claim 1(vi).

Finally, we define the valuation $\bar{p} \in \mathcal{P}(\omega)^N$ by setting, for each $1 \leq j \leq N$,

$$p_j := \{n \in \omega \mid \sigma(n) = q(x) \text{ for some } x \in a_j\}.$$

The fact that $(\star)$ in the statement of Lemma 3.8 is true for this choice of $\bar{p}$ will follow from the following claim and the fact that σ is surjective, by property (ii) of σ.

Claim 2. For any $s \in \Gamma$,

for any $x \in X, n \in \omega$, if $q(x) = \sigma(n)$ and $x \in s^A(\bar{a})$, then $n \in s^{\mathcal{P}(\omega)}(\bar{p})$. $\quad (P_s)$

Proof of Claim 2. By induction on s, which is in negation normal form by assumption.

($s = v_j$.) If $q(x) = \sigma(n)$ and $x \in a_j$ then $n \in p_j$ by definition.

($s = \neg v_j$.) If $q(x) = \sigma(n)$ and $x \in \neg v_j$, then $x \notin a_j$. Since $v_j \in \Gamma$, for any x' with $\sigma(n) = q(x')$, we have $x' \sim x$, so $x' \notin a_j$. Hence, $n \notin p_j$, so $n \in \neg p_j$.

($s = I$.) If $q(x) = \sigma(n)$ and $x \in I$ then $x = x_0$. Then $\sigma(n) = q(x_0) = y_0$, since $q(x_0) \preceq q(x)$ for all x by Claim 1(x). If we would have $n > 0$, then we would get $\sigma(n-1)Fy_0$, which is impossible by Claim 1(x). So $n = 0 \in I$.

($s = \neg I$.) If $q(x) = \sigma(n)$ and $x \in \neg I$ then $x \neq x_0$. By Claim 1(ix), $q(x) \neq y_0$. Thus, $\sigma(n) \neq y_0$, so in particular $n \neq 0$, so $n \in \neg I$.

$(s = s_1 \vee s_2)$ and $(s = s_1 \wedge s_2)$ are straight-forward.

$(s = \mathbf{X}t.)$ Suppose that $q(x) = \sigma(n)$ and $x \in \mathbf{X}t$. Since $\sigma(n)F\sigma(n+1)$, pick x' such that $q(x') = \sigma(n) = q(x)$ and $q(f(x')) = \sigma(n+1)$. Since $x \in \mathbf{X}t$, $\mathbf{X}t \in \Gamma$, and $q(x) = q(x')$, we have $x' \in \mathbf{X}t$, so $f(x') \in t$. By the induction hypothesis (P_t), we get $n+1 \in t$, so $n \in \mathbf{X}t = s$.

$(s = \Box t.)$ Suppose that $q(x) = \sigma(n)$ and $x \in \Box t$. Let $m \geq n$ be arbitrary, and pick x' such that $q(x') = \sigma(m)$. It follows from property (i) of σ that $q(x) = \sigma(n) \leq_Y \sigma(m) = q(x')$. By Claim 1(iv), we have $x' \in t$. Applying the induction hypothesis (P_t) to x' and m, we obtain $m \in t$. Since m was arbitrary, we conclude that $n \in \Box t$.

$(s = \Diamond t.)$ Suppose that $q(x) = \sigma(n)$ and $x \in \Diamond t$. We need to prove that $n \in \Diamond t$, i.e., that there exists $k \geq 0$ such that $n + k \in t$. We first prove that, *for any k, if $n + i \notin t$ for all $0 \leq i \leq k$, then $q^{-1}(\sigma(n+k)) \subseteq \Diamond t$.* For $k = 0$, this is clear, because $x \in \Diamond t \in \Gamma$. For the induction step, if the statement holds for some k, assume that $q(x') = \sigma(n+k+1)$ and $n + i \notin t$ for all $0 \leq i \leq k+1$. Pick some x'' such that $q(x'') = \sigma(n+k)$. Then $x'' \in \Diamond t$ by the statement for k, and $x'' \notin t$, for otherwise (P_t) would give $n + k \in t$. Therefore, by Def. 2.1(i), $x'' \in \mathbf{X}\Diamond t$, and $\mathbf{X}\Diamond t \in \Gamma$ since Γ is filterable. By Claim 1(ii) and the fact that $q(x'') = \sigma(n+k)F\sigma(n+k+1) = q(x')$, we obtain $x' \in \Diamond t$.

Now, since σ visits y_α infinitely often, pick $\ell \geq 0$ such that $\sigma(n+\ell) = y_\alpha$. If $n + i \in t$ for some $0 \leq i \leq \ell$, then we are done. Otherwise, pick some $u \in X$ such that $q(u) = y_\alpha$; by the previous paragraph, $u \in \Diamond t$. By Claim 1(vii), pick $w \in X$ such that $y_\alpha \preceq q(w)$ and $w \in t$. Then $q(w) = y_{\alpha+j}$ for some $0 \leq j \leq \beta - 1$. Since σ visits $y_{\alpha+j}$ infinitely often, pick $k \geq \ell$ such that $\sigma(n+k) = y_{\alpha+j} = q(w)$. Then $n + k \in t$ by (P_t). $\dashv$

We are now in a position to prove Theorem 3.2, which, by the remarks preceding it, finishes the proof that T^* is a companion of T.

PROOF OF THEOREM 3.2. Let $t(v_1, \ldots, v_N)$ be any $\mathcal{L}$-term, and assume that $\mathcal{P}(\omega) \models t = \top$. By Lemma 3.6, let u be an NNF $\mathcal{L}$-term that is equivalent to $\neg t$. Let $\mathrm{Sub}(u)$ denote the set of subterms of the term u, and let Γ be the filterable set

$$\Gamma := \{s \mid s \in \mathrm{Sub}(u)\} \cup \{\mathbf{X}\Diamond s \mid \Diamond s \in \mathrm{Sub}(u)\} \cup \{\mathbf{X}\Box s \mid \Box s \in \mathrm{Sub}(u)\} \cup \{I\}.$$

Indeed, Γ is finite, and any element of Γ is in negation normal form.

Let A be a $(\Diamond, \mathbf{X}, I)$-algebra and let $\bar{a} \in A^N$ be arbitrary; we prove that $t^A(\bar{a}) = \top^A$. By Theorem 3.4, there exists a $(\Diamond, \mathbf{X}, I)$-space $(X, \leq, f, x_0)$ with dual algebra (isomorphic to) A. Choose $\bar{p} \in \mathcal{P}(\omega)^n$ with property $(\star)$ of Lemma 3.8. Let $x \in X$ be arbitrary. Pick $n_x \in \omega$ as in Lemma 3.8. Then $n_x \in t^{\mathcal{P}(\omega)}(\bar{p})$, since $\mathcal{P}(\omega) \models t = \top$, so $n_x \notin (\neg t)^{\mathcal{P}(\omega)}(\bar{p}) = u^{\mathcal{P}(\omega)}(\bar{p})$. Hence, since $u \in \Gamma$, we also have $x \notin u^A(\bar{a}) = (\neg t)^A(\bar{a})$, by the property $(\star)$ in Lemma 3.8. Therefore, $x \in t^A(\bar{A})$. Since x was arbitrary, it follows that $t^A(\bar{a}) = X = \top^A$. $\dashv$

§4. T^* **is model complete.** Let φ be a first-order formula in the signature of $(\Diamond, \mathbf{X}, I)$-algebras. We need to show that φ is equivalent to an existential formula φ' in T^*, the first-order theory of $\mathcal{P}(\omega)$. We will proceed according to the following scheme.

(i) Syntactically transform φ into a formula $\widehat{\varphi}$ of the monadic second-order logic S1S (Proposition 4.2).
(ii) Associate to the S1S-formula $\widehat{\varphi}$ a Büchi automaton $\mathcal{A}_{\widehat{\varphi}}$ (Theorem 4.5).
(iii) Associate to the Büchi automaton $\mathcal{A}_{\widehat{\varphi}}$ a $(\Diamond, \mathbf{X}, I)$-term $t_{\mathcal{A}_{\widehat{\varphi}}}$ representing it (Proposition 4.7).
(iv) Use an appropriate renaming of variables to obtain an existential formula φ' equivalent to φ (Subsection 4.4).

We now perform each of the steps in Subsections 4.1–4.4 below.

4.1. From first-order to S1S. Let us recall a definition of the syntax and semantics of the monadic second-order logic S1S (see, e.g., [11, Ch. 12] for more details and background).

DEFINITION 4.1. The set **S1S**$(\mathcal{V})$ of *S1S-formulae* with free variables contained in a set $\mathcal{V}$ is the smallest set such that:

(i) for each $P, Q \in \mathcal{V}$, the formulae $P \subseteq Q$ and $S(P, Q)$ are in **S1S**$(\mathcal{V})$,
(ii) if φ and ψ are in **S1S**$(\mathcal{V})$, then $\varphi \vee \psi$ and $\neg\varphi$ are in **S1S**$(\mathcal{V})$,
(iii) if φ is in **S1S**$(\mathcal{V})$ and R is in $\mathcal{V}$ then $\exists R\varphi$ is in **S1S**$(\mathcal{V} \setminus \{R\})$.

The *satisfaction relation* $\models \subseteq \mathcal{P}(\omega)^{\mathcal{V}} \times$ **S1S**$(\mathcal{V})$ is defined inductively by

(i) $v \models P \subseteq Q$ if, and only if, $v(P) \subseteq v(Q)$,
    $v \models S(P, Q)$ if, and only if, there exists $n \in P$ such that $n + 1 \in Q$,
(ii) $v \models \varphi \vee \psi$ if, and only if, $v \models \varphi$ or $v \models \psi$,
    $v \models \neg\varphi$ if, and only if, $v \models \varphi$ does not hold,
(iii) $v \models \exists R\varphi$ if, and only if, there exists a variant v' of v, differing from v only for the value of R, such that $v' \models \varphi$.

Recall [11, Def. 12.5] that a quantification over an individual variable, $\exists x\varphi$, can be encoded in S1S as a quantification $\exists x(\mathsf{sing}(x) \wedge \varphi)$, where $\mathsf{sing}(x)$ is an S1S-formula expressing that $v(x)$ must be a singleton for any valuation v. We use the convention that lowercase letters x, y, ... are individual variables, while capital letters P, Q, R, ... are set variables. We will also make use of the standard abbreviations $x \in P$, $P = Q$, $x \leq y$, $S(x, y)$, as in [11, Ch. 12].

Notation. In order to avoid confusion, we need to distinguish the function symbols $\vee$, $\neg$, $\perp$ in the first-order signature $\mathcal{L}$ of $(\Diamond, \mathbf{X}, I)$-algebras from the symbols $\vee$, $\neg$, $\perp$ that occur as connectives in first-order $\mathcal{L}$-formulae and in S1S-formulae. Therefore, throughout this section, we use an alternative first-order signature $\mathcal{L}^* := \{\cup, -, \emptyset, \Diamond, \mathbf{X}, I\}$ for $(\Diamond, \mathbf{X}, I)$-algebras, as well as abbreviations $a \cap b := -(-a \cup -b)$ and $a \Rightarrow b := -a \cup b$.

We now come to the translation from first-order $\mathcal{L}^*$-formulas to formulas in the logic S1S. This translation is a variant of the so-called *standard translation* of modal logic into monadic second-order logic, cf., e.g., [1, Prop. 3.12].

PROPOSITION 4.2. *For any first-order $\mathcal{L}^*$-formula $\varphi(p_1, \ldots, p_n)$, there exists an S1S-formula $\widehat{\varphi}(P_1, \ldots, P_n)$ such that, for any $v \in \mathcal{P}(\omega)^n$,[3]*

$$\mathcal{P}(\omega), v \models \varphi \text{ if, and only if, } v \models \widehat{\varphi}. \tag{1}$$

PROOF. We first inductively define, for any $\mathcal{L}^*$-term $t(p_1, \ldots, p_n)$, an S1S-formula $t^\bullet(P_1, \ldots, P_n, x)$, where x is a fresh individual variable, as follows:

- $(p_i)^\bullet := x \in P_i$,
- $(t \cup u)^\bullet := t^\bullet(x) \vee u^\bullet(x)$,
- $(-t)^\bullet := \neg(t^\bullet)(x)$,
- $\emptyset^\bullet := \neg(x = x)$,
- $(\Diamond t)^\bullet := \exists y(x \leq y \wedge t^\bullet(P_1, \ldots, P_n, y))$,
- $(\mathbf{X}t)^\bullet := \exists y(S(x, y) \wedge t^\bullet(P_1, \ldots, P_n, y))$,
- $I^\bullet := \forall z(x \leq z)$.

Note that, for any $i \in \omega$ and $v \in \mathcal{P}(\omega)^n$, we have

$$i \in t^{\mathcal{P}(\omega)}(v(1), \ldots, v(n)) \text{ if, and only if, } v \models t^\bullet(v(1), \ldots, v(n), i). \tag{2}$$

Now, for any $\mathcal{L}^*$-formula $\varphi(p_1, \ldots, p_n)$, define $\widehat{\varphi}(P_1, \ldots, P_n)$ by replacing any atomic formula $t = u$ by $\forall x(t^\bullet(P_1, \ldots, P_n, x) \leftrightarrow u^\bullet(P_1, \ldots, P_n, x))$, and any occurrence of $\exists p$ by $\exists P$. The claimed equivalence is proved by an easy induction on the complexity of φ, using (2) for the base case. $\dashv$

4.2. From S1S-formula to Büchi automaton. We briefly recall the definition of finite automata and the Büchi acceptance condition. See, e.g., [11, Ch. 1] for more details.

DEFINITION 4.3. Let Σ be a finite alphabet. A (finite, non-deterministic) *automaton* is a tuple $\mathcal{A} = (Q, \delta, q_0, F)$ where Q is a finite set, whose elements are called *states*, $q_0 \in Q$ is a distinguished element called *the initial state*, $F \subseteq Q$ is a set, whose elements are called *final states*, and $\delta \subseteq Q \times \Sigma \times Q$ is a ternary relation, called the *transition relation*. For an infinite word $w \in \Sigma^\omega$, a function $\rho \colon \omega \to Q$ is called a *run of $\mathcal{A}$ on w* if $\rho(0) = q_0$ and, for any $i \in \omega$, $(\rho(i), w(i), \rho(i+1)) \in \delta$. For any word $w \in \Sigma^\omega$, the automaton $\mathcal{A}$ *Büchi-accepts* w if, and only if, there exist a run ρ of $\mathcal{A}$ on w and a final state $q_f \in F$ such that the set $\{i \in \omega \mid \rho(i) = q_f\}$ is infinite.

We next recall how valuations of variables naturally define infinite words.

DEFINITION 4.4. Let $n \geq 0$ and $\Sigma := \mathcal{P}(\{1, \ldots, n\})$. For any $v \in \mathcal{P}(\omega)^n$, we define the infinite word $w_v \in \Sigma^\omega$ by

$$w_v(i) := v^{-1}(\{u \in \Sigma \mid i \in u\}) = \{k \mid i \in v(k)\}, \quad i \in \omega.$$

The following theorem, originally due to Büchi [2], states that any S1S-formula can be converted into an automaton.

---

[3]Note that the symbol $\models$ is used with two different meanings in (1): on the left-hand-side, it denotes the usual satisfaction relation of first-order logic, while on the right-hand-side it denotes the satisfaction relation of S1S of Definition 4.1. Similarly, the tuple $v \in \mathcal{P}(\omega)^n$ is regarded on the left as a valuation of the first-order variables $p_1, \ldots, p_n$, and on the right as a valuation of the second-order variables $P_1, \ldots, P_n$.

THEOREM 4.5. *For any S1S-formula $\psi(P_1, \ldots, P_n)$, there exists an automaton $\mathcal{A}_\psi$ on the finite alphabet $\Sigma := \mathcal{P}(\{1, \ldots, n\})$ such that, for any $v \in \mathcal{P}(\omega)^n$,*

$$v \models \psi \text{ if, and only if, } \mathcal{A}_\psi \text{ Büchi-accepts } w_v, \tag{3}$$

PROOF. See, e.g., [11, Thm. 12.15] or [15, Thm. 5.9].   ⊣

4.3. From automata to existential formulae. It is well-known that a Büchi automaton can be transformed into a formula of S1S which starts with a block of existential monadic set quantifiers, but also has some quantifications over individual variables after that. As we remarked in the introduction to the paper, we require slightly more, namely that the quantifications over individual variables are not needed in the signature $\mathcal{L}^*$.

In order to state precisely the translation back from a Büchi automaton to an existential $\mathcal{L}^*$-formula, we need to recall how infinite words in Σ^ω yield valuations of Σ in $\mathcal{P}(\omega)$, by the reverse process to Definition 4.4.

DEFINITION 4.6. Let Σ be a finite alphabet. For any $w \in \Sigma^\omega$, we define the valuation $v_w : \Sigma \to \mathcal{P}(\omega)$ by

$$v_w(a) := w^{-1}(a) = \{i \in \omega : w(i) = a\}, \quad a \in \Sigma.$$

Let $\mathcal{A} = (Q, \delta, q_0, F)$ be an automaton on a finite alphabet Σ. We now define a number of $\mathcal{L}^*$-terms with variables from the set $Q \cup \Sigma$. First define the $\mathcal{L}^*$-terms:

- $\mathsf{Init} := I \Rightarrow q_0$,
- $\mathsf{Trans} := \bigcap_{q \in Q} \left(q \Rightarrow \bigcup_{(q,a,q') \in \delta} (a \cap \mathbf{X}q') \right)$,
- $\mathsf{Part} := \bigcup_{q \in Q} (q \cap \bigcap_{\substack{q' \in Q \\ q' \neq q}} -q')$,
- $\mathsf{Accept} := \bigcup_{q \in F} \Diamond q$,

and define the $\mathcal{L}^*$-term $t_\mathcal{A}$ by:

$$t_\mathcal{A}(\overline{a}, \overline{q}) := \mathsf{Part} \cap \mathsf{Init} \cap \mathsf{Trans} \cap \mathsf{Accept}. \tag{4}$$

PROPOSITION 4.7. *Let $\mathcal{A}$ be an automaton on a finite alphabet Σ. For any $w \in \Sigma^\omega$,*

$$\mathcal{A} \text{ Büchi-accepts } w \text{ if, and only if, } \mathcal{P}(\omega), v_w \models \exists \overline{q} \, (t_\mathcal{A} = \top). \tag{5}$$

PROOF. Clear from the definitions.   ⊣

4.4. Proof that T^* is model complete. Let $\varphi(p_1, \ldots, p_n)$ be a first-order $\mathcal{L}^*$-formula. Let $\widehat{\varphi}(P_1, \ldots, P_n)$ be the S1S-formula given by Proposition 4.2. Let $\mathcal{A}_{\widehat{\varphi}}$ be the automaton on the alphabet $\Sigma := \mathcal{P}(1, \ldots, n) = \{a_1, \ldots, a_{2^n}\}$ given by Theorem 4.5. Let $t_{\mathcal{A}_{\widehat{\varphi}}}$ be the $\mathcal{L}^*$-term defined in (4). We define the existential $\mathcal{L}^*$-formula

$$\varphi'(p_1, \ldots, p_n) := \exists \overline{a} \, \exists \overline{q} \left((t_{\mathcal{A}_{\widehat{\varphi}}} = \top) \wedge \bigwedge_{\ell=1}^{2^n} \left(a_\ell = \bigcap_{k \in a_\ell} p_k \cap \bigcap_{k \notin a_\ell} -p_k \right) \right).$$

Claim. $T^* \vdash \varphi \leftrightarrow \varphi'$.

PROOF. It suffices to prove that $\mathcal{P}(\omega), v \models \varphi$ if, and only if, $\mathcal{P}(\omega), v \models \varphi'$, for any $v \in \mathcal{P}(\omega)^n$. Let $v \in \mathcal{P}(\omega)^n$ be arbitrary. We obtain:

$$
\begin{aligned}
\mathcal{P}(\omega), v \models \varphi \;\Longleftrightarrow\; & v \models \widehat{\varphi} && \text{(by (1) in Prop. 4.2)} \\
\Longleftrightarrow\; & \mathcal{A}_{\widehat{\varphi}} \text{ accepts } w_v && \text{(by (3) in Thm. 4.5)} \\
\Longleftrightarrow\; & \mathcal{P}(\omega), v_{w_v} \models \exists \overline{q}\,(t_{\mathcal{A}_{\widehat{\varphi}}} = \top) && \text{(by (5) in Prop. 4.7)} \\
\Longleftrightarrow\; & \mathcal{P}(\omega), v \models \varphi',
\end{aligned}
$$

where the last equivalence follows from the fact that, for each $a \in \Sigma = \mathcal{P}(1, \ldots, n)$, we have

$$
v_{w_v}(a) = \{i \in \omega : a = \{k : i \in v(k)\}\} = \bigcap_{k \in a} v(k) \cap \bigcap_{k \notin a} -v(k).
$$

$\dashv$

§5. Conclusions and future work. In this paper, we characterized monadic second order logic on infinite words as the model companion of the universal class of $(\Diamond, \mathbf{X}, I)$-algebras. This is not a stand-alone result; we indicate a few of the possible directions of further research here.

First, we expect that it is possible to give a similar characterization of monadic second order logic on finite words, by adjusting the axiomatization of $(\Diamond, \mathbf{X}, I)$-algebras in the appropriate manner. A harder, but equally natural question, is whether a version of our result holds for monadic second order logics on tree structures, such as S2S, or finite trees.

Our result (and/or its incarnation for finite words) is likely to be related to the duality theory for regular languages and logics that is being developed in a series of papers including [6, 7]. It would be interesting to make these connections explicit, in order to determine if the two approaches can benefit from each other.

Riba [13] gives a model-theoretic proof of the completeness of Siefkes' axiomatization of S1S. Our result in this paper entails that S1S, viewed as a first-order theory, coincides with the theory of the existentially closed $(\Diamond, \mathbf{X}, I)$-algebras [16, Thm. B]. Therefore, we suggest that an alternative axiomatization of S1S, and completeness proof for it, could be sought by axiomatizing existentially closed $(\Diamond, \mathbf{X}, I)$-algebras. In a similar direction, we note that Gheerbrant and Ten Cate [8] used modal logic techniques to axiomatize monadic second order logic on finite trees, and fragments. An extension of our results in this paper to finite trees could also be connected to the results in [8].

REFERENCES

[1] P. BLACKBURN, M. DE RIJKE, and Y. VENEMA, *Modal logic*, Cambridge Tracts in Theoretical Computer Science, vol. 53, Cambridge University Press, 2001.

[2] J. R. BÜCHI, *On a decision method in restricted second-order arithmetic*, **Proceedings 1960 International Congress for Logic, Methodology and Philosophy of Science** (E. Nagel, P. Suppes, and A. Tarski, editors), Stanford University Press, 1962, pp. 1–11.

[3] L. ESAKIA, *Topological Kripke models*, **Soviet Mathematics Doklady**, vol. 15 (1974), pp. 147–151.

[4] K. FINE, *Some connections between elementary and modal logic*, **Proceedings of the Third Scandinavian Logic Symposium**, North-Holland, 1975, pp. 15–31.

[5] M. GEHRKE, *Canonical extensions, Esakia spaces, and universal models*, **Leo Esakia on Duality in Modal and Intuitionistic Logics** (G. Bezhanishvili, editor), Springer, Dordrecht, 2014, pp. 9–42.

[6] M. GEHRKE, S. GRIGORIEFF, and J.-É. PIN, *Duality and equational theory of regular languages*, **35th International Colloquium on Automata, Languages and Programming (ICALP 2008)** (L. Aceto et al., editor), Lecture Notes in Computer Science, no. 5126, Springer, 2008, pp. 246–257.

[7] M. GEHRKE, A. KREBS, and J.-É. PIN, *From ultrafilters on words to the expressive power of a fragment of logic*, **Descriptional Complexity of Formal Systems (DCFS 2014)**, Lecture Notes in Computer Science, vol. 8614, Springer, 2014, pp. 138–149.

[8] A. GHEERBRANT and B. TEN CATE, *Complete axiomatizations of fragments of monadic second-order logic on finite trees*, **Logical Methods in Computer Science**, vol. 8 (2012), no. 4:12.

[9] R. GOLDBLATT, *Varieties of complex algebras*, **Annals of Pure and Applied Logic**, vol. 44 (1989), pp. 173–242.

[10] R. GOLDBLATT, **Logics of time and computation**, 2nd ed., CSLI Lecture Notes, no. 7, Stanford University, 1992.

[11] E. Grädel, W. Thomas, and T. Wilke (editors), *Automata logics, and infinite games: A guide to current research*, Lecture Notes in Computer Science, vol. 2500, Springer, 2002.

[12] A. PNUELI, *The temporal logic of programs*, **Proceedings of the 18th Annual Symposium on Foundations of Computer Science (FOCS)**, 1977, pp. 46–57.

[13] C. RIBA, *A model theoretic proof of completeness of an axiomatization of monadic second-order logic on infinite words*, **Theoretical Computer Science 2012**, Lecture Notes in Computer Science, vol. 7604, Springer, 2012, pp. 310–324.

[14] A. ROBINSON, **Introduction to model theory and to the metamathematics of algebra**, Studies in logic and the foundations of mathematics, North-Holland, 1963.

[15] W. THOMAS, *Languages, automata, and logic*, **Technical Report 9607**, Institut für Informatik und Praktische Mathematik, Christian-Albrechts-Universität, Kiel, 1997.

[16] W. H. WHEELER, *Model-companions and definability in existentially complete structures*, **Israel Journal of Mathematics**, vol. 25 (1976), pp. 305–330.