Extended ω-Regular Languages and Interval Temporal Logic[⋆] [⋆⋆]

Dario Della Monica[1], Angelo Montanari[1], and Pietro Sala[2]

[1] University of Udine, Italy {dario.dellamonica|angelo.montanari}@uniud.it
[2] University of Verona, Italy pietro.sala@univr.it

Abstract. Some extensions of ω-regular languages have been proposed in the literature to express asymptotic properties of ω-words which are not captured by ω-regular languages. Formal definitions of extended ω-regular languages have been given in terms of both suitable classes of automata and extended ω-regular expressions. On the contrary, satisfactory temporal logic counterparts are still missing. In this paper, we give a characterization of them in terms of interval temporal logics.

1 Introduction

In this paper, we explore the relationships between extended ω-regular languages and temporal logic by providing a characterization of language expressions in terms of formulas of suitable interval temporal logics.

ω-regular languages are a natural setting for the specification and verification of nonterminating finite-state systems. Since the seminal work by Büchi, McNaughton, and Rabin in the sixties, much has been done on the theory and the applications of ω-regular languages. Equivalent characterizations of ω-regular languages have been given in terms of formal languages, automata, classical and temporal logic. However, while the consensus on what features regular languages of finite words must exhibit is unanimous (it largely relies on Myhill-Nerode theorem), the notion of ω-regular languages is more controversial. In the last years, it has been shown that ω-languages can be extended in meaningful ways, preserving their decidability and (some of their) closure properties [3–6]. The proposed extensions pair the Kleene star $(.)^*$ with some variants of it. The bounding exponent B of ωB-regular languages, denoted by $(.)^B$, constrains the language L in the expression L^B to be iterated only a bounded number of times, the bound being fixed for the whole ω-word [6].[3] The unbounding exponent S of ωS-regular languages, denoted by $(.)^S$, when applied to a language L, forces

⋆ Partially supported by the GNCS 2020 project "Ragionamento Strategico e Sintesi Automatica di Sistemi Multi-Agente".

[3] We are obviously assuming that there are no other occurrences of L in the ωB-regular expression. The same holds for the other extended ω-regular languages.

the number of iterations of L to tend to infinity, i.e., for every $k > 0$, it constrains the number of times the argument L is repeated at most k times to be finite [6]. $(.)^B$ and $(.)^S$ can be freely mixed in ωBS-regular languages (the combination of ωB- and ωS-regular ones) [6]. ωB- and ωS-regular languages are properly included in ωBS-regular ones [5], as witnessed by the ωBS-regular language $\mathcal{L} = (a^B b + a^S b)^\omega$ consisting of those ω-words w featuring infinitely many occurrences of b and such that there are only finitely many numbers occurring infinitely often in the sequence of exponents of a in w. The existence of non-ωBS-regular languages that are the complements of some ωBS-regular ones and express natural asymptotic behaviours motivated the search for other classes of extended ω-regular languages. In [2], ωT-regular languages, which are based on a different extension of $(.)^*$, denoted by $(.)^T$, and include meaningful non-ωBS-regular languages, like, e.g., the complement of $\mathcal{L}$ above, have been studied. Besides those in terms of ωB-, ωS-, ωBS-, and ωT-regular expressions, equivalent characterizations of the above languages have been given in terms of automata and classical logic (extensions of the monadic second-order theory of one successor S1S). Temporal logic counterparts are still missing. As a matter of fact, interval temporal logic counterparts of ωB- and ωS-regular languages were proposed in [15] and [14], respectively. Unfortunately, both of them are flawed. Here, we provide a fix, and, in addition, give an interval temporal logic characterization of ωT-regular languages.

Interval temporal logic (ITL) is a general framework for representing and reasoning about time. ITLs are characterized by high expressiveness (they overcome various limitations of point-based temporal logics) and high computational complexity (formulas translate into binary relations over the underlying linear order). One of the first ITLs proposed in the literature is Moszkowski's Propositional ITL (PITL), which was successfully applied to hardware specification and verification [17]. The application of interval-based formalisms to temporal reasoning in AI was first investigated by Allen [1]. A systematic logical study of interval reasoning started with Halpern and Shoham's work on the logic HS featuring one modality for each Allen relation [10]. While decidability is a common feature of point-based temporal logics, undecidability rules over ITLs. The first such undecidability results were obtained for PITL by Moszkowski [16]. General undecidability results for HS are given in [10] and further sharpened in [11]. For a long time, these results have discouraged the search for practical applications and further theoretical investigation on ITLs. This bleak picture started lightening up in the last few years when various non-trivial decidable fragments of HS have been identified (see, e.g., [7, 8]). In this paper, we focus on the interval logic AB, whose modalities correspond to Allen's relations *meets* (modality $\langle A \rangle$) and *begun by* (modality $\langle B \rangle$) [13], and some extensions of it with modalities for the inverse relations *met by* (modality $\langle \bar{A} \rangle$) and *begins* (modality $\langle \bar{B} \rangle$). In [15], Montanari and Sala have proved that regular (resp., ω-regular) languages can be defined in $AB\bar{B}$, interpreted over finite linear orders (resp., $\mathbb{N}$).[4] Here, we show that extended ω-regular languages can be captured by suitable extensions of AB.

[4] In fact, $\langle \bar{B} \rangle$ simplifies the encoding, but it is not necessary; thus, we do not use it.

In particular, we show that (i) ωB-regular languages can be expressed in $AB\bar{A}$, that extends AB with the past modality $\langle \bar{A} \rangle$ corresponding to Allen's relation *met by* [12], (ii) ωS-regular languages can be encoded in AB enriched with an equivalence relation $\sim$ ($AB\sim$) [14], and (iii) ωT-regular languages are captured by $AB\bar{A}\sim$. A distinctive feature of the encodings is that they do not resort to any counter, that is, checking the satisfaction of boundedness/unboundedness conditions in interval temporal logic does not require the precision in length measurements given by counters (in fact, some abstraction over counters, that allows one to consider orders of magnitude rather than exact values, is applied also in the automaton-based characterizations of extended ω-regular languages).

The paper is organized as follows. First, we provide some background knowledge. Then, we enrich the encoding of ω-regular languages into AB given in [15] to capture the increased expressive power of extended ω-regular languages.

2 Preliminaries

Extended ω-regular languages. We give a short account of extended ω-regular languages in terms of the extended ω-regular expressions that define them. For a detailed one, we refer the reader to [2]. Extended ω-regular expressions are built on top of the corresponding extended regular ones, just as ω-regular expressions are built on top of regular ones. Let Σ be a finite, nonempty alphabet. An extended *regular expression* over Σ is defined by (a subset of) the grammar: $e ::= \emptyset \mid a \mid e \cdot e \mid e + e \mid e^* \mid e^B \mid e^S \mid e^T$, where $a \in \Sigma$.

Extended regular expressions differ from regular ones as they allow constructors from the set $\{(.)^B, (.)^S, (.)^T\}$. Their semantics is given in terms of languages of infinite sequences of finite words, by imposing suitable constraints that capture the intended meaning of $(.)^B$, $(.)^S$, and $(.)^T$.

Let $\mathbb{N}$ be the set of natural numbers and $\mathbb{N}^+ = \mathbb{N} \setminus \{0\}$. For an infinite sequence $\boldsymbol{u}$ of finite words over Σ and $i \in \mathbb{N}^+$, we denote by u_i its i-th element. The basic shuffle operation is defined as follows [2]. Let $\boldsymbol{v}^1 = (v^1_1, v^1_2, \ldots)$ and $\boldsymbol{v}^2 = (v^2_1, v^2_2, \ldots)$ be two infinite word sequences, and let $g : \mathbb{N}^+ \rightarrow \{1, 2\}$ be a selection function. We define the g-shuffle of $\boldsymbol{v}^1$ and $\boldsymbol{v}^2$ as the word $\boldsymbol{v} = (v_1, v_2, \ldots)$, where $v_i = v^{g(i)}_{|\{j \in \mathbb{N}^+ | j \leq i \text{ and } g(j)=g(i)\}|}$ for all $i \in \mathbb{N}^+$. We say that an infinite word sequence $\boldsymbol{v}$ is a *shuffle* of $\boldsymbol{v}^1$ and $\boldsymbol{v}^2$ if there is a selection function g such that $\boldsymbol{v}$ is the g-shuffle of them. Notice that the set of selection functions includes those g where there exists $k \in \mathbb{N}^+$ such that $g(x) = 1$ (resp., $g(x) = 2$), for all $x > k$.

The semantics of extended regular expressions over Σ is defined as follows:

- $\mathcal{L}(\emptyset) = \emptyset$;
- for $a \in \Sigma$, $\mathcal{L}(a)$ only contains the infinite sequence of the one-letter word a, that is, $\mathcal{L}(a) = \{(a, a, a, \ldots)\}$;
- $\mathcal{L}(e_1 \cdot e_2) = \{\boldsymbol{w} \mid \forall i. w_i = u_i \cdot v_i,\ \boldsymbol{u} \in \mathcal{L}(e_1),\ \boldsymbol{v} \in \mathcal{L}(e_2)\}$;
- $\mathcal{L}(e_1 + e_2) = \{\boldsymbol{w} \mid \boldsymbol{w}$ is a shuffle of $\boldsymbol{u}$ and $\boldsymbol{v}$, for some $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{L}(e_1) \cup \mathcal{L}(e_2)\}$;
- $\mathcal{L}(e^*) = \{(u_{f(0)} u_2 \ldots u_{f(1)-1}, u_{f(1)} \ldots u_{f(2)-1}, \ldots) \mid \boldsymbol{u} \in \mathcal{L}(e)$ and $f : \mathbb{N} \rightarrow \mathbb{N}^+$ is a nondecreasing function with $f(0) = 1\}$;

- $\mathcal{L}(e^B) = \{(u_{f(0)}u_2 \dots u_{f(1)-1}, u_{f(1)} \dots u_{f(2)-1}, \dots) \mid \boldsymbol{u} \in \mathcal{L}(e)$ and $f : \mathbb{N} \to \mathbb{N}^+$ is a nondecreasing function, with $f(0) = 1$, such that $\exists n \in \mathbb{N}\ \forall i \in \mathbb{N}.(f(i+1) - f(i) < n)\}$;
- $\mathcal{L}(e^S) = \{(u_{f(0)}u_2 \dots u_{f(1)-1}, u_{f(1)} \dots u_{f(2)-1}, \dots) \mid \boldsymbol{u} \in \mathcal{L}(e)$ and $f : \mathbb{N} \to \mathbb{N}^+$ is a nondecreasing function, with $f(0) = 1$, such that $\forall n \in \mathbb{N}\ \exists k \in \mathbb{N}\ \forall i > k(f(i+1) - f(i) > n)\}$;
- $\mathcal{L}(e^T) = \{(u_{f(0)}u_2 \dots u_{f(1)-1}, u_{f(1)} \dots u_{f(2)-1}, \dots) \mid \boldsymbol{u} \in \mathcal{L}(e)$ and $f : \mathbb{N} \to \mathbb{N}^+$ is a nondecreasing function, with $f(0) = 1$, such that $\exists^\omega n \in \mathbb{N}\ \forall k \in \mathbb{N}\ \exists i > k.(f(i+1) - f(i) = n)\}$.

Given a sequence $\boldsymbol{v} = (u_{f(0)}u_2 \dots u_{f(1)-1}, u_{f(1)} \dots u_{f(2)-1}, \dots) \in e^{op}$, with $\boldsymbol{u} \in \mathcal{L}(e)$ and $op \in \{*, B, S, T\}$, we define the *sequence of exponents of e in $\boldsymbol{v}$*, denoted by $N(\boldsymbol{v})$, as the sequence $\big(f(i+1) - f(i)\big)_{i \in \mathbb{N}}$. While the $*$-constructor does not impose any constraint on $N(\boldsymbol{v})$, the B-constructor forces it to be bounded, the S-constructor forces it to be strongly unbounded, that is, its limit inferior is infinite (equivalently, no exponent occurs infinitely often in the sequence), and the T-constructor requires infinitely many exponents to occur infinitely often.

Let e be a BST-regular expression. The *ω-constructor* turns languages of infinite word sequences into languages of ω-words (flattening) as follows:

- $\mathcal{L}(e^\omega) = \{w \mid |w| = \infty$ and $w = u_1u_2u_3 \dots$ for some $\boldsymbol{u} \in \mathcal{L}(e)\}$.

ωBST-expressions are defined by the following grammar, where we denote languages of word sequences (resp., words) by lowercase (resp., uppercase) letters $e, e_1, \dots,$ (resp., $E, E_1, \dots, R, R_1, \dots$): $E ::= E{+}E \mid R{\cdot}E \mid e^\omega$ where R is a regular expression, e is a BST-regular expression, and $+$ and $\cdot$ respectively denote union and concatenation of word languages (formally, $\mathcal{L}(E_1{+}E_2){=}\mathcal{L}(E_1){\cup}\mathcal{L}(E_2)$ and $\mathcal{L}(E_1{\cdot}E_2){=}\{u{\cdot}v \mid u \in \mathcal{L}(E_1), v \in \mathcal{L}(E_2)\}$).[5] As we did for languages of word sequences, we will sometimes omit the operator $\cdot$ between word languages.

Interval temporal logics AB, $AB\bar{A}$, $AB{\sim}$, and $AB\bar{A}{\sim}$. Syntax and semantics of AB, $AB\bar{A}$, $AB{\sim}$, and $AB\bar{A}{\sim}$ are defined as follows. AB features modalities $\langle A\rangle$ and $\langle B\rangle$, that correspond to Allen's relations *meets* (denoted by A) and *begun by* (B), respectively. Its satisfiability problem is EXPSPACE-complete over both finite linear orders and $\mathbb{N}$ [13]. Formally, given a set $\mathcal{P}rop$ of proposition letters, formulas of AB are defined as follows: $\varphi := p \mid \varphi \vee \varphi \mid \neg\varphi \mid \langle A\rangle\varphi \mid \langle B\rangle\varphi$, where $p \in \mathcal{P}rop$. We use the shorthands $\varphi \wedge \psi$ for $\neg(\neg\varphi \vee \neg\psi)$, $[X]\varphi$ for $\neg\langle X\rangle(\neg\varphi)$, with $X \in \{A, B\}$, $\bot$ for $p \wedge \neg p$, and $\top$ for $p \vee \neg p$. Formulas of AB are interpreted in interval temporal structures over $\mathbb{N}$ endowed with Allen's relations A and B. We identify any given ordinal $N \leq \omega$ with the prefix of length N of $\mathbb{N}$, and we accordingly define $\mathbb{I}(N)$ as the set of all closed intervals $[i, j]$, with $i, j \in N$ and $i \leq j$. A special role will be played by point intervals (intervals $[i, i]$, for each $i \in N$) and unit intervals (intervals $[i, i+1]$), which are captured by the formulas $\pi = [B]\bot$ and $unit = \langle B\rangle\top \wedge [B][B]\bot$, respectively. Allen's relations A and B are defined as follows. Given two intervals $[i, j], [i', j'] \in \mathbb{I}(N)$, we say that: *(a)* $[i, j]A[i', j']$ if and only if $j = i'$; *(b)* $[i, j]B[i', j']$ if and only if $i = i'$ and $j > j'$. AB semantics is given in terms of interval models $M = \langle \mathbb{I}(N), A, B, V\rangle$,

[5] Notice the abuse of notation with the previous definition of the operators $+$ and $\cdot$ over languages of word sequences.

where $V : \mathbb{I}(N) \to \mathcal{P}(\mathcal{P}rop)$ is the valuation function that assigns to every interval the set of proposition letters that are true on it. Truth of AB formulas is inductively defined as follows: (i) clauses for proposition letters and Boolean connectives are defined as usual; (ii) $M, [i, j] \models \langle X \rangle \varphi$, for $X \in \{A, B\}$, if and only if there exists an interval $[i', j']$ such that $[i, j]X[i', j']$ and $M, [i', j'] \models \varphi$. Given $M = \langle \mathbb{I}(N), A, B, V \rangle$ and φ, M satisfies φ if there is $[i, j] \in \mathbb{I}(N)$ such that $M, [i, j] \models \varphi$, and φ is satisfiable if there is an interval model M that satisfies it.

$AB\bar{A}$ is obtained from AB by adding the (past) modality $\langle \bar{A} \rangle$ for the Allen relation *met by* ($\bar{A}$). Unlike what happens with point-based temporal logics, the addition of past operators to interval ones usually increases both their expressiveness and their computational complexity (see, e.g., [9]). This is the case with $AB\bar{A}$: its satisfiability problem is still decidable, but non-primitive recursive, over finite linear orders, and undecidable over $\mathbb{N}$ [12]. $AB\bar{A}$ syntax extends that of AB in the obvious way. As for its semantics, for any pair of intervals $[i, j], [i', j'] \in \mathbb{I}(N)$, $[i, j]\bar{A}[i', j']$ if and only if $i = j'$. $AB\bar{A}$ formulas are interpreted on models $M = \langle \mathbb{I}(N), A, B, \bar{A}, V \rangle$. The semantics is defined as expected.

$AB{\sim}$ is obtained from AB by adding an equivalence relation $\sim$ over the points of the model. Similarly to $AB\bar{A}$, the satisfiability problem for $AB{\sim}$ remains decidable over finite linear orders, but it becomes non-primitive recursive, while decidability is lost over $\mathbb{N}$ [14]. Formally, the language of AB is extended with a new symbol $\sim$, and formulas are built according to the syntax: $\varphi := p \mid \sim \mid \varphi \vee \varphi \mid \neg \varphi \mid \langle A \rangle \varphi \mid \langle B \rangle \varphi \mid \langle \bar{B} \rangle \varphi$, where $p \in \mathcal{P}rop$. $AB{\sim}$ formulas are interpreted on models $M = \langle \mathbb{I}(N), A, B, \sim, V \rangle$, where $\sim$ is an equivalence relation on N. Truth is defined as for AB formulas, with an additional semantic clause for $\sim$: $M, [i, j] \models \sim$ if and only if $i \sim j$. Syntax and semantics of $AB\bar{A}{\sim}$ are obtained from those of $AB\bar{A}$ and $AB{\sim}$ by merging them in the obvious way.

Hereafter, we use modalities $[G]$ (*globally in the future*) and $[init]$(*every initial interval*), which are defined as follows: *(i)* $[G]\varphi$ iff $[B][A]\varphi \wedge [A][A]\varphi$, and *(ii)* $[init]\varphi$ iff $[B](\pi \to [A]\varphi) \wedge (\pi \to [A]\varphi)$. Both of them are definable in all the above logics. When evaluated on $[x, y]$, $[G]\varphi$ forces φ to be true over all $[w, z]$ with $w \geq x$; in particular, when evaluated on $[0, y]$, it forces φ to be true on all intervals. When evaluated on $[x, y]$, $[init]\varphi$ forces φ to be true on all $[x, z]$; in particular, when evaluated on $[0, y]$, it forces φ to be true on all initial intervals.

3 Encoding ωB-, ωS-, and ωT-regular languages

Building on the encoding of regular and ω-regular expressions in AB given in [15] (for the convenience of the reader, they are reported in the appendix), we provide an encoding of ωB-, ωS-, and ωT-regular ones into suitable extensions of AB. As already noticed, an encoding of ωB-regular (resp., ωS-regular) expressions in $AB\bar{B}\bar{A}$ (resp., $AB\bar{B}\sim$) was proposed in [15] (resp., [14]). Unfortunately, both encodings are flawed. Let us focus on ωB-regular expressions. Let E be an expression and e_i be a sub-expression of the form e_j^B. In [15], two formulas of $AB\bar{B}\bar{A}$ are exploited to encode the B-constructor: a local one, which is basically the same used for the Kleene star, and a global one, that constrains the size of e_i

blocks to be bounded. The latter formula says that, for any ω-word belonging to the language, it is possible to define an infinite sequence of positions (*milestones*) such that (i) no milestone is properly contained in an e_i block, and (ii) the window between two consecutive milestones contains a non-increasing number of occurrences of e_j. In the following, we show that such a claim is wrong.

Let $\sigma = (a_n)_{n \in \mathbb{N}}$ be a sequence of natural numbers. We define a grouping of σ as a sequence ρ of natural numbers $(b_n)_{n \in \mathbb{N}}$ such that $b_n = \sum_{i=f(n)}^{f(n+1)-1} a_i$, where $f : \mathbb{N} \to \mathbb{N}$, with $f(0) = 0$, is an increasing function (that identifies the milestones). The above claim can be reformulated as follows: every bounded sequence admits a non-increasing grouping. Since we are dealing with $\mathbb{N}$, the only sequences that satisfy such a condition are the definitively constant ones. A counterexample is given by the sequence $\sigma = 1, 2, 1, 2, 2, 1, 2, 2, 2, 1, 2, 2, 2, 2, 1, 2, 2, 2, 2, 2, 1, \ldots$ (we would like to thank David Barozzini for it). By contradiction, assume that there is a non-increasing grouping $\rho = (b_n)_{n \in \mathbb{N}}$ of σ. Then, for every $i \in \mathbb{N}$, it holds that $f(i+1) - f(i) \leq b_0$. We show that ρ is not definitively constant. Suppose that, for some i, b_i is odd. For i large enough, we can assume that the sequence $a_{f(i)}, a_{f(i)+1}, \ldots, a_{f(i+1)-1}$ contains exactly one occurrence of 1 (as the difference between of two consecutive terms of f is bounded, while the distance between two consecutive occurrences of 1 is not). Moreover, if i is large enough, we can also assume that the sequence $a_{f(i+1)}, a_{f(i+1)+1}, \ldots, a_{f(i+1)+b_0}$ contains no occurrence of 1. It follows that b_{i+1} is even, and thus $b_{i+1} < b_i$. Suppose now that, for some i, b_i is even. For i large enough, we can assume that the sequence $a_{f(i)}, a_{f(i)+1}, \ldots, a_{f(i+1)-1}$ contains no occurrence of 1. It follows that there exists $j > 0$ such that $a_{f(i+j)}, a_{f(i+j)+1}, \ldots, a_{f(i+j+1)-1}$ contains exactly one occurrence of 1. Thus, b_{i+j} is odd, and $b_{i+j} < b_i$.

Let us consider now ωB-, and ωS-, and ωT-regular expressions. Given an expression E, we list its sub-expressions $e_1, \ldots, e_n (= E)$ in increasing order of complexity. Then, for all i, we introduce two proposition letters $expr_i$ and $expr_i^{end}$, and then we define inductively a formula φ_{expr_i}. Finally, we capture the language $\mathcal{L}(E)$ with the formula $\varphi_E = \bigwedge_{i=1}^{n} \varphi_{expr_i} \wedge \bigwedge_{i=1}^{n} \varphi_{expr_i}^{end} \wedge \bigwedge_{i=1}^{n} \varphi_{expr_i}^{\not\nearrow}$. The only missing ingredient is a way to recursively define formulas φ_{expr_i} for e_i of the forms e_j^B, e_j^S, and e_j^T. These formulas are conjunctions of two sub-formulas, a local one, that is the same we defined in the case $e_i = e_j^*$, and a global one, which guarantees the constraints imposed by the B-, S-, and T-constructors.

It is worth remarking that if, for a sub-expression $e_i = e_j^B$ of E, there are only finitely many occurrences of $expr_i$ intervals in the model, then we do not need to guarantee the satisfaction of the boundedness constraint imposed by the B-constructor (the same holds for $e_i = e_j^S$ and $e_i = e_j^T$). This is the case, for instance, with expressions like $(a^B b + a^* b)^\omega$, which is, in fact, equivalent to $(a^* b)^\omega$ due to the *shuffle* operator that, from a given position on, can postpone forever the selection of occurrences of $a^B b$. Thus, formulas we are going to build in the next sections are assumed to be (and, in the end, will be) guarded by the requirement that infinitely many $expr_i$ intervals occur.

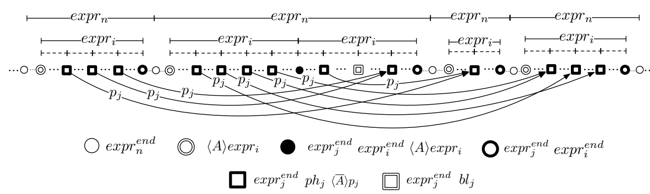

Fig. 1: Example of the structure we are enforcing by means of formula $\Phi_B^{(i,j)}$ for an expression $E = (e_n)^\omega$, where e_n contains the sub-expression $e_i = e_j^B$ (dashed intervals represent $expr_j$ intervals).

ωB-regular languages in $AB\bar{A}$. Let $B(E)=\{(i,j) : e_i=e_j^B$ is a sub-expression of $E\}$. To force the proper behaviour of the B-constructor, for every $(i,j)\in B(E)$, we introduce the additional proposition letters ph_j, bl_j, and p_j, which can be exploited to express (by means of suitable $AB\bar{A}$ formulas) the following properties:

1. ph_j and bl_j may only label left endpoints of $expr_j$ intervals which are not left endpoints of $expr_i$ ones, but they cannot label the same points: $[G](ph_j \vee bl_j \rightarrow \pi \wedge \langle A\rangle expr_j \wedge \neg\langle A\rangle expr_i) \wedge [G]((ph_j \rightarrow \neg bl_j) \wedge (bl_j \rightarrow \neg ph_j))$;
2. there exists $n \in \mathbb{N}$ such that every $n' > n$ which is the left endpoint of an $expr_j$ interval, but not the left endpoint of an $expr_i$ interval, is labeled with either ph_j or bl_j: $[G](\langle A\rangle[A](\langle A\rangle expr_j \wedge [A]\neg expr_i \rightarrow \langle A\rangle(ph_j \vee bl_j)))$;
3. in between two consecutive bl_j points x, y, with $x < y$, there exists at least one point z, with $x < z < y$, such that z is the left endpoint of an $expr_i$ interval: $[G](\langle B\rangle bl_j \wedge \langle A\rangle bl_j \rightarrow \langle B\rangle\langle A\rangle expr_i)$;
4. every ph_j point is the left endpoint of exactly one p_j interval: $[G](ph_j \rightarrow \langle A\rangle p_j) \wedge [G](p_j \rightarrow \neg\langle B\rangle p_j)$;
5. every p_j interval is begun by a ph_j point and strictly contains exactly one bl_j point: $[G](p_j \rightarrow \langle B\rangle ph_j \wedge \langle B\rangle\langle A\rangle bl_j \wedge [B](\langle A\rangle bl_j \rightarrow \neg\langle B\rangle\langle A\rangle bl_j))$;
6. every ph_j point x for which there exists a bl_j point y such that $y < x$ is the right endpoint of at least one p_j interval: $[G](\langle A\rangle ph_j \wedge \langle B\rangle bl_j \rightarrow \langle A\rangle\langle\bar{A}\rangle p_j)$.

Fig. 1 gives a graphical account of the above properties. Let us assume that infinitely many $expr_i$ intervals occur. Properties 1–2 guarantee that, from a point on, say it n, every point which is the left endpoint of an $expr_j$ interval, but not of an $expr_i$ one, is labeled with either ph_j or bl_j. By properties 4–5, every ph_j point is followed by a bl_j one. Thus, the suffix starting at n can be seen as a (possibly finite or even empty) sequence of slices $[n_0, n_1], [n_1, n_2] \ldots$, where $\{n_0, n_1, \ldots\}$ is the set of bl_j points greater than n. Now, let $[n_k, n_{k+1}]_{ph_j}$ be the set of all and only those ph_j points x, with $n_k < x < n_{k+1}$, that, by properties 1–2, are left endpoints of $expr_j$ intervals, but not of $expr_i$ ones. By properties 4–6, p_j encodes a series of surjective functions $f_k : [n_k, n_{k+1}]_{ph_j} \rightarrow [n_{k+1}, n_{k+2}]_{ph_j}$, with $k \geq 0$, linking the ph_j points of pairs of consecutive slices.[6] It follows that $|[n_0, n_1]_{ph_j}| \geq |[n_1, n_2]_{ph_j}| \geq \ldots$, i.e., the sequence is not increasing. Finally, property 3 imposes that, for every k, there is at least one point x, with $n_k < x < n_{k+1}$, which is

[6] As a matter of fact, the image of one such function f_k might also include elements not belonging to $[n_{k+1}, n_{k+2}]_{ph_j}$; however, properties 4–6 guarantee that $[n_{k+1}, n_{k+2}]_{ph_j}$ is included in the image, which is enough for our purposes.

the left endpoint of an $expr_i$ interval. Then, every $expr_i$ interval starting after n spans at most two adjacent slices, and thus it contains at most $|[n_0, n_1]_{ph_j}| * 2$ many $expr_j$ intervals, thus providing a bound, as required by the B-constructor.

Now, for every $(i, j) \in B(E)$, let $\Psi_B^{(i,j)}$ be the conjunction of the above formulas and $\Phi_B^{(i,j)} = [G]\langle A\rangle\langle A\rangle expr_i \rightarrow \Psi_B^{(i,j)}$.

Theorem 1. *Let E be an ωB-regular expression over Σ. Then, $\mathcal{L}(E) = \{w \in \Sigma^\omega : w \approx M$ for some model M such that $M, [0, n] \models \varphi_E \wedge \varphi_\Sigma \wedge \bigwedge_{(i,j)\in B(E)} \Phi_B^{(i,j)}$ for some $n \in \mathbb{N}\}$.*

ωS-regular languages in $AB\sim$. Let $S(E)=\{(i,j) : e_i=e_j^S$ is a sub-expression of $E\}$. To force the proper behaviour of the S-constructor, we make use of the equivalence relation $\sim$ and, for all $(i, j) \in S(E)$, we introduce the proposition letters ph_j and new_j, that allow us to express (by means of suitable $AB\sim$ formulas) the following properties:

1. ph_j may only label left endpoints of $expr_j$ intervals which are neither left nor right endpoints of $expr_i$ ones (this implies that a ph_j point must occur inside an $expr_i$ interval): $[G](ph_j \rightarrow \pi \wedge \langle A\rangle expr_j \wedge \neg\langle A\rangle expr_i \wedge \neg\langle A\rangle expr_i^{end})$;
2. if two ph_j points x and y, with $x < y$, belong to the same $expr_i$ interval, then $x \not\sim y$, or, equivalently, if $x \sim y$, then they belong to two distinct $expr_i$ intervals: $[G](\sim \wedge\langle B\rangle ph_j \rightarrow \langle A\rangle ph_j \wedge \langle B\rangle\langle A\rangle expr_i^{end})$;
3. for every $expr_i$ interval $[n, n']$ that contains at least one ph_j point, there is an $expr_i$ interval $[\overline{n}, \overline{n}']$ with $\overline{n} \geq n'$. Moreover, if $\overline{n}$ is the smallest point such that $\overline{n} \geq n'$ and there is $\overline{n}' > \overline{n}$ for which $[\overline{n}, \overline{n}']$ is an $expr_i$ interval, then, for every ph_j point x, with $n < x < n'$, there is a ph_j point y, with $\overline{n} < y < \overline{n}'$, such that $x \sim y$ (notice that property 2 forces such a point y to be uniquely determined): $[G](ph_j \rightarrow \langle A\rangle(\neg\pi\wedge \sim \wedge[B](\langle A\rangle expr_i^{end} \rightarrow [B][A]\neg expr_i^{end})))$;
4. for all $expr_i$ interval $[n, n']$, let $|[n, n']|_{ph_j} = |\{x : n < x < n',\ x$ is a ph_j point $\}|$ be the number of ph_j points inside $[n, n']$. If the model features an infinite sequence $[n_0, n_0'], [n_1, n_1'], \ldots$ of $expr_i$ intervals, then the sequence $|[n_0, n_0']|_{ph_j}, |[n_1, n_1']|_{ph_j}, \ldots$ is non-decreasing and unbounded. It follows that there are infinitely many classes of ph_j points. This is expressed by means of the auxiliary proposition letter new_j as follows: (i) new_j may only appear in a labeling that already contains ph_j; (ii) for a new_j point x, there is no point y, with $y < x$, such that $y \sim x$; (iii) for every ph_j point x we have that there exists a new_j point $y > x$:
$$[G](new_j \rightarrow ph_j) \wedge [G](\neg\pi \wedge \sim \rightarrow [A]\neg new_j) \wedge [G](ph_j \rightarrow \langle A\rangle(\neg\pi \wedge \langle A\rangle new_j)).$$

A graphical account of the properties imposed by the above formulas is given in Fig. 2. First, we observe that $expr_i$ intervals may contain a different number of $expr_j$ intervals. From one $expr_i$ interval to the next one, such a number may increase, decrease, or remain the same. However, according to the semantics of the S-constructor, for every $n \in \mathbb{N}$, such a number is forced to be greater than n for a suffix of the model. This is done by means of proposition letter ph_j. Once the left endpoint of an $expr_j$ interval included in an $expr_i$ one is labeled with ph_j, properties 2 and 3 guarantee that, in every future $expr_i$ interval, there is

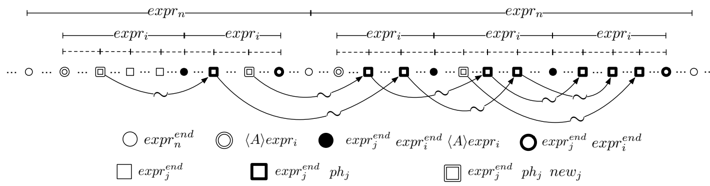

Fig. 2: Example of the structure we are enforcing by means of formula $\Phi_S^{(i,j)}$ for an expression $E = (e_n)^\omega$, where e_n contains the sub-expression $e_i = e_j^S$ (dashed intervals represent $expr_j$ intervals).

exactly one ph_j point belonging to the same equivalence class. Every equivalence class can thus be seen as an infinite chain (with a starting point) of ph_j points belonging to consecutive $expr_i$ intervals. It follows that the number n of distinct ph_j points belonging to an $expr_i$ interval forces all the following $expr_i$ intervals to contain at least n distinct ph_j points, and thus, by property 1, at least n $expr_j$ intervals. Now, let us observe, as shown in Fig. 2, that not all the left endpoints of the $expr_j$ intervals belonging to an $expr_i$ one must be labeled with ph_j. In this way, their number may fluctuate, while ph_j points ensure that such a number does not go below a certain threshold. Finally, property 4 guarantees that the equivalence relation of ph_j points is of infinite index, and thus the behaviour of the S-constructor is correctly captured.

To complete the encoding, we need to guarantee the presence of ph_j points and the behaviour induced by them whenever there are infinitely many $expr_i$ intervals in the model. Indeed, if e_j is matched to ε infinitely often, i.e., if there are infinitely many occurrences of point intervals labeled with $expr_j$, then there is no need to guarantee the satisfaction of the unboundedness constraint imposed by the S-constructor, because ε can "hide" arbitrarily many repetitions of e_j.

To this end, it suffices to add a formula that constrains the model to feature at least one ph_j point if it features infinitely many $expr_i$ intervals, but only finitely many point intervals labeled with $expr_j$. Last but not least, we need to prevent the case in which there are infinitely many occurrences of point intervals labeled with $expr_i$ but not $expr_j$, which corresponds to infinitely many instantiations of e_i with zero repetitions of e_j. Such a condition is encoded by the second conjunct of the consequent of the implication: $[G]\langle A\rangle\langle A\rangle expr_i \wedge \langle A\rangle[A][A](expr_j \rightarrow \neg\pi) \rightarrow \langle B\rangle\langle A\rangle\langle A\rangle ph_j \wedge \langle A\rangle[A][A](expr_i \rightarrow \neg\pi)$.

For all $(i, j) \in S(E)$, let $\Phi_S^{(i,j)}$ be the conjunction of the above formulas.

Theorem 2. *Let E be an ωS-regular expression over Σ. Then, $\mathcal{L}(E) = \{w \in \Sigma^\omega : w \approx M$ for some model M such that $M, [0, n] \models \varphi_E \wedge \varphi_\Sigma \wedge \bigwedge_{(i,j)\in S(E)} \Phi_S^{(i,j)}$ for some $n \in \mathbb{N}\}$.*

ωT-regular languages in $AB\bar{A}\sim$. Let $T(E)=\{(i, j) : e_i{=}e_j^T$ is a sub-expression of $E\}$. To encode ωT-regular languages in $AB\bar{A}\sim$, we first show that a particular class of models over $\mathbb{N}$ can be captured by $AB\bar{A}\sim$ formulas $\Phi_\infty^{(i,j)}$, for $(i, j) \in T(E)$. Then, we use such a formula to constrain the behaviour of $(.)^T$.

By making use of proposition letters ph_j, bl_j, p_j, q_j, and $conf_j$, and of the proposition letter $\sim$, representing an equivalence relation over $\mathbb{N}$, we want to characterize, through $\Phi_\infty^{(i,j)}$, the models that satisfy the following properties:

1. ph_j, bl_j, and $conf_j$ only appear as labels of points, ph_j and bl_j never occur together in the same labeling, and $conf_j$ only appears in a labeling containing also bl_j, that is, a *configuration* (an interval whose endpoints are consecutive $conf_j$ points) features one or more *blocks* (intervals whose endpoints are consecutive bl_j points): $[G]((bl_j \vee ph_j \to \pi) \wedge (conf_j \to bl_j) \wedge (ph_j \to \neg bl_j))$;
2. there are infinitely many $conf_j$ points, that is, there are infinitely many configurations: $[G]\langle A\rangle\langle A\rangle conf_j$;
3. between two consecutive bl_j points there is at least one ph_j point and all of them belong to the same equivalence class, that is, each block is associated with exactly one equivalence class of ph_j points: $[G](\langle B\rangle bl_j \ \wedge \ \langle A\rangle bl_j \ \to \langle B\rangle\langle A\rangle ph_j) \ \wedge \ [G](\langle B\rangle ph_j \wedge \langle A\rangle ph_j \wedge [B][A]\neg bl_j \to \sim)$;
4. let x, y, and z, with $x < y < z$, be three consecutive bl_j points and y be not labeled with $conf_j$; then, there are more ph_j points between x and y than between y and z, that is, the sequence of the numbers of ph_j points featured in blocks of the same configuration is strictly decreasing:
$$[G](p_j \to \langle B\rangle ph_j \wedge \langle A\rangle ph_j \wedge [B]\neg p_j \wedge [B][A]\neg conf_j \wedge \\ \langle B\rangle\langle A\rangle bl_j \wedge [B](\langle A\rangle bl_j \to [B][A]\neg bl_j)) \wedge \\ [G](ph_j \wedge [A](\neg\pi \wedge \sim \to \langle B\rangle\langle A\rangle bl_j) \to [A]\neg p_j) \wedge \\ [G](ph_j \wedge \langle \bar{A}\rangle(\langle B\rangle bl_j \wedge [B][A]\neg conf_j) \to \langle \bar{A}\rangle p_j);$$
5. for every pair of ph_j points x and y, with $x < y$, if there is a bl_j point but no $conf_j$ point between them, then $x \not\sim y$, that is, pairs of distinct blocks in the same configuration represent distinct equivalence classes of ph_j points:
$$[G](\sim \to \ (\langle B\rangle ph_j \ \leftrightarrow \ \langle A\rangle ph_j)) \wedge \\ [G](\sim \wedge \langle B\rangle ph_j \ \wedge \ \langle B\rangle\langle A\rangle bl_j \to \langle B\rangle\langle A\rangle conf_j);$$
6. for every ph_j point x there is a ph_j point $y > x$ such that $x \sim y$ and there is exactly one $conf_j$ point between x and y, that is, an equivalence class in a configuration is witnessed in all the following configurations:
$$[G](ph_j \to \langle A\rangle(\sim \wedge \langle B\rangle\langle A\rangle conf_j \wedge [B](\langle A\rangle conf_j \to [B][A]\neg conf_j)));$$
7. let x, y, and z be three consecutive $conf_j$ points, with $x < y < z$; then, there are less bl_j points between x and y than between y and z, that is, the sequence of the numbers of blocks in configurations is strictly increasing:
$$[G](conf_j \to \langle A\rangle(\langle A\rangle ph_j \wedge [B](\neg\pi \to [A]\neg conf_j) \wedge \langle A\rangle[\bar{A}]\neg \sim));$$
8. if (x, y) and (x', y') are two pairs of bl_j points, with $x < y < x' < y'$, both witnessing the same equivalence class, then the number of ph_j points between x and y is greater than or equal to the number of those between x' and y', that is, the sequence of blocks of ph_j points in the same equivalence class is non-increasing in the number of ph_j points in every block:
$$[G](q_j \to \sim \wedge [B]\neg q_j \wedge \langle B\rangle\langle A\rangle conf_j \wedge [B](\langle A\rangle conf_j \to [B][A]\neg conf_j)) \wedge \\ [G](ph_j \wedge \langle \bar{A}\rangle(\sim \wedge \langle B\rangle\langle A\rangle bl_j) \to \langle \bar{A}\rangle q_j).$$

Let $\Phi_\infty^{(i,j)}$ be the conjunction of the above formulas. A graphical account of the structure enforced by $\Phi_\infty^{(i,j)}$ is given in Fig. 3. Notice that there may be points whose labeling do not contain any of the proposition letters ph_j, bl_j, and $conf_j$.

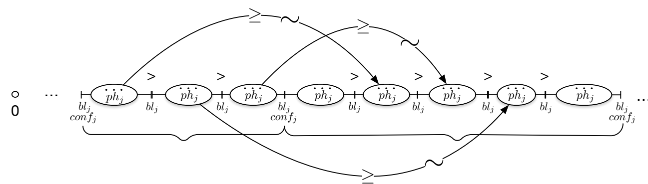

Fig. 3: Example of the type of structure we enforce by means of formula $\Phi_\infty^{(i,j)}$.

Thanks to $\Phi_\infty^{(i,j)}$, a model can be seen as an infinite sequence of configurations $[conf_j^0, conf_j^1], [conf_j^1, conf_j^2], \ldots$ For every $x \in \mathbb{N}$, $conf_j^x$ contains a finite sequence of $n(x)+1$, with $n : \mathbb{N} \to \mathbb{N}$, sets $Sbl_j^{x,0}, \ldots, Sbl_j^{x,n(x)}$ of ph_i points each one associated with exactly one equivalence class, i.e., points in $Sbl_j^{x,y}$ belong to the same equivalence class, for every $y \in \{0, \ldots, n(x)\}$. Formally, $Sbl_j^{x,y} = \{n \in \mathbb{N} : M, [n,n] \models ph_j, |\{n' < n : M, [n',n'] \models conf_j\}| = x+1, |\{n' < n : M, [n',n'] \models bl_j, \forall n''(n' < n'' < n \to M, [n'',n''] \not\models conf_j)\}| = y\}$. Intuitively, $n(x)+1$ is the number of blocks in $[conf_x, conf_{x+1}]$ and $|Sbl_j^{x,y}|$ is the number of ph_j points in the yth block of $[conf_x, conf_{x+1}]$. The following properties hold:

(P1) the function $n(x)$ is strictly increasing (property 7);

(P2) for all $x, y, y' \in \mathbb{N}$, with $0 \leq y < y' \leq n(x)$, $|Sbl_j^{x,y}| > |Sbl_j^{x,y'}|$ (property 4);

(P3) for every ph_j point w it is possible to identify an infinite sequence of pairs of indexes $(x, y_0), (x+1, y_1), \ldots$ such that $[w]_\sim = \bigcup_{k \in \mathbb{N}} Sbl_j^{x+k,y_k}$ (property 6) and $|Sbl_j^{x,y_0}| \geq |Sbl_j^{x+1,y_1}| \geq \ldots$ (property 8);

$(P3)$ states that, for every equivalence class $[w]_\sim$ of ph_j points, there is a configuration such that $[w]_\sim$ is witnessed exactly in all the successive configurations. Moreover, it states that the blocks that witness $[w]_\sim$ feature a non-increasing number of points. Let $(x, y_0), (x+1, y_1), \ldots$ be such that $[w]_\sim = \bigcup_{k \in \mathbb{N}} Sbl_j^{x+k,y_k}$. Since the number of points in each block is finite, there is $k' \in \mathbb{N}$ for which $|Sbl_j^{x+k',y_{k'}}| = |Sbl_j^{x+k'+1,y_{k'+1}}| = \ldots$, i.e., the sequence of Sbl_j^{x+k,y_k} cardinalities converges to a single value (we call it the value of the equivalence class), denoted by $val([w]_\sim)$. By $(P2)$, it holds that for any two distinct equivalence classes $[w]_\sim$ and $[w']_\sim$ of ph_j points, $val([w]_\sim) \neq val([w']_\sim)$; otherwise, it would be possible to find a configuration featuring two distinct blocks with the same number of ph_j points, which contradicts $(P2)$. Finally, $(P1)$ guarantees that the number of distinct equivalence classes is infinite. The following lemma holds.

Lemma 1. *If $M, [0,0] \models \Phi_\infty^{(i,j)}$, then there exists an infinite subset N of $\mathbb{N}$ such that for every $n \in N$ there is $w \in \mathbb{N}$ with $val([w]_\sim) = n$.*

An *instantiation* of an equivalence class with an $expr_i$ interval is a correspondence between the number of ph_j points in a block associated with that class and the number of $expr_j$ intervals in the $expr_i$ interval. This is the case when the set of ph_j points in the block and the set of points starting an $expr_j$ interval within the $expr_i$ interval coincide. The T-constructor forces all the equivalence classes to be instantiated infinitely many times. Indeed, if an equivalence class is instantiated infinitely often, the infinitely many $expr_i$ intervals are matched by repetitions of $val([w]_\sim)$ many $expr_j$ intervals in an ω-iteration. Since the number of equivalence classes is infinite and all of them feature distinct values $val([w]_\sim)$, the behaviour of the T-constructor is correctly encoded.

However, there are cases in which we do not need to satisfy the constraint imposed by the T-constructor or it suffices to satisfy a weaker version of it. This is the case when (i) there are only finitely many $expr_i$ intervals, as for B- and S-constructors; (ii) there are infinitely many point intervals labeled with both $expr_i$ and $expr_j$; this corresponds to e_i being matched by occurrences of $expr_j$ matched, in turn, by the empty string ε; in this case, such an $expr_i$ interval can be thought of as featuring any possible number of $expr_j$ intervals; (iii) there are infinitely many $expr_j$ point intervals but only finitely many labeled with both $expr_i$ and $expr_j$; in this case, it suffices to impose that at least one equivalence class of ph_j points is instantiated infinitely often.

When none of the cases above applies, we force all the equivalence classes to be instantiated infinitely many times. For an $expr_i$ interval $[x, y]$, let $points_j([x, y]) = \{z : x \leq z \leq y, \exists z'(M, [z, z'] \models expr_j)\}$, and, for a ph_j point w, let $Seq([w]_\sim) = (x, y_0), (x+1, y_1), \ldots$ be the sequence such that $[w]_\sim = \bigcup_{k\in\mathbb{N}} Sbl_j^{x+k, y_k}$. In what follow, we define a formula Φ^{in_j}, which uses proposition letter in_j to force that, for every equivalence class $[w]_\sim$ of ph_j points, there is an infinite sub-sequence $(\overline{x}_0, \overline{y}_0), (\overline{x}_1, \overline{y}_1), \ldots$ of $Seq([w]_\sim)$ such that for every $h \in \mathbb{N}$ there is a distinct $expr_i$ interval $[x'_h, y'_h]$ with $points_j([x'_h, y'_h]) = |Sbl_j^{\overline{x}_h, \overline{y}_h}|$, i.e., each equivalence class is instantiated infinitely many times:

- in_j appears only as the label of ph_j points that begin $expr_j$ intervals:
  $$[G](in_j \rightarrow ph_j \wedge \langle A\rangle expr_j);$$
- ph_j points that share the same block of an in_j point are labeled with in_j:
  $$[G]((\langle B\rangle ph_j \wedge \langle A\rangle ph_j \wedge [B][A]\neg bl_j) \rightarrow (\langle A\rangle in_j \leftrightarrow \langle B\rangle in_j));$$
- every block of in_j labeled points encloses exactly an $expr_i$ labeled interval:
  $$[G](expr_i \wedge \langle B\rangle\langle A\rangle in_j \rightarrow [B][A]\neg bl_j \wedge [\bar{A}]([B][A]\neg bl_j \rightarrow [B][A]\neg in_j) \wedge [A]([B][A]\neg bl_j \rightarrow [B][A]\neg in_j));$$
- if an $expr_i$ interval contains an in_j point, then the $expr_j$ intervals within it begin with an in_j point: $[G](expr_i \wedge \langle B\rangle\langle A\rangle in_j \rightarrow [B](\langle A\rangle expr_j \rightarrow \langle A\rangle in_j))$.

For $(i, j) \in S(E)$, let Φ^{in_j} be the conjunction of the above formulas and $\Phi_T^{(i,j)}$ be

$$\big(\langle A\rangle[A][A]\neg expr_i\big) \vee \big([G]\langle A\rangle\langle A\rangle(\pi \ \wedge \ expr_i \wedge expr_j)\big) \vee$$
$$\big(\Phi_\infty^{(i,j)} \wedge \Phi^{in_j} \ \wedge \ [G]\langle A\rangle\langle A\rangle(\pi \ \wedge \ expr_j) \qquad \wedge \ \langle B\rangle\langle A\rangle\langle A\rangle in_j \wedge$$
$$\wedge [G](in_j \rightarrow \langle A\rangle(\langle B\rangle\langle A\rangle conf_j \ \wedge \ in_j))\big) \vee$$
$$\big(\Phi_\infty^{(i,j)} \wedge \Phi^{in_j} \ \wedge \ \langle A\rangle[A][A](expr_j \rightarrow \neg\pi) \ \wedge \ [G](ph_j \rightarrow \langle A\rangle(\neg\pi \wedge \ \sim \wedge \langle A\rangle in_j))\big)$$

Theorem 3. *Let E be a ωT-regular expression over Σ. Then, $\mathcal{L}(E) = \{w \in \Sigma^\omega : w \approx M$ for some model M such that $M, [0, n] \models \varphi_E \wedge \varphi_\Sigma \wedge \bigwedge_{(i,j)\in T(E)} \Phi_T^{(i,j)}$ for some $n \in \mathbb{N}\}$.*

4 Conclusions

In this paper, we filled a gap in the study of extended ω-regular languages by providing a temporal logic characterization of ωB-, ωS-, and ωT-regular languages. We showed how to turn ωB-, ωS-, and ωT-regular expressions into formulas of suitable interval temporal logics. As for future work, we are looking for syntactic and/or semantic fragments of the considered interval temporal logics, that preserve (un)satisfiability of the resulting formulas and behave better from a computational point of view.

References

1. Allen, J.F.: Maintaining knowledge about temporal intervals. Comm. of the ACM **26**(11), 832–843 (1983). https://doi.org/10.1145/182.358434
2. Barozzini, D., de Frutos-Escrig, D., Della Monica, D., Montanari, A., Sala, P.: Beyond ω-regular languages: ωT-regular expressions and their automata and logic counterparts. Theor. Comput. Sci. **813**, 270–304 (2020)
3. Bojańczyk, M.: A bounding quantifier. In: CSL. LNCS, vol. 3210, pp. 41–55. Springer (2004). https://doi.org/10.1007/978-3-540-30124-0_7
4. Bojańczyk, M.: Weak MSO with the unbounding quantifier. Theory of Computing Systems **48**(3), 554–576 (2011). https://doi.org/10.1007/s00224-010-9279-2
5. Bojańczyk, M., Colcombet, T.: Bounds in ω-regularity. In: LICS. pp. 285–296 (2006). https://doi.org/10.1109/LICS.2006.17
6. Bojańczyk, M., Colcombet, T.: Boundedness in languages of infinite words. Logical Methods in Computer Science **Volume 13, Issue 4** (Oct 2017)
7. Bresolin, D., Della Monica, D., Montanari, A., Sala, P., Sciavicco, G.: Interval temporal logics over strongly discrete linear orders: Expressiveness and complexity. Theor. Comput. Sci. **560**, 269–291 (2014)
8. Bresolin, D., Della Monica, D., Montanari, A., Sala, P., Sciavicco, G.: Decidability and complexity of the fragments of the modal logic of allen's relations over the rationals. Inf. Comput. **266**, 97–125 (2019). https://doi.org/10.1016/j.ic.2019.02.002
9. Della Monica, D., Montanari, A., Sala, P.: The importance of the past in interval temporal logics: The case of propositional neighborhood logic. In: Logic Programs, Norms and Action - Essays in Honor of Marek J. Sergot on the Occasion of His 60th Birthday. LNCS, vol. 7360, pp. 79–102. Springer (2012). https://doi.org/10.1007/978-3-642-29414-3_6
10. Halpern, J.Y., Shoham, Y.: A propositional modal logic of time intervals. Journal of the ACM **38**(4), 935–962 (Oct 1991). https://doi.org/10.1145/115234.115351
11. Lodaya, K.: Sharpening the undecidability of interval temporal logic. In: Proc. of the 6th Asian Computing Science Conference – Advances in Computing Science – ASIAN. LNCS, vol. 1961, pp. 290–298. Springer (2000). https://doi.org/10.1007/3-540-44464-5_21
12. Montanari, A., Puppis, G., Sala, P.: Maximal decidable fragments of Halpern and Shoham's modal logic of intervals. In: Proc. of the 37th ICALP, Part II. LNCS, vol. 6199, pp. 345–356. Springer (2010). https://doi.org/10.1007/978-3-642-14162-1_29
13. Montanari, A., Puppis, G., Sala, P., Sciavicco, G.: Decidability of the interval temporal logic ABB over the natural numbers. In: Proc. of the 27th STACS. LIPIcs, vol. 5, pp. 597–608. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2010). https://doi.org/10.4230/LIPIcs.STACS.2010.2488
14. Montanari, A., Sala, P.: Adding an equivalence relation to the interval logic ABB: complexity and expressiveness. In: Proc. of the 28th LICS. pp. 193–202. IEEE Computer Society (2013). https://doi.org/10.1109/LICS.2013.25

15. Montanari, A., Sala, P.: Interval logics and ωb-regular languages. In: Proc. of the 7th LATA. LNCS, vol. 7810, pp. 431–443. Springer (2013). https://doi.org/10.1007/978-3-642-37064-9_38
16. Moszkowski, B.C.: Reasoning About Digital Circuits. Ph.D. thesis, Department of Computer Science, Stanford University, Stanford, CA (1983)
17. Moszkowski, B.C., Manna, Z.: Reasoning in interval temporal logic. In: Proc. of Workshop on Logic of Programs. LNCS, vol. 164, pp. 371–382. Springer (1983). https://doi.org/10.1007/3-540-12896-4_374

A Encoding regular and ω-regular languages in AB

In the following, we describe the encoding of regular and ω-regular languages in AB (they are basically those given in [15]).

To begin with, we show how to interpret words and ω-words as interval temporal models, and vice versa. For a word $w = w_0 w_1 \ldots$ over a finite alphabet Σ and a model $M = \langle \mathbb{I}(N), A, B, V\rangle$, we say that M and w are *compatible*, denoted by $w \approx M$ (or, equivalently, $M \approx w$), if $N = |w| + 1$, $\Sigma \subseteq \mathcal{P}rop$, and $V : \mathbb{I}(N) \to \mathcal{P}(\mathcal{P}rop)$ is such that on each unit interval only the proper letter holds, that is, $V([i, i+1]) \cap \Sigma = \{w_i\}$ for every $i < |w|$.

Let R be a regular expression on Σ. We show how to encode R into an AB-formula over the finite set of proposition letter $\mathcal{P}rop$, which includes Σ. First, we force proposition letters in Σ to hold true only on unit intervals, and constrain each unit interval to satisfy exactly one proposition letter in Σ:

$$\varphi_\Sigma = [G]\left(\left(unit \leftrightarrow \bigvee_{a\in\Sigma} a\right) \wedge \bigwedge_{a\in\Sigma}\left(a \to \bigwedge_{b\in\Sigma\setminus\{a\}} \neg b\right)\right).$$

The regular expression R can be given a tree structure, whose leaves and internal nodes belong to Σ and $\{+, \cdot, ^*\}$, respectively. Each sub-tree identifies a sub-expression of R. Let $e_1, \ldots, e_n$ be all the sub-expressions of R, including elements in Σ. For each e_i, we introduce two new proposition letters $expr_i$ and $expr_i^{end}$. Notice that two occurrences e_i and e_j of the same sub-expression are associated with two different pairs of proposition letters ($expr_i/expr_i^{end}$ and $expr_j/expr_j^{end}$). For each i, we force $expr_i^{end}$ to hold only at point intervals, and if there is an interval where $expr_i$ holds true, then $expr_i^{end}$ holds on its right endpoint, and on no point strictly included in the interval. Moreover, every $expr_i^{end}$ is the ending point of an $expr_i$ interval. This is formalized by the following formula:

$$\begin{aligned}\varphi^{end}_{expr_i} = {} & [G]((expr_i^{end} \to \pi) \wedge (expr_i \to \langle A\rangle expr_i^{end} \wedge \\ & [B](\neg\pi \to [A]\neg expr_i^{end}))) \wedge \\ & [init](\langle A\rangle expr_i^{end} \to \langle A\rangle(\pi \wedge expr_i) \vee \langle B\rangle\langle A\rangle(\neg\pi \wedge expr_i)) \wedge \\ & [G](\langle A\rangle expr_i^{end} \wedge \langle B\rangle(\neg\pi \wedge expr_i) \to \\ & \langle B\rangle(\neg\pi \wedge \langle A\rangle(\neg\pi \wedge expr_i))).\end{aligned}$$

Finally, we prevent two $expr_i$ intervals from intersecting (except for intersections of a single point):

$$\varphi^{\not\cap}_{expr_i} = [G](expr_i \to [B](\neg\pi \to [A](\neg expr_i \wedge \neg expr_i^{end}))).$$

Let $e_1, \ldots, e_n$ be ordered according to their complexity with $e_n = R$, that is, if e_i is a sub-expression of e_j, then $i \le j$. We define formulas φ_{expr_i} by induction on the complexity of expressions e_i.

- If $e_i = a$ for some $a \in \Sigma$, we put $\varphi_{expr_i} = [G](expr_i \to a)$.
- If $e_i = \varepsilon$, we put $\varphi_{expr_i} = [G](expr_i \to \pi)$.
- If $e_i = e_j + e_k$, we put $\varphi_{expr_i} = [G](expr_i \leftrightarrow (expr_j \vee expr_k))$.
- If $e_i = e_j e_k$, we distinguish two cases, depending on whether or not the string matched by e_j is the empty string. The first conjunct of the formula below states that every interval on which $expr_i$ holds can be split in two ordered parts on which $expr_j$ and $expr_k$ respectively hold. It distinguishes two cases: either $expr_j$ holds on the left sub-interval (possibly a point interval) and

$expr_k$ holds over the right one (necessarily not a point interval), or $expr_j$ holds on the whole interval and $expr_k$ holds on its right endpoint (a point interval). Notice that the latter covers the case where both e_i and e_j match the empty string. The second conjunct constrains every $expr_j$ interval to occur as a (not necessarily strict) prefix of an $expr_i$ interval and to be followed by an $expr_k$ one. Similarly, the third conjunct constrains every $expr_k$ interval to occur as a (not necessarily strict) suffix of an $expr_i$ interval and to be preceded by an $expr_j$ one. In addition, the formula guarantees that $expr_j$ and $expr_k$ do not intersect (except for intersections of a single point).

$$\begin{aligned}\varphi_{expr_i} = {} & [G](expr_i \rightarrow \langle B\rangle(expr_j \wedge \langle A\rangle(expr_k \wedge \langle A\rangle expr_i^{end} \wedge \\ & \quad [B][A]\neg expr_i^{end})) \vee (expr_j \wedge \langle A\rangle(\pi \wedge expr_k))) \wedge \\ & [G]((\langle A\rangle expr_j \rightarrow \langle A\rangle expr_i) \wedge (\langle A\rangle(expr_j \wedge \neg\pi) \rightarrow \\ & \quad \langle A\rangle(expr_i \wedge \neg\pi)) \wedge (expr_j \rightarrow [B](\neg\pi \rightarrow [A]\neg expr_i^{end}) \wedge \\ & \quad \langle A\rangle expr_k)) \wedge \\ & [G](expr_k \rightarrow \langle A\rangle expr_i^{end} \wedge \\ & \quad [B](\neg\pi \rightarrow [A](\neg expr_i^{end} \wedge \neg expr_j \wedge \neg expr_j^{end})) \wedge \\ & \quad (\pi \leftrightarrow expr_j^{end}) \wedge (\neg\pi \leftrightarrow \langle B\rangle expr_j^{end}) \wedge \\ & \quad (\langle B\rangle expr_i^{end} \rightarrow expr_i)).\end{aligned}$$

- If $e_i = e_j^*$, we distinguish three cases, depending on the number of repetitions of the sub-string matched by e_j in the string matched by e_i, namely zero, one, or more than one. They are encoded by the three disjuncts in the first conjunct of the formula below. The rest of the formula guarantees that every interval on which $expr_j$ holds occurs inside an interval on which $expr_i$ holds.

$$\begin{aligned}\varphi_{expr_i} = {} & [G](expr_i \rightarrow \pi \vee expr_j \vee (\langle B\rangle expr_j \wedge [B](\langle A\rangle expr_j^{end} \rightarrow \\ & \qquad \langle A\rangle(\neg\pi \wedge expr_j)) \wedge \langle A\rangle expr_j^{end})) \wedge \\ & [init](\langle A\rangle expr_j \rightarrow \langle A\rangle expr_i \vee \langle B\rangle\langle A\rangle(\neg\pi \wedge expr_i)) \wedge \\ & [G](\langle A\rangle expr_j \wedge \langle B\rangle(\neg\pi \wedge expr_i) \rightarrow \langle B\rangle(\neg\pi \wedge \langle A\rangle(\neg\pi \wedge expr_i))) \\ & \wedge [G](\langle A\rangle expr_j \wedge \langle A\rangle expr_i^{end} \rightarrow \langle A\rangle expr_i) \\ & \wedge [G](\langle A\rangle(\neg\pi \wedge expr_j) \wedge \langle A\rangle expr_i \rightarrow \langle A\rangle(\neg\pi \wedge expr_i)) \\ & \wedge [G](expr_j \rightarrow [B](\neg\pi \rightarrow [A]\neg expr_i^{end})).\end{aligned}$$

Let φ_R be the formula: $expr_n \wedge [A]\pi \wedge \bigwedge_{i=1}^{n} \varphi_{expr_i} \wedge \bigwedge_{i=1}^{n} \varphi_{expr_i}^{end} \wedge \bigwedge_{i=1}^{n} \varphi_{expr_i}^{\emptyset}$. The following theorem holds [15].

Theorem 4. *Let R be a regular expression over Σ. Then, $\mathcal{L}(R) = \{w \in \Sigma^* : w \approx M$ for some model M such that $M, [0, N] \models \varphi_R \wedge \varphi_\Sigma\}$.*

The encoding of regular expressions can be lifted to ω-regular ones. Let E be an ω-regular expression. We can give it a finite tree structure in the same way we did it for regular ones. As before, we list all the regular and ω-regular sub-expressions $e_1, \ldots, e_n$ in increasing order of complexity with $e_n = E$.

Since we are forced to work with finite intervals, the formula encoding an ω-regular expression intuitively behaves as follows. An ω-regular expression E can be seen as the alternation $(+)$ of a finite number of expressions of the form Re^ω, i.e., $E = R_1 e_1^\omega + \ldots + R_k e_k^\omega$, where, for all i, R_i is regular. The formula encoding the expression $E_i = Re^\omega$ holds true on a certain finite prefix of $\mathbb{N}$, that represents the finite word captured by R, and it uses modality $\langle A\rangle$ to describe properties of the infinite suffix. The encoding of E consists of the disjunction of the formulas encoding the sub-expressions E_i. Formally, the encoding of an ω-regular expression E into an AB formula is inductively defined as follows. As for the regular sub-expressions, we proceed as above. Thus, we only need to specify how to handle the ω-constructor, and alternation and concatenation when one of the operands is an ω-regular expression.

- If $e_i = e_j + e_k$, where e_j and e_k are ω-regular expressions, then $\varphi_{expr_i} = \varphi_{expr_j} \vee \varphi_{expr_k}$.
- If $e_i = e_j e_k$, where e_j is a regular expression and e_k is an ω-regular one, then $\varphi_{expr_i} = \varphi_{expr_j} \wedge \langle A \rangle \varphi_{expr_k}$.
- If $e_i = e_j^{\omega}$, where e_j is a regular expression, then $\varphi_{expr_i} = expr_j \wedge \langle A \rangle(\neg\pi \wedge expr_j) \wedge [A](\langle A \rangle expr_j^{end} \rightarrow \langle A \rangle(\neg\pi \wedge expr_j))$.

Now, let φ_E be the formula: $\bigwedge_{i=1}^{n} \varphi_{expr_i} \wedge \bigwedge_{i=1}^{n} \varphi_{expr_i}^{end} \wedge \bigwedge_{i=1}^{n} \varphi_{expr_i}^{\not\emptyset}$. The following theorem holds [15].

Theorem 5. *Let E be an ω-regular expression over Σ. Then, $\mathcal{L}(E) = \{w \in \Sigma^{\omega} : w \approx M$ for some model M such that $M, [0, n] \models \varphi_E \wedge \varphi_{\Sigma}$ for some $n \in \mathbb{N}\}$.*