Satisfiability of CTL* with constraints*

Claudia Carapelle, Alexander Kartzow, and Markus Lohrey

Institut für Informatik, Universität Leipzig, Germany

Abstract. We show that satisfiability for CTL^* with equality-, order-, and modulo-constraints over $\mathbb{Z}$ is decidable. Previously, decidability was only known for certain fragments of CTL^*, e.g., the existential and positive fragments and EF.

1 Introduction

Temporal logics like LTL, CTL or CTL^* are nowadays standard languages for specifying system properties in model-checking. They are interpreted over node labeled graphs (Kripke structures), where the node labels (also called atomic propositions) represent abstract properties of a system. Clearly, such an abstracted system state does in general not contain all the information of the original system state. Consider for instance a program that manipulates two integer variables x and y. A useful abstraction might be to introduce atomic propositions $v_{-2^{32}}, \ldots, v_{2^{32}}$ for $v \in \{x, y\}$, where the meaning of v_k for $-2^{32} < k < 2^{32}$ is that the variable $v \in \{x, y\}$ currently holds the value k, and $v_{-2^{32}}$ (resp., $v_{2^{32}}$) means that the current value of v is at most -2^{32} (resp., at least 2^{32}). It is evident that such an abstraction might lead to incorrect results in model-checking.

To overcome these problems, extensions of temporal logics with constraints have been studied. Let us explain the idea in the context of LTL. For a fixed relational structure $\mathcal{A}$ (typical examples for $\mathcal{A}$ are number domains like the integers or rationals extended with certain relations) one adds atomic formulas of the form $r(\mathsf{X}^{i_1}x_1, \ldots, \mathsf{X}^{i_k}x_k)$ (so called constraints) to standard LTL. Here, r is (a name of) one of the relations of the structure $\mathcal{A}$, $i_1, \ldots, i_k \geq 0$, and $x_1, \ldots, x_k$ are variables that range over the universe of $\mathcal{A}$. An LTL-formula containing such constraints is interpreted over (infinite) paths of a standard Kripke structure, where in addition every node (state) associates with each of the variables $x_1, \ldots, x_k$ an element of $\mathcal{A}$ (one can think of $\mathcal{A}$-registers attached to the system states). A constraint $r(\mathsf{X}^{i_1}x_1, \ldots, \mathsf{X}^{i_k}x_k)$ holds in a path $s_0 \to s_1 \to s_2 \to \cdots$ if the tuple $(a_1, \ldots, a_k)$, where a_j is the value of variable x_j at state s_{i_j}, belongs to the $\mathcal{A}$-relation r. In this way, the values of variables at different system states can be compared. In our example from the first paragraph, one might choose for $\mathcal{A}$ the structure $(\mathbb{Z}, <, =, (=_a)_{a\in\mathbb{Z}})$, where $=_a$ is the unary predicate that only holds for a. This structure has infinitely many predicates, which is not a problem; our main result will actually talk about an expansion of $(\mathbb{Z}, <, =, (=_a)_{a\in\mathbb{Z}})$. Then, one might for instance write down a formula $(<(x, \mathsf{X}^1 y))\mathsf{U}(=_{100}(y))$ which holds on a path if and only if there is a point of time where variable y holds the value 100 and for all previous points of time t, the value of x at time t is strictly smaller than the value of y at time $t + 1$.

* **Omitted proofs can be found in [4].** This work is supported by the DFG Research Training Group 1763 (QuantLA). The second author is supported by the DFG research project GELO.

In [9], Demri and Gascon studied LTL extended with constraints from a language IPC*. If we disregard succinctness aspects, these constraints are equivalent to constraints over the structure

$$\mathcal{Z} = (\mathbb{Z}, <, =, (=_a)_{a\in\mathbb{Z}}, (\equiv_{a,b})_{0\le a<b}), \tag{1}$$

where $=_a$ denotes the unary relation $\{a\}$ and $\equiv_{a,b}$ denotes the unary relation $\{a + xb \mid x \in \mathbb{Z}\}$ (expressing that an integer is congruent to a modulo b). The main result from [9] states that satisfiability of LTL with constraints from $\mathcal{Z}$ is decidable and in fact PSPACE-complete, and hence has the same complexity as satisfiability for LTL without constraints. We should remark that the PSPACE upper bound from [9] even holds for the succinct IPC*-representation of constraints used in [9].

In the same way as outlined for LTL above, constraints can be also added to CTL and CTL* (then, constraints $r(\mathsf{X}^{i_1}x_1, \ldots, \mathsf{X}^{i_k}x_k)$ are path formulas). A weak form of CTL* with constraints from $\mathcal{Z}$ (where only integer variables and the same state can be compared) was first introduced in [5], where it is used to describe properties of infinite transition systems, represented by relational automata. There it is shown that the model checking problem for CTL* over relational automata is undecidable.

Demri and Gascon [9] asked whether satisfiability of CTL* with constraints from $\mathcal{Z}$ over Kripke structures is decidable. This problem was investigated in [3,10], where several partial results where shown: If we replace in $\mathcal{Z}$ the binary predicate $<$ by unary predicates $<_c = \{x \mid x < c\}$ for $c \in \mathbb{Z}$, then satisfiability for CTL* is decidable by [10]. While, for the full structure $\mathcal{Z}$ satisfiability is decidable for the CTL* fragment CEF$^+$ (which contains the existential and universal fragment of CTL* as well as EF) [3].

In this paper we prove that CTL* with constraints over $\mathcal{Z}$ is decidable. Our proof is divided into two steps. The first step provides a tool to prove decidability of CTL* with constraints over any structure $\mathcal{A}$ over a countable (finite or infinite) signature $\mathcal{S}$ (the structure $\mathcal{A}$ has to satisfy the additional property that the complement of any of its relations has to be definable in positive existential first-order logic over $\mathcal{A}$). Let $\mathcal{L}$ be a logic that satisfies the following two properties: (i) satisfiability of a given $\mathcal{L}$-sentence over the class of infinite node-labeled trees is decidable, and (ii) $\mathcal{L}$ is closed under boolean combinations with monadic second-order formulas (MSO). A typical such logic is MSO itself. By Rabin's seminal tree theorem, satisfiability of MSO-sentences over infinite node-labeled trees is decidable. Assuming $\mathcal{L}$ has these two properties, we prove that satisfiability of CTL* with constraints over $\mathcal{A}$ is decidable if one can compute from a given finite subsignature $\sigma \subseteq \mathcal{S}$ an $\mathcal{L}$-sentence ψ_σ (over the signature σ) such that for every countable σ-structure $\mathcal{B}$: $\mathcal{B} \models \psi_\sigma$ if and only if there exists a homomorphism from $\mathcal{B}$ to $\mathcal{A}$ (i.e., a mapping from the domain of $\mathcal{B}$ to the domain of $\mathcal{A}$ that preserves all relations from σ). We say that the structure $\mathcal{A}$ has the property EHomDef($\mathcal{L}$) if such a computable function $\sigma \mapsto \psi_\sigma$ exists. EHomDefstands for "existence of homomorphism is definable". For instance, the structure $(\mathbb{Q}, <, =)$ has the property EHomDef(MSO), see Example 7.

It is not clear whether $\mathcal{Z}$ from (1) has the property EHomDef(MSO) (we conjecture that it does not). Hence, we need a different logic. It turns out that $\mathcal{Z}$ has the property EHomDef(WMSO+B), where WMSO+B is the extension of weak monadic second-order logic (where only quantification over finite subsets is allowed) with the bounding

quantifier B. A formula $\mathsf{B}X : \varphi$ holds in a structure $\mathcal{A}$ if and only if there exists a bound $b \in \mathbb{N}$ such that for every finite subset B of the domain of $\mathcal{A}$ with $\mathcal{A} \models \varphi(B)$ we have $|B| \leq b$. Recently, Bojańczyk and Toruńczyk have shown that satisfiability of $\mathsf{WMSO}+\mathsf{B}$ over infinite node-labeled trees is decidable [1]. The next problem is that $\mathsf{WMSO}+\mathsf{B}$ is not closed under boolean combinations with MSO-sentences. But fortunately, the decidability proof for $\mathsf{WMSO}+\mathsf{B}$ can be extended to boolean combinations of MSO-sentences and ($\mathsf{WMSO}+\mathsf{B}$)-sentences, see Section 3 for details. This finally shows that satisfiability of CTL^* with constraints from $\mathcal{Z}$ is decidable.

While it would be extremely useful to add successor constraints ($y = x + 1$) to $\mathcal{Z}$, this would lead to undecidability even for LTL [8] and the very basic description logic $\mathcal{ALC}$ [13], which is basically multi-modal logic. Nonetheless $\mathcal{Z}$ allows qualitative representation of increment, for example $x = y + 1$ can be abstracted by $(y > x) \wedge (\equiv_{1,2^k} (y))$ where k is a large natural number. This is why temporal logics extended with constraints over $\mathcal{Z}$ seem to be a good compromise between (unexpressive) total abstraction and (undecidable) high concretion.

In the area of knowledge representation, extensions of description logics with constraints from so called concrete domains have been intensively studied, see [11] for a survey. In [12], it was shown that the extension of the description logic $\mathcal{ALC}$ with constraints from $(\mathbb{Q}, <, =)$ has a decidable ($\mathsf{EXPTIME}$-complete) satisfiability problem with respect to general TBoxes (also known as general concept inclusions). Such a TBox can be seen as a second $\mathcal{ALC}$-formula that has to hold in all nodes of a model. Our decidability proof is partly inspired by the construction from [12], which in contrast to our proof is purely automata-theoretic. Further results for description logics and concrete domains can be found in [13,14].

Unfortunately, our proof does not yield any complexity bound for satisfiability of CTL^* with constraints from $\mathcal{Z}$. The boolean combinations of ($\mathsf{WMSO}+\mathsf{B}$)-sentences and MSO sentences that have to be checked for satisfiability (over infinite trees) are of a simple structure, in particular their quantifier depth is not high. But no complexity statement for satisfiability of $\mathsf{WMSO}+\mathsf{B}$ is made in [1], and it seems to be difficult to analyze the algorithm from [1] (but it seems to be elementary for a fixed quantifier depth). It is based on a construction for cost functions over finite trees from [6], where the authors only note that their construction seems to have very high complexity.

2 Preliminaries

Let $[1, d] = \{1, \ldots, d\}$. For a word $w = a_1 a_2 \cdots a_l \in [1, d]^*$ and $k \leq l$ we define $w[: k] = a_1 a_2 \cdots a_k$; it is the prefix of w of length k.

Let P be a countable set of (atomic) propositions. A Kripke structure over P is a triple $\mathcal{K} = (D, \rightarrow, \rho)$, where (i) D is an arbitrary set of nodes (or states), (ii) $\rightarrow$ is a binary relation on D such that for every $u \in D$ there exists $v \in D$ with $u \rightarrow v$, and (iii) $\rho : D \rightarrow 2^{\mathsf{P}}$ assigns to every node the set of propositions that hold in the node. We require that $\bigcup_{v \in D} \rho(v)$ is finite, i.e., only finitely many propositions appear in $\mathcal{K}$. A *$\mathcal{K}$-path* is an infinite sequence $\pi = (v_0, v_1, v_2, \ldots)$ such that $v_i \rightarrow v_{i+1}$ for all $i \geq 0$. For $i \geq 0$ we define the state $\pi(i) = v_i$ and the path $\pi^i = (v_i, v_{i+1}, v_{i+2}, \ldots)$. A *Kripke d-tree* is a Kripke structure of the form $\mathcal{K} = ([1, d]^*, \rightarrow, \rho)$, where $\rightarrow$ contains all pairs

(u, ui) with $u \in [1,d]^*$ and $1 \leq i \leq d$, i.e., $([1,d]^*, \to)$ is a tree with root ε where every node has d children.

A signature is a countable (finite or infinite) set $\mathcal{S}$ of relation symbols. Every relation symbol $r \in \mathcal{S}$ has an associated arity $\mathsf{ar}(r) \geq 1$. An $\mathcal{S}$-structure is a pair $\mathcal{A} = (A, I)$, where A is a non-empty set and I maps every $r \in \mathcal{S}$ to an $\mathsf{ar}(r)$-ary relation over A. Quite often, we will identify the relation $I(r)$ with the relation symbol r, and we will specify an $\mathcal{S}$-structure as $(A, r_1, r_2, \ldots)$ where $\mathcal{S} = \{r_1, r_2, \ldots\}$. The $\mathcal{S}$-structure $\mathcal{A} = (A, I)$ is *negation-closed* if there exists a computable function that maps a relation symbol $r \in \mathcal{S}$ to a positive existential first-order formula $\varphi_r(x_1, \ldots, x_{\mathsf{ar}(r)})$ (i.e., a formula that is built up from atomic formulas using $\wedge$, $\vee$, and $\exists$) such that $A^{\mathsf{ar}(r)} \setminus I(r) = \{(a_1, \ldots, a_{\mathsf{ar}(r)}) \mid \mathcal{A} \models \varphi_r(a_1, \ldots, a_{\mathsf{ar}(r)})\}$. In other words, the complement of every relation $I(r)$ must be effectively definable by a positive existential first-order formula.

Example 1. The structure $\mathcal{Z}$ from (1) is negation-closed (we will write $x = a$ instead of $=_a(x)$ and similarly for $\equiv_{a,b}$). We have for instance:

- $x \neq y$ if and only if $x < y$ or $y < x$.
- $x \neq a$ if and only if $\exists y \in \mathbb{Z} : y = a \wedge (x < y \vee y < x)$.
- $x \not\equiv a \bmod b$ if and only if $x \equiv c \bmod b$ for some $0 \leq c < b$ with $a \neq c$.

For a subsignature $\sigma \subseteq \mathcal{S}$, a σ-structure $\mathcal{B} = (B, J)$ and an $\mathcal{S}$-structure $\mathcal{A} = (A, I)$, a *homomorphism* $h : \mathcal{B} \to \mathcal{A}$ is a mapping $h : B \to A$ such that for all $r \in \sigma$ and all tuples $(b_1, \ldots, b_{\mathsf{ar}(r)}) \in J(r)$ we have $(h(b_1), \ldots, h(b_{\mathsf{ar}(r)})) \in I(r)$. We write $\mathcal{B} \preceq \mathcal{A}$ if there is a homomorphism from $\mathcal{B}$ to $\mathcal{A}$.

3 MSO and WMSO+B

Recall that *monadic second-order logic* (MSO) is the extension of first-order logic where also quantification over subsets of the underlying structure is allowed. We assume that the reader has some familiarity with MSO. *Weak monadic second-order logic* (WMSO) has the same syntax as MSO but second-order variables only range over finite subsets of the underlying structure. Finally, WMSO+B is the extension of WMSO by the additional quantifier $\mathsf{B}X : \varphi$ (the *bounding quantifier*). The semantics of $\mathsf{B}X : \varphi$ in the structure $\mathcal{A} = (A, I)$ is defined as follows: $\mathcal{A} \models \mathsf{B}X : \varphi(X)$ if and only if there is a bound $b \in \mathbb{N}$ such that $|B| \leq b$ for every finite subset $B \subseteq A$ with $\mathcal{A} \models \varphi(B)$.

Example 2. For later use, we state some example formulas. Let $\varphi(x, y)$ be a WMSO-formula with two free first-order variables x and y. Let $\mathcal{A} = (A, I)$ be a structure and let $E = \{(a, b) \in A \times A \mid \mathcal{A} \models \varphi(a, b)\}$ be the binary relation defined by $\varphi(x, y)$. We define the WMSO-formula $\mathsf{reach}_\varphi(a, b)$ to be

$$\exists X \, \forall Y \big(a \in Y \wedge \forall x \forall y ((x \in Y \wedge y \in X \wedge \varphi(x, y)) \to y \in Y) \to b \in Y \big)$$

It is straightforward to prove that $\mathcal{A} \models \mathsf{reach}_\varphi(a, b)$ if and only if $(a, b) \in E^*$. Note that reach_φ is the standard MSO-formula for reachability but restricted to some finite induced subgraph. Clearly, b is reachable from a in the graph (A, E) if and only if it is in some finite subgraph of (A, E).

Let $\mathsf{ECycle}_\varphi = \exists x\, \exists y(\mathsf{reach}_\varphi(x,y) \wedge \varphi(y,x))$ be the WMSO-formula expressing that there is a cycle in (A,E).

Given a second-order variable Z, we define $\mathsf{reach}^Z_\varphi(a,b)$ to be

$$a \in Z \wedge \forall Y \subseteq Z\big(a \in Y \wedge \forall x \forall y((x \in Y \wedge y \in Z \wedge \varphi(x,y)) \rightarrow y \in Y) \rightarrow b \in Y\big).$$

We have $\mathcal{A} \models \mathsf{reach}^Z_\varphi(a,b)$ iff b is reachable from a in the subgraph of (A,E) induced by the (finite) set Z. Note that $\mathcal{A} \models \mathsf{reach}^Z_\varphi(a,b)$ implies $\{a,b\} \subseteq Z$.

For the next examples we restrict our attention the case that the graph (A,E) defined by $\varphi(x,y)$ is acyclic. Hence, the reflexive transitive closure E^* is a partial order on A. Note that a finite set $F \subseteq A$ is an E-path from $a \in F$ to $b \in F$ if and only if $(F, (E \cap (F \times F))^*)$ is a finite linear order with all elements between a and b. Define the WMSO-formula $\mathsf{Path}_\varphi(a,b,Z)$ as

$$\forall x \in Z\ \forall y \in Z\ (\mathsf{reach}^Z_\varphi(x,y) \vee \mathsf{reach}^Z_\varphi(y,x))\ \wedge \mathsf{reach}^Z_\varphi(a,x) \wedge \mathsf{reach}^Z_\varphi(x,b).$$

For every acyclic (A,E) we have $\mathcal{A} \models \mathsf{Path}_\varphi(a,b,P)$ if and only if P contains exactly the nodes along an E-path from a to b.

We finally define the WMSO+B-formula $\mathsf{BPaths}_\varphi(x,y) = \mathsf{B}Z : \mathsf{Path}_\varphi(x,y,Z)$. By definition of the quantifier B, if (A,E) is acyclic, then $\mathcal{A} \models \mathsf{BPaths}_\varphi(a,b)$ if and only if there is a bound $k \in \mathbb{N}$ on the length of any E-path from a to b.

Next, let $\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}+\mathsf{B})$ be the set of all Boolean combinations of MSO-formulas and (WMSO+B)-formulas. We will use the following result.

Theorem 3 (cf. [1]). *One can decide whether for a given $d \in \mathbb{N}$ and a formula $\varphi \in \mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}+\mathsf{B})$ there exists a Kripke d-tree $\mathcal{K}$ such that $\mathcal{K} \models \varphi$.*

Proof. This theorem follows from results of Bojańczyk and Toruńczyk [1,2]. They introduced puzzles which can be seen as pairs $P = (A,C)$, where A is a parity tree automaton and C is an unboundedness condition C which specifies a certain set of infinite paths labeled by states of A. A puzzle accepts a tree $\mathcal{T}$ if there is an accepting run ρ of A on $\mathcal{T}$ such that for each infinite path π occurring in ρ, $\pi \in C$ holds. In particular, ordinary parity tree automata can be seen as puzzles with trivial unboundedness condition. The proof of our theorem combines the following results.

Lemma 4 ([1]). *From a given (WMSO+B)-formula φ and $d \in \mathbb{N}$ one can construct a puzzle P_φ such that φ is satisfied by some Kripke d-tree iff P_φ is nonempty.*

Lemma 5 ([1]). *Emptiness of puzzles is decidable.*

Lemma 6 (Lemma 17 of [2]). *Puzzles are effectively closed under intersection.*

Let $\varphi \in \mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}+\mathsf{B})$. First, φ can be effectively transformed into a disjunction $\bigvee_{i=1}^n (\varphi_i \wedge \psi_i)$ where $\varphi_i \in \mathsf{MSO}$ and $\psi_i \in \mathsf{WMSO}+\mathsf{B}$ for all i. By Lemma 4, we can construct a puzzle P_i for ψ_i. It is known that the MSO-formula φ_i can be translated into a parity tree automaton A_i. Let P_i' be a puzzle recognizing the intersection of P_i and A_i (cf. Lemma 6). Now φ is satisfiable over Kripke d-trees if and only if there is an i such that $\varphi_i \wedge \psi_i$ is satisfiable over Kripke d-trees if and only if there is an i such that P_i' is nonempty. By Lemma 5, the latter condition is decidable which concludes the proof of the theorem. □

Let $\mathcal{L}$ be a logic (e.g. MSO or $\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B})$). An $\mathcal{S}$-structure $\mathcal{A}$ has the property $\mathsf{EHomDef}(\mathcal{L})$ (existence of homomorphisms to $\mathcal{A}$ is $\mathcal{L}$-definable) if there is a computable function that maps a finite subsignature $\sigma \subseteq \mathcal{S}$ to an $\mathcal{L}$-sentence φ_σ such that for every countable σ-structure $\mathcal{B}$: $\mathcal{B} \preceq \mathcal{A}$ if and only if $\mathcal{B} \models \varphi_\sigma$.

Example 7. The structure $\mathcal{Q} = (\mathbb{Q}, <, =)$ has the property $\mathsf{EHomDef}(\mathsf{WMSO})$ (and $\mathsf{EHomDef}(\mathsf{MSO})$). In [12] it is implicitly shown that for a countable $\{<, =\}$-structure $\mathcal{B} = (B, I)$, $\mathcal{B} \preceq \mathcal{Q}$ if and only if there does not exist $(a, b) \in I(<)$ such that $(b, a) \in (I(<) \cup I(=) \cup I(=)^{-1})^*$. This condition can be easily expressed in WMSO using the reach-construction from Example 2. Note that $I(=)$ is not required to be the identity relation on B.

4 CTL^* with constraints

Let us fix a countably infinite set of atomic propositions P and a countably infinite set of variables V for the rest of the paper. Let $\mathcal{S}$ be a signature. We define an extension of CTL^* with constraints over the signature $\mathcal{S}$. We define $\mathsf{CTL}^*(\mathcal{S})$-state formulas φ and $\mathsf{CTL}^*(\mathcal{S})$-path formulas ψ by the following grammar, where $p \in \mathsf{P}$, $r \in \mathcal{S}$, $k = \mathsf{ar}(r)$, $i_1, \ldots, i_k \geq 0$, and $x_1, \ldots, x_k \in \mathsf{V}$:

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \mathsf{E}\psi$$
$$\psi ::= \varphi \mid \neg\psi \mid (\psi \wedge \psi) \mid \mathsf{X}\psi \mid \psi\mathsf{U}\psi \mid r(\mathsf{X}^{i_1}x_1, \ldots, \mathsf{X}^{i_k}x_k)$$

A formula of the form $R := r(\mathsf{X}^{i_1}x_1, \ldots, \mathsf{X}^{i_k}x_k)$ is also called an *atomic constraint* and we define $d(R) = \max\{i_1, \ldots, i_k\}$ (the depth of R). The syntactic difference between $\mathsf{CTL}^*(\mathcal{S})$ and ordinary CTL^* lies in the presence of atomic constraints.

Formulas of $\mathsf{CTL}^*(\mathcal{S})$ are interpreted over triples $\mathcal{C} = (\mathcal{A}, \mathcal{K}, \gamma)$, where $\mathcal{A} = (A, I)$ is an $\mathcal{S}$-structure (also called the *concrete domain*), $\mathcal{K} = (D, \to, \rho)$ is a Kripke structure over P, and $\gamma : D \times \mathsf{V} \to A$ assigns to every $(v, x) \in D \times \mathsf{V}$ a value $\gamma(v, x)$ (the value of variable x at node v). We call such a triple $\mathcal{C} = (\mathcal{A}, \mathcal{K}, \gamma)$ an *$\mathcal{A}$-constraint graph*. An $\mathcal{A}$-constraint graph $\mathcal{C} = (\mathcal{A}, \mathcal{K}, \gamma)$ is an *$\mathcal{A}$-constraint d-tree* if $\mathcal{K}$ is a Kripke d-tree.

We now define the semantics of $\mathsf{CTL}^*(\mathcal{S})$. For an $\mathcal{A}$-constraint graph $\mathcal{C} = (\mathcal{A}, \mathcal{K}, \gamma)$ with $\mathcal{A} = (A, I)$ and $\mathcal{K} = (D, \to, \rho)$, a state $v \in D$, a $\mathcal{K}$-path π, a state formula φ, and a path formula ψ we write $(\mathcal{C}, v) \models \varphi$ if φ holds in $(\mathcal{C}, v)$ and $(\mathcal{C}, \pi) \models \psi$ if ψ holds in $(\mathcal{C}, \pi)$. This is inductively defined as follows (for the boolean connectives $\neg$ and $\wedge$ the definitions are as usual and we omit them):

- $(\mathcal{C}, v) \models p$ iff $p \in \rho(v)$.
- $(\mathcal{C}, v) \models \mathsf{E}\psi$ iff there is a $\mathcal{K}$-path π with $\pi(0) = v$ and $(\mathcal{C}, \pi) \models \psi$.
- $(\mathcal{C}, \pi) \models \varphi$ iff $(\mathcal{C}, \pi(0)) \models \varphi$.
- $(\mathcal{C}, \pi) \models \mathsf{X}\psi$ iff $(\mathcal{C}, \pi^1) \models \psi$.
- $(\mathcal{C}, \pi) \models \psi_1\mathsf{U}\psi_2$ iff there exists $i \geq 0$ such that $(\mathcal{C}, \pi^i) \models \psi_2$ and for all $0 \leq j < i$ we have $(\mathcal{C}, \pi^j) \models \psi_1$.
- $(\mathcal{C}, \pi) \models r(\mathsf{X}^{i_1}x_1, \ldots, \mathsf{X}^{i_n}x_n)$ iff $(\gamma(\pi(i_1), x_1), \ldots, \gamma(\pi(i_n), x_n)) \in I(r)$.

Note that the role of the concrete domain $\mathcal{A}$ and of the valuation function γ is restricted to the semantic of atomic constraints. CTL^*-formulas are interpreted over Kripke structures, and to obtain their semantics it is sufficient to replace $\mathcal{C}$ by $\mathcal{K}$ in the rules above and to remove the last line.

We use the usual abbreviations: $\theta_1 \vee \theta_2 := \neg(\neg\theta_1 \wedge \neg\theta_2)$ (for both state and path formulas), $\mathsf{A}\psi := \neg\mathsf{E}\neg\psi$ (universal path quantifier), $\psi_1 \mathsf{R} \psi_2 := \neg(\neg\psi_1 \mathsf{U} \neg\psi_2)$ (the release operator). Note that $(\mathcal{C}, \pi) \models \psi_1 \mathsf{R} \psi_2$ iff ($(\mathcal{C}, \pi^i) \models \psi_2$ for all $i \geq 0$ or there exists $i \geq 0$ such that $(\mathcal{C}, \pi^i) \models \psi_1$ and $(\mathcal{C}, \pi^j) \models \psi_2$ for all $0 \leq j \leq i$).

Using this extended set of operators we can put every formula into a semantically equivalent *negation normal form*, where $\neg$ only occurs in front of atomic propositions or atomic constraints. Let $\#_{\mathsf{E}}(\theta)$ be the the number of different subformulas of the form $\mathsf{E}\psi$ in the $\mathsf{CTL}^*(\mathcal{S})$-formula θ. Then $\mathsf{CTL}^*(\mathcal{S})$ has the following tree model property:

Theorem 8 (cf. [10]). *Let φ be a $\mathsf{CTL}^*(\mathcal{S})$-state formula in negation normal form and let $\mathcal{A} = (A, I)$ be an $\mathcal{S}$-structure. Then φ is $\mathcal{A}$-satisfiable if and only if there exists an $\mathcal{A}$-constraint $(\#_{\mathsf{E}}(\varphi) + 1)$-tree $\mathcal{C}$ with $(\mathcal{C}, \varepsilon) \models \varphi$.*

Note that for checking $(\mathcal{A}, \mathcal{K}, \gamma) \models \varphi$ we may ignore all propositions $p \in \mathsf{P}$ that do not occur in φ. Similarly, only those values $\gamma(u, x)$, where x is a variable that appears in φ, are relevant. Hence, if V_φ is the finite set of variables that occur in φ, then we can consider γ as a mapping from $D \times \mathsf{V}_\varphi$ to the domain of $\mathcal{A}$. Intuitively, we assign to each node $u \in D$ registers that store the values $\gamma(u, x)$ for $x \in \mathsf{V}_\varphi$.

5 Satisfiability of constraint CTL^* over a concrete domain

When we talk about satisfiability for $\mathsf{CTL}^*(\mathcal{S})$ our setting is as follows: We fix a concrete domain $\mathcal{A} = (A, I)$. Given a $\mathsf{CTL}^*(\mathcal{S})$-state formula φ, we say that φ is $\mathcal{A}$-satisfiable if there is an $\mathcal{A}$-constraint graph $\mathcal{C} = (\mathcal{A}, \mathcal{K}, \gamma)$ and a node v of $\mathcal{K}$ such that $(\mathcal{C}, v) \models \varphi$. With $\mathsf{SAT}\mathsf{CTL}^*(\mathcal{A})$ we denote the following computational problem: *Is a given state formula $\varphi \in \mathsf{CTL}^*(\mathcal{S})$ $\mathcal{A}$-satisfiable?* The main result of this section is:

Theorem 9. *Let $\mathcal{A}$ be a negation-closed $\mathcal{S}$-structure, which moreover has the property $\mathsf{EHomDef}(\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B}))$. Then the problem $\mathsf{SAT}\mathsf{CTL}^*(\mathcal{A})$ is decidable.*

We say that a $\mathsf{CTL}^*(\mathcal{S})$-formula φ is in *strong negation normal form* if negations only occur in front of atomic propositions (i.e., φ is in negation normal form and there is no subformula $\neg R$ where R is an atomic constraint).

Let us fix a $\mathsf{CTL}^*(\mathcal{S})$-state formula φ in negation normal form and a negation-closed $\mathcal{S}$-structure $\mathcal{A}$ for the rest of this section. We want to check whether φ is $\mathcal{A}$-satisfiable. First, we reduce to formulas in strong negation normal form:

Lemma 10. *Let $\mathcal{A} = (A, I)$ be a negation-closed $\mathcal{S}$-structure. From a given $\mathsf{CTL}^*(\mathcal{S})$-state formula φ one can compute a $\mathsf{CTL}^*(\mathcal{S})$-state formula $\hat{\varphi}$ in strong negation normal form such that φ is $\mathcal{A}$-satisfiable iff $\hat{\varphi}$ is $\mathcal{A}$-satisfiable.*

From now on let us assume that φ is in strong negation normal form. Let $d = \#_{\mathsf{E}}(\varphi)+1$. Let $R_1, \ldots, R_n$ be a list of all atomic constraints that are subformulas of φ, and let V_φ

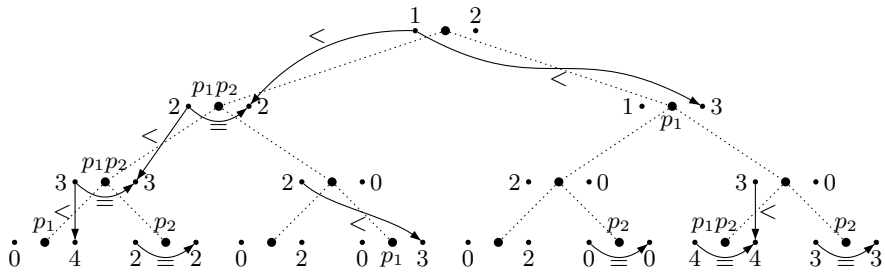

Fig. 1. The $(\mathbb{N}, <, =)$-constraint 2-tree $\mathcal{C}$ from Ex. 11, the Kripke 2-tree $\mathcal{T} = \mathcal{C}^a$, and the structure $\mathcal{G}_\mathcal{T}$.

be the finite set of variables that occur in φ. Let us fix new propositions $p_1, \ldots, p_n$ (one for each R_i) that do not occur in φ. Let $d_i = d(R_i)$ be the depth of the constraint R_i. We denote with φ^a the (ordinary) CTL^*-formula obtained from φ by replacing every occurrence of a constraint R_i by $\mathsf{X}^{d_i} p_i$. Given an $\mathcal{A}$-constraint d-tree $\mathcal{C} = (\mathcal{A}, \mathcal{K}, \gamma)$, where $\mathcal{K} = ([1,d]^*, \to, \rho)$ and $\rho(v) \cap \{p_1, \ldots, p_n\} = \emptyset$ for all $v \in [1,d]^*$, we define a Kripke d-tree $\mathcal{C}^a = ([1,d]^*, \to, \rho^a)$, where $\rho^a(v)$ contains

- all propositions from $\rho(v)$ and
- all propositions p_i ($1 \leq i \leq n$) such that the following holds, where we assume that R_i has the form $r(\mathsf{X}^{j_1} x_1, \ldots, \mathsf{X}^{j_k} x_k)$ with $k = \mathsf{ar}(r)$ (hence, $d_i = \max\{j_1, \ldots, j_k\}$):
  - $v = su$ with $|u| = d_i$
  - $(\gamma(su_1, x_1), \ldots, \gamma(su_k, x_k)) \in I(r)$, where $u_l = u[: j_l]$ for $1 \leq l \leq k$.

Hence, the fact that proposition p_i labels node su with $|u| = d_i$ means that the constraint R_i holds along every path that starts in node s and descends in the tree down via node su. The superscript "a" in $\mathcal{C}^a$ stands for "abstracted" since we abstract from the concrete constraints and replace them by new propositions.

Moreover, given a Kripke d-tree $\mathcal{T} = ([1,d]^*, \to, \rho)$ (where the new propositions $p_1, \ldots, p_n$ are allowed to occur in $\mathcal{T}$) we define a countable $\mathcal{S}$-structure $\mathcal{G}_\mathcal{T} = ([1,d]^* \times \mathsf{V}_\varphi, J)$ as follows: The interpretation $J(r)$ of the relation symbol $r \in \mathcal{S}$ contains all k-tuples (where $k = \mathsf{ar}(r)$) $((su_1, x_1), \ldots, (su_k, x_k))$ for which there exist $1 \leq i \leq n$ and $u \in [1,d]^*$ with $|u| = d_i$ such that $p_i \in \rho(su)$, $R_i = r(\mathsf{X}^{j_1} x_1, \ldots, \mathsf{X}^{j_k} x_k)$, and $u_t = u[: j_t]$ for $1 \leq t \leq k$.

Example 11. Figure 1 shows an example, where we assume that $d = 2$ and $n = 2$, $R_1 = [<(x_1, \mathsf{X}x_2)]$, and $R_2 = [=(\mathsf{X}x_1, \mathsf{X}x_2)]$. The figure shows an initial part of an $(\mathbb{N}, <, =)$-constraint 2-tree $\mathcal{C} = ((\mathbb{N}, <, =), \mathcal{K}, \gamma)$. The edges of the Kripke 2-tree $\mathcal{K}$ are dotted. We assume that $\mathcal{K}$ is defined over the empty set of propositions. The node to the left (resp., right) of a tree node u is labeled by the value $\gamma(u, x_1)$ (resp. $\gamma(u, x_2)$). The figure shows the labeling of tree nodes with the two new propositions p_1 and p_2 (corresponding to R_1 and R_2) as well as the $\{<, =\}$-structure $\mathcal{G}_\mathcal{T}$ for $\mathcal{T} = \mathcal{C}^a$.

Lemma 12. *Let φ be a $\mathsf{CTL}^*(\mathcal{S})$-state formula in strong negation normal form. The formula φ is $\mathcal{A}$-satisfiable if and only if there exists a Kripke $(\#_{\mathsf{E}}(\varphi)+1)$-tree $\mathcal{T}$ such that $(\mathcal{T},\varepsilon) \models \varphi^a$ and $\mathcal{G}_\mathcal{T} \preceq \mathcal{A}$.*

Let $\theta = \varphi^a$ for the further discussion. Hence, θ is an ordinary CTL^*-state formula, where negations only occur in front of propositions from $\mathsf{P} \setminus \{p_1,\ldots,p_m\}$, and $d = \#_{\mathsf{E}}(\theta)+1$. By Lemma 12, we have to check, whether there exists a Kripke d-tree $\mathcal{T}$ such that $(\mathcal{T},\varepsilon) \models \theta$ and $\mathcal{G}_\mathcal{T} \preceq \mathcal{A}$.

Let $\sigma \subseteq \mathcal{S}$ be the finite subsignature consisting of all predicate symbols that occur in our initial $\mathsf{CTL}^*(\mathcal{S})$-formula φ. Note that $\mathcal{G}_\mathcal{T}$ is actually a σ-structure. Since the concrete domain $\mathcal{A}$ has the property $\mathsf{EHomDef}(\mathsf{Bool}(\mathsf{MSO},\mathsf{WMSO}{+}\mathsf{B}))$, one can compute from σ a $\mathsf{Bool}(\mathsf{MSO},\mathsf{WMSO}{+}\mathsf{B})$-formula α such that for every countable σ-structure $\mathcal{B}$ we have $\mathcal{B} \models \alpha$ if and only if $\mathcal{B} \preceq \mathcal{A}$. Hence, our new goal is to decide, whether there exists a Kripke d-tree $\mathcal{T}$ such that $(\mathcal{T},\varepsilon) \models \theta$ and $\mathcal{G}_\mathcal{T} \models \alpha$ (note that $\mathcal{G}_\mathcal{T}$ is countable). It is well known that every CTL^*-state formula can be effectively transformed into an equivalent MSO-formula with a single free first-order variable. Since the root ε of a tree is first-order definable, we get an MSO-sentence ψ such that $(\mathcal{T},\varepsilon) \models \theta$ if and only if $\mathcal{T} \models \psi$. Hence, we have to check whether there exists a Kripke d-tree $\mathcal{T}$ such that $\mathcal{T} \models \psi$ and $\mathcal{G}_\mathcal{T} \models \alpha$. If we can translate the $\mathsf{Bool}(\mathsf{MSO},\mathsf{WMSO}{+}\mathsf{B})$-formula α back into a $\mathsf{Bool}(\mathsf{MSO},\mathsf{WMSO}{+}\mathsf{B})$-formula α' such that ($\mathcal{G}_\mathcal{T} \models \alpha \Leftrightarrow \mathcal{T} \models \alpha'$), then we can finish the proof.

Recall the construction of $\mathcal{G}_\mathcal{T}$: For every node $v \in D$ of $\mathcal{T} = (D,\to,\rho)$ we introduce $m := |\mathsf{V}_\varphi|$ copies (v,x) for $x \in \mathsf{V}_\varphi$. The $\mathcal{S}$-relations between these nodes are determined by the propositions $p_1,\ldots,p_n$: The interpretation of $r \in \mathcal{S}$ contains all k-tuples ($k = \mathsf{ar}(r)$) $((su_1,y_1),\ldots,(su_k,y_k))$ for which there exist $1 \le i \le n$ and $u \in [1,d]^*$ with $|u| = d_i$, $p_i \in \rho(su)$, $R_i = r(\mathsf{X}^{j_1}y_1,\ldots,\mathsf{X}^{j_k}y_k)$, and $u_t = u[{:}\ j_t]$ for $1 \le t \le k$. This is a particular case of an MSO-transduction [7] with copy number m. It is therefore possible to compute from a given MSO-sentence η over the signature $\mathcal{S}$ an MSO-sentence η' such that $\mathcal{G}_\mathcal{T} \models \eta \Leftrightarrow \mathcal{T} \models \eta'$. But the problem is that in our situation η is the $\mathsf{Bool}(\mathsf{MSO},\mathsf{WMSO}{+}\mathsf{B})$-formula α, and it is not clear whether MSO-transductions (or even first-order interpretations) are compatible with the logic $\mathsf{WMSO}{+}\mathsf{B}$. Nevertheless, there is a simple solution. Let $\mathsf{V}_\varphi = \{x_1,\ldots,x_m\}$. From a Kripke d-tree $\mathcal{T} = ([1,d]^*,\to,\rho)$ we build an extended $(d+m)$-Kripke tree $\mathcal{T}^e = ([1,d+m]^*,\to,\rho^e)$ as follows: Let us fix new propositions $q_1,\ldots,q_m$ (one for each variable x_i) that do not occur in the MSO-sentence ψ and such that $\rho(v) \cap \{q_1,\ldots,q_m\} = \emptyset$ for all $v \in [1,d]^*$. We define the new labeling function ρ^e as follows:

$$\begin{aligned}
\rho^e(v) &= \rho(v) \text{ for } v \in [1,d]^* \\
\rho^e(vi) &= \{q_{i-d}\} \text{ for } v \in [1,d]^*, d+1 \le i \le d+m \\
\rho^e(viu) &= \emptyset \text{ for } v \in [1,d]^*, d+1 \le i \le d+m, u \in [1,d+m]^+
\end{aligned}$$

It is easy to write down an MSO-sentence β such that for every $(d+m)$-Kripke tree $\mathcal{T}'$ we have $\mathcal{T}' \models \beta$ if and only if $\mathcal{T}' \cong \mathcal{T}^e$ for some Kripke d-tree $\mathcal{T}$. Moreover, since the old Kripke d-tree $\mathcal{T}$ is MSO-definable within $\mathcal{T}^e$, we can construct from the MSO-sentence ψ a new MSO-sentence ψ^e such that $\mathcal{T} \models \psi$ if and only if $\mathcal{T}^e \models \psi^e$. Finally, let $q(x) = \bigvee_{i=1}^m q_i(x)$. Then, the nodes of $\mathcal{G}_\mathcal{T}$ are in a natural bijection with

the nodes of $\mathcal{T}^e$ that satisfy $q(x)$: If $\mathcal{T}^e \models q(u)$ for $u \in [1, d+m]^*$, then there is a unique $i \in [1, m]$ such that $\mathcal{T}^e \models q_i(u)$ and $u = v(i+d)$. Then we associate the node u with node (v, x_i) of $\mathcal{G}_\mathcal{T}$. By relativizing all quantifiers in the $\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B})$-formula α to $q(x)$, we can construct a $\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B})$-formula α^e such that $\mathcal{G}_\mathcal{T} \models \alpha$ if and only if $\mathcal{T}^e \models \alpha^e$.

It follows that there is a Kripke d-tree $\mathcal{T}$ such that $\mathcal{T} \models \psi$ and $\mathcal{G}_\mathcal{T} \models \alpha$ if and only if there is a Kripke $(d+m)$-tree $\mathcal{T}'$ such that $\mathcal{T}' \models (\beta \wedge \psi^e \wedge \alpha^e)$. Since $\beta \wedge \psi^e \wedge \alpha^e$ is a $\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B})$-formula, the latter is decidable by Thm. 3.

6 Concrete domains over the integers

The main technical result of this section is:

Proposition 13. *$\mathcal{Z}$ from* (1) *has the property* $\mathsf{EHomDef}(\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B}))$.

Since $\mathcal{Z}$ is negation-closed (see Ex. 1) our main result follows by Thm. 9:

Theorem 14. $\mathsf{SATCTL}^*(\mathcal{Z})$ *is decidable.*

We prove Prop. 13 in three steps. First, we show that the structure $(\mathbb{Z}, <)$ has the property $\mathsf{EHomDef}(\mathsf{WMSO}{+}\mathsf{B})$. Then we extend this result to the structure $(\mathbb{Z}, <, =)$ and, finally, to the full structure $\mathcal{Z}$.

Proposition 15. *$(\mathbb{Z}, <)$ has the property* $\mathsf{EHomDef}(\mathsf{WMSO}{+}\mathsf{B})$.

As a preparation of the proof, we first define some terminology and then we characterize structures that allow homomorphisms to $(\mathbb{Z}, <)$ in terms of their paths. Let $\mathcal{A} = (A, I)$ be a countable $\{<\}$-structure. We identify $\mathcal{A}$ with the directed graph (A, E) where $E = I(<)$. When talking about paths, we always refer to finite directed E-paths. The length of a path $(a_0, a_1, \ldots, a_n)$ (i.e., $(a_{i-1}, a_i) \in E$ for $1 \leq i \leq n$) is n. For $S \subseteq A$ and $x \in A \setminus S$, a path from x to S is a path from x to some node $y \in S$. A path from S to x is defined in a symmetric way.

Lemma 16. *We have $\mathcal{A} \preceq (\mathbb{Z}, <)$ if and only if*

(H1) $\mathcal{A}$ does not contain cycles, and
(H2) for all $a, b \in A$ there is $c \in \mathbb{N}$ such that the length of all paths from a to b is bounded by c.

Proof. Let us first show the "only if" direction of the lemma. Suppose h is a homomorphism from $\mathcal{A}$ to $(\mathbb{Z}, <)$. The presence of a cycle $(a_0, \ldots a_{k-1})$ in $\mathcal{A}$ ($k \geq 1$, $(a_i, a_{i+1 \bmod k}) \in E$ for $0 \leq i \leq k-1$) would imply the existence of integers $z_0, \ldots z_{k-1}$ with $z_i < z_{i+1 \bmod k}$ for $0 \leq i \leq k-1$ (where $z_i = h(a_i)$), which is not possible. Hence, (H1) holds.

Suppose now that $a, b \in A$ are such that for every n there is a path of length at least n from a to b. If $d = h(b) - h(a)$, we can find a path $(a_0, a_1 \ldots, a_k)$ with $a_0 = a$, $a_k = b$ and $k > d$. Since h is a homomorphism, this path will be mapped to an increasing sequence of integers $h(a) = h(a_0) < h(a_1) < \cdots < h(a_k) = h(b)$. But this contradicts $h(b) - h(a) = d < k$. Hence, (H2) holds.

For the "if" direction of the lemma assume that $\mathcal{A}$ is acyclic (property (H1)) and that (H2) holds. Fix an enumeration $a_0, a_1, a_2, \ldots$ of the countable set A. For $n \geq 0$ let $S_n := \{a \in A \mid \exists i, j \leq n : (a_i, a), (a, a_j) \in E^*\}$, which has the following properties:

(P1) S_n is convex w.r.t. the partial order E^*: If $a, c \in S_n$ and $(a, b), (b, c) \in E^*$, then $b \in S_n$.
(P2) For $a \in A \setminus S_n$ all paths between a and S_n are "one-way", i.e., there do not exist $b, c \in S_n$ such that $(b, a), (a, c) \in E^*$. This follows from (P1).
(P3) For all $a \in A \setminus S_n$ there exists a bound $c \in \mathbb{N}$ such that all paths between a and S_n have length at most c. Let $c_n^a \in \mathbb{N}$ be the smallest such bound (hence, we have $c_n^a = 0$ if there do not exist paths between a and S_n).

To see (P3), assume that there only exist paths from S_n to a but not the other way round (see (P2)); the other case is symmetric. If there is no bound on the length of paths from S_n to a, then by definition of S_n, there is no bound on the length of paths from $\{a_0, \ldots, a_n\}$ to a. By the pigeon principle, there exists $0 \leq i \leq n$ such that there is no bound on the length of paths from a_i to a. But this contradicts property (H2).

We build our homomorphism h inductively. For every $n \geq 0$ we define functions $h_n : S_n \to \mathbb{Z}$ such that the following invariants hold for all $n \geq 0$.

(I1) If $n > 0$ then $h_n(a) = h_{n-1}(a)$ for all $a \in S_{n-1}$
(I2) $h_n(S_n)$ is bounded in $\mathbb{Z}$, i.e., there exist $z_1, z_2 \in \mathbb{Z}$ such that $h_n(S_n) \subseteq [z_1, z_2]$.
(I3) h_n is a homomorphism from the subgraph $(S_n, E \cap (S_n \times S_n))$ to $(\mathbb{Z}, <)$.

For $n = 0$ we have $S_0 = \{a_0\}$. We set $h_0(a_0) = 0$ (any other integer would be also fine). Properties (I1)–(I3) are easily verified. For $n > 0$, there are four cases.

Case 1. $a_n \in S_{n-1}$, thus $S_n = S_{n-1}$. We set $h_n = h_{n-1}$. Clearly, (I1)–(I3) hold for n.

Case 2. $a_n \notin S_{n-1}$ and there is no path from a_n to S_{n-1} or vice versa. We set $h_n(a_n) := 0$ (and $S_n = S_{n-1} \cup \{a_n\}$). In this case (I1)–(I3) follow easily from the induction hypothesis.

Case 3. $a_n \notin S_{n-1}$ and there exist paths from a_n to S_{n-1}. Then, by (P2) there do not exist paths from S_{n-1} to a_n. Hence, we have

$$S_n = S_{n-1} \cup \{a \in A \mid \exists b \in S_{n-1} : (a_n, a), (a, b) \in E^*\}.$$

We have to assign a value $h_n(a)$ for all $a \in A \setminus S_{n-1}$ that lie along a path from a_n to S_{n-1}. By (I2) there exist $z_1, z_2 \in \mathbb{Z}$ with $h_{n-1}(S_{n-1}) \subseteq [z_1, z_2]$. Recall the definition of c_{n-1}^a from (P3). For all $a \in A \setminus S_{n-1}$ that lie on a path from a_n to S_{n-1}, we set $h_n(a) := z_1 - c_{n-1}^a$. Since there are paths from a to S_{n-1}, we have $c_{n-1}^a > 0$. Hence, for all $a \in S_n \setminus S_{n-1}$, $h_n(a) < z_1$. Let us check that $h_n : S_n \to \mathbb{Z}$ satisfy (I1)– (I3): Invariant (I1) holds by definition of h_n. For (I2) note that $h_n(S_n) \subseteq [z_1 - c_{n-1}^{a_n}, z_2]$.

It remains to show (I3), i.e., that h_n is a homomorphism from $(S_n, E \cap (S_n \times S_n))$ to $(\mathbb{Z}, <)$. Hence, we have to show that $h(b_1) < h(b_2)$ for all $(b_1, b_2) \in E \cap (S_n \times S_n)$.

– If $b_1, b_2 \in S_{n-1}$, then $h_n(b_1) = h_{n-1}(b_1) < h_{n-1}(b_2) = h_n(b_2)$ by induction hypothesis.
– If $b_1 \in S_n \setminus S_{n-1}$ and $b_2 \in S_{n-1}$, we know that $h_n(b_2) = h_{n-1}(b_2) \geq z_1$ while $h_n(b_1) < z_1$ by construction. This directly implies $h_n(b_1) < h_n(b_2)$.

– If $b_2 \in S_n \setminus S_{n-1}$ and $b_1 \in S_{n-1}$, then $(b_1, b_2) \in E$ and by assumption b_2 must be on a path from a_n to S_{n-1} which contradicts (P2).
– If both b_1 and b_2 belong to $S_n \setminus S_{n-1}$ then $h_n(b_i) := z_1 - c_{n-1}^{b_i}$ for $i \in \{1, 2\}$ Since $(b_1, b_2) \in E$, we have $c_{n-1}^{b_1} > c_{n-1}^{b_2}$. This implies $h_n(b_1) < h_n(b_2)$.

Case 4. $a_n \notin S_{n-1}$ and there exist paths from S_{n-1} to a_n. For all $a \in S_n \setminus S_{n-1} = \{a \in A \setminus S_{n-1} \mid a \text{ belongs to a path from } S_{n-1} \text{ to } a_n\}$, set $h_n(a) = z_2 + c_{n-1}^a$. The rest of the argument goes analogously to Case 3.

This concludes the construction of h_n. By (I1) limit function $h = \bigcup_{i \in \mathbb{N}} h_i$ exists. By (I3) and $A = \bigcup_{i \in \mathbb{N}} S_i$, h is a homomorphism from $\mathcal{A}$ to $(\mathbb{Z}, <)$. □

Proof of Prop. 15. We translate the conditions (H1) and (H2) from Lemma 16 into $\mathsf{WMSO}+\mathsf{B}$. Cycles are excluded by the sentence $\neg\mathsf{ECycle}_<$ (Example 2). Moreover, for an acyclic $\{<\}$-structure $\mathcal{A}$ we have $\mathcal{A} \models \forall x \forall y\ \mathsf{BPaths}_<(x, y)$ (see also Example 2) if and only if for all $a, b \in A$ there is a bound $b \in \mathbb{N}$ on the length of paths from a to b. Thus, $\mathcal{A} \preceq (\mathbb{Z}, <)$ if and only if $\mathcal{A} \models \neg\mathsf{ECycle}_< \wedge \forall x \forall y\ \mathsf{BPaths}_<(x, y)$. □

Next, we extend Prop. 15 to the negation-closed structure $(\mathbb{Z}, <, =)$. To do so let us fix a countable $\{<, =\}$-structure $\mathcal{A} = (A, I)$. Note that $I(=)$ is not necessarily the identity relation on A. Let $\sim = (I(=) \cup I(=)^{-1})^*$ be the smallest equivalence relation on A that contains $I(=)$. Since $\sim$ is the reflexive and transitive closure of the first-order definable relation $I(=) \cup I(=)^{-1}$, we can construct a WMSO-formula $\tilde{\varphi}(x, y)$ (using the reach-construction from Ex. 2) that defines $\sim$. Let

$$E_< = \sim \circ\, I(<) \circ \sim \text{ i.e., the relation defined by the formula} \tag{2}$$

$$\varphi_<(x, y) = \exists u\ \exists v\ (\tilde{\varphi}(x, u) \wedge u < v \wedge \tilde{\varphi}(v, y)). \tag{3}$$

With $\tilde{\mathcal{A}} = (\tilde{A}, \tilde{I})$ we denote the $\sim$-quotient of $\mathcal{A}$: It is a $\{<\}$-structure, its domain is the set $\tilde{A} = \{[a]_\sim \mid a \in A\}$ of all $\sim$-equivalence classes, and for two equivalence classes $[a]_\sim$ and $[b]_\sim$ we have $([a]_\sim, [b]_\sim) \in \tilde{I}(<)$ iff there are $a' \sim a$ and $b' \sim b$ such that $(a', b') \in I(<)$. Let us write $[a]$ for $[a]_\sim$. We have:

Lemma 17. *$\mathcal{A} \preceq (\mathbb{Z}, <, =)$ if and only $\tilde{\mathcal{A}} \preceq (\mathbb{Z}, <)$.*

In the next lemma, we translate the conditions for the existence of a homomorphism from $\tilde{\mathcal{A}}$ to $(\mathbb{Z}, <)$ into conditions in terms of $\mathcal{A}$.

Lemma 18. *The following conditions are equivalent:*

– *$\tilde{\mathcal{A}}$ satisfies the conditions (H1) and (H2) from Lemma 16.*
– *The graph $(A, E_<)$ is acyclic and for all $a, b \in A$ there is a bound $c \in \mathbb{N}$ such that all $E_<$-paths from a to b have length at most c.*

Proposition 19. *The concrete domains $(\mathbb{Z}, <, =)$, $(\mathbb{N}, <, =)$ and $(\mathbb{Z} \setminus \mathbb{N}, <, =)$ have property $\mathsf{EHomDef}(\mathsf{WMSO}+\mathsf{B})$.*

Proof. We only proof the proposition for $(\mathbb{Z}, <, =)$. The other two cases are similar. We want to find a $(\mathsf{WMSO}+\mathsf{B})$-formula φ such that for all $\{<, =\}$-structures $\mathcal{A}$, $\mathcal{A} \models \varphi$ if and only if $\mathcal{A} \preceq (\mathbb{Z}, <, =)$. Let $\mathcal{A} = (A, I)$ be a $\{<, =\}$-structure. We use the

notations introduced before Lemma 17. By Lemma 17 and 18 we have to construct a $(\mathsf{WMSO}{+}\mathsf{B})$-formula expressing that $\mathcal{A}$ has no $E_<$-cycles and for all $a, b \in A$ there is a bound $c \in \mathbb{N}$ on the length of $E_<$-paths from a to b. For this, we can use the formula constructed in the proof of Prop. 15 with $<$ replaced by the formula $\varphi_<$ from (3). $\square$

In the rest of this section, we prove Prop. 19 for the full structure $\mathcal{Z}$ from (1), which is defined over the infinite signature $\mathcal{S} = \{<, =\} \cup \{=_c |\ c \in \mathbb{Z}\} \cup \{\equiv_{a,b} |\ 0 \le a < b\}$. By the definition of $\mathsf{EHomDef}(\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B}))$ we have to compute from a finite subsignature $\sigma \subseteq \mathcal{S}$ a $\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B})$-sentence φ_σ that defines the existence of a homomorphism to $\mathcal{Z}$ when interpreted over a σ-structure $\mathcal{A}$. Hence, let us fix a finite subsignature $\sigma \subseteq \mathcal{S}$. We can assume that $\sigma = \{<, =\} \cup \{=_c |\ c \in C\} \cup \{\equiv_{a,b} |\ b \in D, 0 \le a < b\}$ for finite non-empty sets $C \subseteq \mathbb{Z}$ and $D \subseteq \mathbb{N} \setminus \{0, 1\}$. Define $m = \min(C)$ and $M = \max(C)$. W.l.o.g. we can assume that $m \le 0$ and $M \ge 0$. Let $\mathcal{A} = (A, I)$ be a countable σ-structure. In order to not confuse the relation $I(=)$ with the identity relation on A, we write in the following $E_=(x, y)$ for the atomic formula expressing that (x, y) belongs to the relation $I(=)$. Similarly, we write $E_c(x)$ for the atomic formula expressing that $x \in I(=_c)$. Instead of $\equiv_{a,b}(x)$ we write $x \equiv a \bmod b$.

Define $x \le y \Leftrightarrow (x < y \vee E_=(x, y) \vee E_=(y, x))$ and the MSO-formula

$$\varphi_{\text{bounded}}(x) = \exists y\, \exists z \big(\bigvee_{c \in C} E_c(y) \wedge \bigvee_{c \in C} E_c(z) \wedge \mathsf{reach}_{\le}(y, x) \wedge \mathsf{reach}_{\le}(x, z) \big).$$

Let $B = \{a \in A \mid \mathcal{A} \models \varphi_{\text{bounded}}(a)\}$. We call the induced substructure $\mathcal{B} := \mathcal{A}{\restriction_B}$ the "bounded" part of $\mathcal{A}$. Every homomorphism from $\mathcal{B}$ to $\mathcal{Z}$ has to map B to the interval $[m, M]$. Thus, a homomorphism $h : \mathcal{B} \to \mathcal{Z}$ can be identified with a partition of B into $M - m + 1$ sets $B_m, \ldots, B_M$, where $B_i = \{a \in B \mid h(a) = i\}$. It follows that:

Lemma 20. *There is an* MSO*-sentence* φ_B *such that for every* $\mathcal{S}$*-structure* $\mathcal{A}$ *with bounded part* $\mathcal{B}$*, we have* $\mathcal{B} \preceq \mathcal{Z}$ *if and only if* $\mathcal{A} \models \varphi_B$*.*

Similar to B we define three other parts of a σ-structure by the WMSO-formulas

$$\begin{aligned}
\varphi_{\text{greater}}(x) &= \neg\varphi_{\text{bounded}}(x) \wedge \exists y\, \big(\varphi_{\text{bounded}}(y) \wedge \mathsf{reach}_{\le}(y, x)\big), \\
\varphi_{\text{smaller}}(x) &= \neg\varphi_{\text{bounded}}(x) \wedge \exists y\, \big(\varphi_{\text{bounded}}(y) \wedge \mathsf{reach}_{\le}(x, y)\big), \\
\varphi_{\text{rest}}(x) &= \neg(\varphi_{\text{bounded}}(x) \vee \varphi_{\text{greater}}(x) \vee \varphi_{\text{smaller}}(x)).
\end{aligned}$$

Moreover, let $G = \{a \in A \mid \mathcal{A} \models \varphi_{\text{greater}}(a)\}$, $S = \{a \in A \mid \mathcal{A} \models \varphi_{\text{smaller}}(a)\}$, and $R = \{a \in A \mid \mathcal{A} \models \varphi_{\text{rest}}(a)\}$. Let $\mathcal{N} = \mathcal{Z}{\restriction_{\mathbb{N}}}$ and $\overline{\mathcal{N}} = \mathcal{Z}{\restriction_{\mathbb{Z}\setminus\mathbb{N}}}$. Then we have:

Lemma 21. $\mathcal{A} \preceq \mathcal{Z}$ *iff* $\big(\mathcal{B} \preceq \mathcal{Z}, \mathcal{A}{\restriction_{G \cup S \cup R}} \preceq \mathcal{Z}, \mathcal{A}{\restriction_G} \preceq \mathcal{N},$ *and* $\mathcal{A}{\restriction_S} \preceq \overline{\mathcal{N}}\big)$.

We need some conventions on modulo constraints. A sequence $(a_1, b_1), \ldots, (a_k, b_k)$ with $0 \le a_i < b_i \in D$ for $1 \le i \le k$ is *contradictory*, if there is no number $n \in \mathbb{N}$ such that $n \equiv a_i \bmod b_i$ for all $1 \le i \le k$. In the following let CS_k denote the set of contradictory sequences of length k. It is straightforward to show that every contradictory sequence contains a contradictory subsequence of length at most $\ell := \max\{2, |D|\}$.

Recall that $\sim$ is the smallest equivalence relation containing $I(=)$ and that $\sim$ is defined by the WMSO-formula $\tilde{\varphi}(x, y)$. We call a σ-structure $\mathcal{A} = (A, I)$ *modulo*

contradicting if there is a $\sim$-class $[c]$, elements $c_1, c_2, \ldots, c_k \in [c]$, and a contradictory sequence $(a_1, b_1), \ldots, (a_k, b_k)$ such that $c_i \in I(\equiv_{a_i,b_i})$ for all $1 \leq i \leq k$.

The following WMSO-formula φ_{modcon} expresses that a σ-structure is modulo contradicting, where we write $s_a(j)$ (resp. $s_b(j)$) for the first (resp. second) entry of the j-th element of the sequence $s \in \mathsf{CS}_k$:

$$\varphi_{\text{modcon}} = \bigvee_{2 \leq k \leq \ell} \bigvee_{s \in \mathsf{CS}_k} \exists x_1 \cdots \exists x_k \bigwedge_{i,j \leq k} \tilde{\varphi}(x_i, x_j) \wedge \bigwedge_{j \leq k} x_j \equiv s_a(j) \bmod s_b(j)$$

Lemma 22. *Let $\sigma' = \sigma \setminus \{=_c \mid c \in \mathbb{Z}\}$. Let $\mathcal{A} = (A, I)$ be a σ'-structure.*

- *$\mathcal{A} \preceq \mathcal{Z}$ iff $\mathcal{A}$ is not modulo contradicting and $(A, I(<), I(=)) \preceq (\mathbb{Z}, <, =)$.*
- *$\mathcal{A} \preceq \mathcal{N}$ iff $\mathcal{A}$ is not modulo contradicting and $(A, I(<), I(=)) \preceq (\mathbb{N}, <, =)$.*

Proof of Prop. 13. Let $\mathcal{A} = (A, I)$ be a σ-structure. We defined a partition of A into $B, G, S,$ and R. Since membership in each of these sets is $(\mathsf{WMSO}{+}\mathsf{B})$-definable, we can relativize any $(\mathsf{WMSO}{+}\mathsf{B})$-formula to any of these sets. For instance, we write φ^G for the relativization of φ to the substructure induced by G. Let φ_B be the MSO-formula from Lemma 20, and for $C \in \{\mathbb{Z}, \mathbb{N}, \mathbb{Z} \setminus \mathbb{N}\}$ let φ_C be a formula that expresses $\mathcal{A} \preceq (C, <, =)$, see Prop. 19. Then $\mathcal{A} \models (\varphi_B \wedge \varphi_{\mathbb{Z}}^{G \cup S \cup R} \wedge \varphi_{\mathbb{N}}^{G} \wedge \varphi_{\mathbb{Z} \setminus \mathbb{N}}^{S} \wedge \neg \varphi_{\text{modcon}})$ iff $\mathcal{A} \preceq \mathcal{Z}$ due to Lemmas 21 and 22. □

7 Extensions, Applications, Open Problems

A simple adaptation of our proof for $\mathcal{Z}$ shows that $\mathcal{Q} = (\mathbb{Q}, <, =, (=_q)_{q \in \mathbb{Q}})$ has the property $\mathsf{EHomDef}(\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B}))$ as well: $\mathcal{A} = (A, I) \preceq \mathcal{Q}$ iff (i) $(A, E_<)$ is acyclic, where $E_<$ is defined as in (2), (ii) there does not exist $(a, b) \in E_<^+$ (the transitive closure of $E_<$) with $a \in I(=_p)$, $b \in I(=_q)$ and $q \leq p$, and (iii) there do not exist $a \sim b$ with $a \in I(=_p)$, $b \in I(=_q)$, and $q \neq p$.

Let us finally state a simple preservation theorem for $\mathcal{A}$-satisfiability for $\mathsf{CTL}^*(\mathcal{S})$. Assume that $\mathcal{A}$ and $\mathcal{B}$ are structures over countable signatures $\mathcal{S}_\mathcal{A}$ and $\mathcal{S}_\mathcal{B}$, respectively, and let B be the domain of $\mathcal{B}$. We say that $\mathcal{A}$ is *existentially interpretable* in $\mathcal{B}$ if there exist $n \geq 1$ and quantifier-free first-order formulas $\varphi(y_1, \ldots, y_l, x_1, \ldots, x_n)$ and

$$\varphi_r(z_1, \ldots, z_{l_r}, x_{1,1}, \ldots, x_{1,n}, \ldots, x_{\mathsf{ar}(r),1}, \ldots, x_{\mathsf{ar}(r),n}) \text{ for } r \in \mathcal{S}_\mathcal{A}$$

over the signature $\mathcal{S}_\mathcal{B}$, where the mapping $r \mapsto \varphi_r$ has to be computable, such that $\mathcal{A}$ is isomorphic to the structure $(\{\bar{b} \in B^n \mid \exists \bar{c} \in B^l \colon \mathcal{B} \models \varphi(\bar{c}, \bar{b})\}, I)$ with

$$I(r) = \{(\bar{b}_1, \ldots, \bar{b}_{\mathsf{ar}(r)}) \in B^{\mathsf{ar}(r)n} \mid \exists \bar{c} \in B^{l_r} \colon \mathcal{B} \models \varphi_r(\bar{c}, \bar{b}_1, \ldots, \bar{b}_{\mathsf{ar}(r)})\} \text{ for } r \in \mathcal{S}_\mathcal{A}.$$

Proposition 23. *If $\mathsf{SATCTL}^*(\mathcal{B})$ is decidable and $\mathcal{A}$ is existentially interpretable in $\mathcal{B}$, then $\mathsf{SATCTL}^*(\mathcal{A})$ is decidable too.*

Examples of structures $\mathcal{A}$ that are existentially interpretable in $(\mathbb{Z}, <, =)$, and hence have a decidable $\mathsf{SATCTL}^*(\mathcal{A})$-problem are (i) $(\mathbb{Z}^n, <_{\text{lex}}, =)$ (for $n \geq 1$), where $<_{\text{lex}}$

denotes the strict lexicographic order on n-tuples of integers, and (ii) the structure $\text{Allen}_{\mathbb{Z}}$, which consists of all $\mathbb{Z}$-intervals together with Allen's relations b (before), a (after), m (meets), *mi* (met-by), o (overlaps), *oi* (overlapped by), d (during), *di* (contains), s (starts), *si* (started by), f (ends), *fi* (ended by). In artificial intelligence, Allen's relations are a popular tool for representing temporal knowledge.

It remains open to determine the complexity of CTL^*-satisfiability with constraints over $\mathcal{Z}$, see the last paragraph in the introduction. Clearly, this problem is 2EXPTIME-hard due to the known lower bound for CTL^*-satisfiability. To get an upper complexity bound, one should investigate the complexity of the emptiness problem for puzzles from [1] (see Lemma 5). An interesting structure for which the decidability status for satisfiability of CTL^* with constraints is open, is $(\{0,1\}^*, \leq_p, \not\leq_p)$, where $\leq_p$ is the prefix order on words, and $\not\leq_p$ is its complement. It is not clear, whether this structure has the property $\mathsf{EHomDef}(\mathsf{Bool}(\mathsf{MSO}, \mathsf{WMSO}{+}\mathsf{B}))$.

Acknowledgments. We are grateful to Szymon Toruńczyk for fruitful discussions.

References

1. M. Bojańczyk and S. Toruńczyk. Weak MSO+U over infinite trees. In *Proc. STACS 2012*, vol. 14 of *LIPIcs*, 648–660. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2012.
2. M. Bojańczyk and S. Toruńczyk. Weak MSO+U over infinite trees (long version). available at `http://www.mimuw.edu.pl/~bojan/papers/wmsou-trees.pdf`
3. L. Bozzelli and R. Gascon. Branching-time temporal logic extended with qualitative Presburger constraints. In *Proc. LPAR 2006*, LNCS 4246, 197–211. Springer, 2006.
4. C. Carapelle, A. Kartzow and M. Lohrey. Satisfiability of CTL* with constraints Technical report, arXiv.org, 2013. `http://arxiv.org/abs/1306.0814`.
5. K.Čerāns. Deciding properties of integral relational automata. In *Proc. ICALP 1994*, LNCS 820, 820:35–46. Springer, 1994.
6. T. Colcombet and C. Löding. Regular cost functions over finite trees. In *Proc. LICS 2010*, 70–79. IEEE Computer Society, 2010.
7. B. Courcelle. The monadic second-order logic of graphs V: On closing the gap between definability and recognizability. *Theor. Comput. Sci.*, 80(2):153–202, 1991.
8. S. Demri and D. D'Souza. An automata-theoretic approach to constraint LTL. *Inf. Comput.*, 205(3):380–415, 2007.
9. S. Demri and R. Gascon. Verification of qualitative $\mathbb{Z}$ constraints. *Theor. Comput. Sci.*, 409(1):24–40, 2008.
10. R. Gascon. An automata-based approach for CTL* with constraints. *Electr. Notes Theor. Comput. Sci.*, 239:193–211, 2009.
11. C. Lutz. Description logics with concrete domains-a survey. In *Advances in Modal Logic 4*, pages 265–296. King's College Publications, 2003.
12. C. Lutz. Combining interval-based temporal reasoning with general TBoxes. *Artificial Intelligence*, 152(2):235 – 274, 2004.
13. C. Lutz. NEXPTIME-complete description logics with concrete domains. *ACM Trans. Comput. Log.*, 5(4):669–705, 2004.
14. C. Lutz and M. Milicic. A tableau algorithm for description logics with concrete domains and general TBoxes. *J. Autom. Reasoning*, 38(1-3):227–259, 2007.