Constraint-Based Abstraction of a Model Checker for Infinite State Systems

Banda, Gourinath; Gallagher, John Patrick

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy
If you believe that this document breaches copyright please contact rucforsk@ruc.dk providing details, and we will remove access to the work immediately and investigate your claim.
Constraint-Based Abstraction of a Model Checker for Infinite State Systems

Gourinath Banda\textsuperscript{1} and John P. Gallagher\textsuperscript{1,2}

\textsuperscript{1} Roskilde University, Denmark
\textsuperscript{2} IMDEA Software, Madrid
Email: \{gnbanda, jpg\}@ruc.dk

Abstract. Abstract interpretation-based model checking provides an approach to verifying properties of infinite-state systems. In practice, most previous work on abstract model checking is either restricted to verifying universal properties, or develops special techniques for temporal logics such as modal transition systems or other dual transition systems. By contrast we apply completely standard techniques for constructing abstract interpretations to the abstraction of a CTL semantic function, without restricting the kind of properties that can be verified. Furthermore we show that this leads directly to implementation of abstract model checking algorithms for abstract domains based on constraints, making use of an SMT solver.

1 Introduction

Model Checking is a widely used technique for verifying properties of reactive systems expressed as formulas in a temporal logic, but is restricted to finite-state systems. Abstraction is an effective technique for handling infinite state spaces, where a finite or infinite number of original states are collectively represented with a single abstract state.

The theory of abstract interpretation formalises abstraction techniques. In abstract interpretation-based analyses, an abstract domain is first constructed and then a Galois connection between the original (or concrete) domain and the abstract domain is defined. Computations over the concrete domain are abstractly interpreted over the abstract domain. Due to the Galois connection, the result from abstract computation will always be an over approximation of the actual result, had the actual analysis been possible.

The present work is part of an attempt to develop a uniform CLP-based formal modelling and verification framework for verifying infinite state reactive systems. The modelling part of this framework was covered in [2] where it is shown (i) how to model linear hybrid automata (LHA) specifications in CLP; (ii) how standard program transformation tools of CLP can be applied to extract the underlying state transition system semantics; and (iii) how to abstract the (infinite) reachable state space with the domain of linear constraints. In this extended abstract, we show the verification part of the framework in which abstract interpretation and model checking are integrated.

\textsuperscript{*} Work partly supported by the Danish Natural Science Research Council project SAFT: Static Analysis Using Finite Tree Automata.
The contributions of this work are threefold. First, we present the definition of a CTL-semantics function in a suitable form, using only monotonic functions and fixpoint operators. Second, we apply the standard abstract interpretation framework to get a precise abstraction of the CTL-semantics function. We do not construct an abstract transition system, which turns out to be an unnecessary restriction. Finally, we show how a constraint-based abstraction can be directly implemented from the abstract semantic function and show how satisfiability modulo theories (SMT)-technology can be exploited to improve the performance of our abstract model checker.

The structure of this paper is as follows. Section 2 introduces the syntax and semantics of CTL, and outlines the theory of abstract interpretation. Section 3 describes abstract model checking, that is, abstract interpretation of the CTL semantics function. Section 4 shows how the framework can be applied in practice to an abstraction based on linear constraints. Section 6 gives some experimental results, and we conclude in Section 8.

2 CTL, Model Checking and Abstract Interpretation

CTL is a formal language used to specify temporal properties. A well formed formula in CTL is constructed from one or more atomic propositions and eight CTL-operators. The syntax of CTL is given below.

**Definition 1 (CTL Syntax).** The set of CTL formulas $\phi$ in negation normal form is inductively defined by the following grammar:

$$
\phi ::= \text{true} \mid p \mid \neg p \mid \phi_1 \land \phi_2 \mid \phi_1 \lor \phi_2 \mid \phi_1 \rightarrow \phi_2 \mid \phi_1 \leftrightarrow \phi_2 \mid \text{AX}\phi \mid \text{EX}\phi \mid \text{AF}\phi \\
\mid \text{EF}\phi \mid \text{AG}\phi \mid \text{EG}\phi \mid \text{AU}\phi_1,\phi_2 \mid \text{AR}\phi_1,\phi_2 \mid \text{ER}\phi_1,\phi_2
$$

where $p$ ranges over a set of atomic formulas $P$.

A CTL-formula is in negation normal form (NNF) if and only if the negations are placed in front of the atomic propositions. Any formula not in NNF can be transformed into NNF by moving negations inwards using equivalence-preserving transformations. The unary operators $\text{AF, AG, AX}$ and the binary operators $\text{AU, AR}$ are called universal CTL operators; while the unary $\text{EF, EG, EX}$ and binary $\text{ER, EU}$ operators are called existential CTL operators.

2.1 CTL Semantics

The semantics of CTL formulas is defined with respect to a Kripke structure, which is a state transition system whose states are labelled with atomic propositions that are true in that state.

**Definition 2 (Kripke structure).** A Kripke structure is a tuple $\langle S, \Delta, I, L, P \rangle$ where $S$ is the set of states, $\Delta \subseteq S \times S$ is the transition relation, $I \subseteq S$ is the set of initial states, $P$ is the set of propositions and $L : S \rightarrow 2^P$ is the labelling function which returns the set of propositions that are true in each state. The set of atomic propositions is closed under negation.
Given the Kripke structure $\langle S, \Delta, I, L, \mathcal{P} \rangle$, the meaning of a formula is the set of states in $S$ where the formula holds; this is itself an abstraction of a more detailed trace-based semantics [9]. We define a function $\llbracket \cdot \rrbracket : CTL \rightarrow 2^S$ that returns the set of states where the formula holds. This function is called the CTL-semantics function.

**Definition 3 (CTL-semantics function).** Given a Kripke structure $K = \langle S, \Delta, I, L, \mathcal{P} \rangle$, the semantic function $\llbracket \cdot \rrbracket : CTL \rightarrow 2^S$ is defined as follows.

$$\llbracket true \rrbracket = S \quad \llbracket false \rrbracket = \emptyset \quad \llbracket \neg p \rrbracket = \text{states}(\neg p)$$
$$\llbracket EX \phi \rrbracket = \text{pred}_3(\llbracket \phi \rrbracket) \quad \llbracket AX \phi \rrbracket = \{ \phi_1 \lor \phi_2 \}$$
$$\llbracket EF \phi \rrbracket = \mu Z.(\llbracket \phi \rrbracket \cup \text{pred}_3(Z)) \quad \llbracket ER \phi_1, \phi_2 \rrbracket = \nu Z.(\llbracket \phi_1 \rrbracket \cap (\llbracket \phi_2 \rrbracket \cup \text{pred}_3(Z)))$$
$$\llbracket AF \phi \rrbracket = \mu Z.(\llbracket \phi \rrbracket \cup \text{pred}_3(Z)) \quad \llbracket AU \phi_1, \phi_2 \rrbracket = \mu Z.(\llbracket \phi_1 \rrbracket \cup (\llbracket \phi_2 \rrbracket \cap \text{pred}_3(Z)))$$
$$\llbracket AG \phi \rrbracket = \nu Z.(\llbracket \phi \rrbracket \cap \text{pred}_3(Z)) \quad \llbracket AU \phi_1, \phi_2 \rrbracket = \nu Z.(\llbracket \phi_2 \rrbracket \cap (\llbracket \phi_1 \rrbracket \cup \text{pred}_3(Z)))$$
$$\llbracket EG \phi \rrbracket = \nu Z.(\llbracket \phi \rrbracket \cap \text{pred}_3(Z)) \quad \llbracket AR \phi_1, \phi_2 \rrbracket = \nu Z.(\llbracket \phi_2 \rrbracket \cap (\llbracket \phi_1 \rrbracket \cup \text{pred}_3(Z)))$$

This semantics function makes use of three subsidiary functions $\text{pred}_3 : 2^S \rightarrow 2^S$, $\text{pred}_4 : 2^S \rightarrow 2^S$ and $\text{states} : \mathcal{P} \rightarrow 2^S$ called the existential predecessor function, the universal predecessor function and allocating function respectively. These three functions are specific for a given Kripke structure $K$ and are monotonic.

**Definition 4.** Given a Kripke structure $K = \langle S, \Delta, I, L, \mathcal{P} \rangle$ we define functions $\text{pred}_3 : 2^S \rightarrow 2^S$, $\text{pred}_4 : 2^S \rightarrow 2^S$ and states : $\mathcal{P} \rightarrow 2^S$ as follows.

- $\text{pred}_3(S') = \{ s | \exists s' \in S' : (s, s') \in \Delta \}$ returns the set of states having at least one of their successors in the set $S' \subseteq S$;
- $\text{pred}_4(S') = \text{pred}_3(S') \setminus \text{pred}_3(\text{compl}(S'))$ returns the set of states all of whose successors are in the set $S' \subseteq S$; the function $\text{compl}(X) = S \setminus X$.
- $\text{states}(p) = \{ s \in S | p \in L(s) \}$ returns the set of states where $p \in \mathcal{P}$ holds.

In the CTL semantic definition, $\mu Z.(F(Z))$ (resp. $\nu Z.(F(Z))$) stands for the least fixed point (resp. greatest fixed point) of the function $\lambda Z.F(Z)$. The Knaster-Tarski fixed point theorem [28] guarantees the existence of least and greatest fixed points for a monotonic function on a complete lattice. All the expressions $F(Z)$ occurring in $\mu Z.(F(Z))$ and $\nu Z.(F(Z))$ in the CTL semantic functions are functions $2^S \rightarrow 2^S$ on the complete lattice $\langle 2^S, \cup, \cap, \subseteq, S, \emptyset \rangle$. They are constructed with monotonic operators ($\cup, \cap$) and monotonic functions ($\text{pred}_3, \text{pred}_4$). Thus the CTL semantics function is well-defined.

### 2.2 Model Checking

Model checking is based on checking that the Kripke structure $K = \langle S, \Delta, I, L, \mathcal{P} \rangle$ possesses a property $\phi$, written $K \models \phi$. This is defined to be true iff $I \subseteq \llbracket \phi \rrbracket$, or equivalently, that $I \cap [\neg \phi] = \emptyset$. (Note that $\neg \phi$ should be converted to negation normal form.) Thus model-checking requires implementing the CTL-semantics function which in essence is a fixed point computation. When the state-space $S$ is finite the greatest and least fixed
point expressions can be evaluated as the limits of Kleene sequences. But when \( S \) is infinite, the fixed point computations might not terminate and hence the model checking of infinite state systems becomes undecidable. In this case we try to approximate \([1]\) using the theory of abstract interpretation.

### 2.3 Abstract Interpretation

In abstract interpretation we replace the “concrete” semantic function by an abstract semantic function, developed systematically from the concrete semantics with respect to a Galois connection. We present the formal framework briefly.

**Definition 5 (Galois Connection).** \( \langle L, \sqsubseteq_L \rangle \overset{\gamma}{\longrightarrow} \langle M, \sqsubseteq_M \rangle \) is a Galois Connection between the lattices \( \langle L, \sqsubseteq_L \rangle \) and \( \langle M, \sqsubseteq_M \rangle \) if and only if \( \alpha: L \rightarrow M \) and \( \gamma: M \rightarrow L \) are monotonic and for all \( l \in L, m \in M \), \( \alpha(l) \sqsubseteq_M m \leftrightarrow l \sqsubseteq_L \gamma(m) \).

In abstract interpretation, \( \langle L, \sqsubseteq_L \rangle \) and \( \langle M, \sqsubseteq_M \rangle \) are the concrete and abstract semantic domains respectively. Given a Galois connection \( \langle L, \sqsubseteq_L \rangle \overset{\gamma}{\longrightarrow} \langle M, \sqsubseteq_M \rangle \) and a monotonic concrete semantics function \( f: L \rightarrow L \), then we define an abstract semantic function \( f^{\sharp}: M \rightarrow M \) such that for all \( m \in M \), \( (\alpha \circ f \circ \gamma)(m) \sqsubseteq_M f^{\sharp}(m) \). Furthermore it can be shown that \( \text{lfp}(f^{\sharp}) \sqsubseteq_L \gamma(\text{lfp}(f^{\sharp})) \) and that \( \text{gfp}(f^{\sharp}) \sqsubseteq_L \gamma(\text{gfp}(f^{\sharp})) \).

Thus the abstract function \( f^{\sharp} \) can be used to compute over-approximations of \( f \), which can be interpreted using the \( \gamma \) function. The case where the abstract semantic function is defined as \( f^{\sharp} = (\alpha \circ f \circ \gamma) \) gives the most precise approximation.

If \( M \) is a finite-height lattice then the non-terminating fixed point computations of \( \text{lfp}(f) \) and \( \text{gfp}(f) \) over \( L \) are approximated with a terminating fixed point computation over the finite lattice \( M \).

We next apply this general framework to abstraction of the CTL semantic function, and illustrate with a specific abstraction in Section 4.

### 3 Abstract Interpretation of the CTL-Semantic function

In this section we consider abstractions based on Galois connections of the form \( \langle 2^S, \subseteq \rangle \overset{\gamma}{\longrightarrow} \langle 2^A, \subseteq \rangle \), where the abstract domain \( 2^A \) consists of sets of abstract states. In fact the abstract domain could be any lattice but for the purposes of this paper we consider such state-based abstractions, which will be further discussed in Section 4.

**Definition 6.** Let \( \text{pred}_\exists: 2^S \rightarrow 2^S \), \( \text{pred}_\forall: 2^S \rightarrow 2^S \), and \( \text{states}: \mathcal{P} \rightarrow 2^S \) be the functions defined in Definition 4 and used in the CTL semantic function. Given a Galois connection \( \langle 2^S, \subseteq \rangle \overset{\gamma}{\longrightarrow} \langle 2^A, \subseteq \rangle \), we define \( \text{apred}_\exists: 2^A \rightarrow 2^A \), \( \text{apred}_\forall: 2^A \rightarrow 2^A \) and \( \text{astates}: \mathcal{P} \rightarrow 2^A \) as

\[
\text{apred}_\exists = \alpha \circ \text{pred}_\exists \circ \gamma \quad \text{apred}_\forall = \alpha \circ \text{pred}_\forall \circ \gamma \quad \text{astates} = \alpha \circ \text{states}
\]

It follows directly from the properties of Galois connections that for all \( S' \subseteq S \), \( \alpha(\text{pred}_\exists(S')) \subseteq \text{apred}_\exists(\alpha(S')) \) and \( \alpha(\text{pred}_\forall(S')) \subseteq \text{apred}_\forall(\alpha(S')) \).
Definition 7 (Abstract CTL semantics function). Given a Galois connection \( (2^S, \subseteq) \xrightarrow{\gamma} (2^A, \subseteq) \), the abstract CTL semantic function \( \llbracket \cdot \rrbracket^a : CTL \rightarrow 2^A \) is defined as follows.

\[
\begin{align*}
\llbracket \text{true} \rrbracket^a &= A \\
\llbracket p \rrbracket^a &= \text{astates}(p) \\
\llbracket EX \phi \rrbracket^a &= \text{apred}_3(\llbracket \phi \rrbracket^a) \\
\llbracket AX \phi \rrbracket^a &= \text{apred}_4(\llbracket \phi \rrbracket^a) \\
\llbracket EF \phi \rrbracket^a &= \mu Z.(\llbracket \phi \rrbracket^a \cup \text{apred}_3(Z)) \\
\llbracket AF \phi \rrbracket^a &= \mu Z.(\llbracket \phi \rrbracket^a \cup \text{apred}_4(Z)) \\
\llbracket AG \phi \rrbracket^a &= \nu Z.(\llbracket \phi \rrbracket^a \cap \text{apred}_3(Z)) \\
\llbracket EG \phi \rrbracket^a &= \nu Z.(\llbracket \phi \rrbracket^a \cap \text{apred}_4(Z)) \\
\llbracket \text{false} \rrbracket^a &= \emptyset \\
\llbracket \neg p \rrbracket^a &= \text{astates}(\neg p) \\
\llbracket \phi_1 \lor \phi_2 \rrbracket^a &= \llbracket \phi_1 \rrbracket^a \cup \llbracket \phi_2 \rrbracket^a \\
\llbracket \phi_1 \land \phi_2 \rrbracket^a &= \llbracket \phi_1 \rrbracket^a \cap \llbracket \phi_2 \rrbracket^a \\
\llbracket \forall \phi \rrbracket^a &= \nu Z.(\llbracket \phi \rrbracket^a \cap \text{apred}_3(Z)) \\
\llbracket \exists \phi \rrbracket^a &= \mu Z.(\llbracket \phi \rrbracket^a \cup \text{apred}_4(Z)) \\
\llbracket \exists \forall \phi \rrbracket^a &= \nu Z.(\llbracket \phi \rrbracket^a \cap \text{apred}_3(Z)) \\
\llbracket \forall \exists \phi \rrbracket^a &= \mu Z.(\llbracket \phi \rrbracket^a \cup \text{apred}_4(Z)) \\
\llbracket \Delta \phi \rrbracket^a &= \llbracket \phi \rrbracket^a \setminus \llbracket \phi \rrbracket^a \\
\llbracket \nu \phi \rrbracket^a &= \{a \in A | \phi(a) \} \\
\llbracket \mu \phi \rrbracket^a &= \{a \in A | \phi(a) \wedge \forall a \in A. \neg \phi(a) \}
\end{align*}
\]

Since all the operators appearing in the abstract CTL-semantic are monotonic, the fixpoint expressions and hence the abstract semantic function is well defined. The following soundness theorem is the basis of our abstract model checking approach.

Theorem 1 (Safety of Abstract CTL Semantics). Let \( K = (S, \Delta, I, L, \mathcal{P}) \) be a Kripke structure, \( (2^S, \subseteq) \xrightarrow{\gamma} (2^A, \subseteq) \) be a Galois connection and \( \phi \) any CTL-formula in negation normal form. Then \( \alpha(\llbracket \phi \rrbracket) \subseteq \llbracket \phi \rrbracket^a \) and \( \gamma(\llbracket \phi \rrbracket^a) \supseteq \llbracket \phi \rrbracket \).

The proof follows from the fact that \( \alpha \) is a join-morphism: that is, that \( \alpha(\llbracket S_1 \cup S_2 \rrbracket) = \alpha(\llbracket S_1 \rrbracket) \cup \alpha(\llbracket S_2 \rrbracket) \) and the fact that \( \alpha(\llbracket S_1 \cap S_2 \rrbracket) \subseteq \alpha(\llbracket S_1 \rrbracket) \cap \alpha(\llbracket S_2 \rrbracket) \).

This theorem provides us with a sound abstract model checking procedure for any CTL formula \( \phi \). As noted previously, \( K \models \phi \) iff \( \gamma(\llbracket \neg \phi \rrbracket^a) \cap I = \emptyset \) (where \( \neg \phi \) is converted to negation normal form). It follows from Theorem 1 that this follows if \( \gamma(\llbracket \neg \phi \rrbracket^a) \cap I = \emptyset \). Of course, if \( \gamma(\llbracket \neg \phi \rrbracket^a) \cap I = \emptyset \) nothing can be concluded.

4 Abstract Model Checking in Constraint-based Domains

The abstract semantics given in Section 3 is not always implementable in practice for a given Galois connection \( (2^S, \subseteq) \xrightarrow{\gamma} (2^A, \subseteq) \). In particular, the function \( \gamma \) yields a value in the concrete domain, which is typically an infinite object such as an infinite set. Thus evaluating the functions \( (\alpha \circ \text{pred}_3 \circ \gamma) \) and \( (\alpha \circ \text{pred}_4 \circ \gamma) \) might not be feasible.

In this section we show that the construction is implementable for transition systems and abstract domains expressed using linear constraints.

4.1 Constraint Representation of Transition Systems

We consider the set of linear arithmetic constraints (hereafter simply called constraints) over the real numbers.

\[
c ::= t_1 \leq t_2 \mid t_1 < t_2 \mid c_1 \wedge c_2 \mid c_1 \lor c_2 \mid \neg c
\]

where \( t_1, t_2 \) are linear arithmetic terms built from real constants, variables and the operators +, \( \times \) and \( - \). The constraint \( t_1 = t_2 \) is an abbreviation for \( t_1 \leq t_2 \wedge t_2 \leq t_1 \). Note
that \(-(t_1 \leq t_2) \equiv t_2 < t_2\) and \(-(t_1 < t_2) \equiv t_2 \leq t_2\), and so the negation symbol \(\neg\) can be eliminated from constraints if desired by moving negations inwards by Boolean transformations and then applying this equivalence.

A constraint is satisfied by an assignment of real numbers to its variables if the constraint evaluates to true under this assignment, and is satisfiable if there exists some assignment that satisfies it. A constraint can be identified with the set of assignments that satisfy it. Thus a constraint over \(n\) real variable represents a set of points in \(R^n\).

A constraint can be projected onto a subset of its variables. Denote by \(\text{proj}_V(c)\) the projection of \(c\) onto the set of variables \(X\).

Let us consider a transition system defined over the state-space \(R^n\). Let \(\vec{x}, \vec{x}_1, \vec{x}_2\) etc. represent \(n\)-tuples of distinct variables, and \(\vec{r}, \vec{r}_1, \vec{r}_2\) etc. represent tuples of real numbers. Let \(\vec{x}/\vec{r}\) represent the assignment of values \(\vec{r}\) to the respective variables \(\vec{x}\). We consider transition systems in which the transitions can be represented as a finite set of transition rules of the form \(\vec{x}_1 \xrightarrow{c(\vec{x}_1, \vec{x}_2)} \vec{x}_2\). This represents the set of all transitions from state \(\vec{r}_1\) to state \(\vec{r}_2\) in which the constraint \(c(\vec{x}_1, \vec{x}_2)\) is satisfied by the assignment \(\vec{x}_1/\vec{r}_1, \vec{x}_2/\vec{r}_2\). Such transition systems can be used to model real-time control systems [18, 2].

4.2 Computation of the CTL semantic function using constraints

A constraint representation of a transition system allows a constraint solver to be used to compute the functions \(\text{pred}_3\), \(\text{pred}_\forall\) and states in the CTL semantics. Let \(T\) be a finite set of transition rules. Let \(c'((\vec{y}))\) be a constraint over variables \(\vec{y}\). It is assumed that the set of propositions in the Kripke structure used in the semantics is the set of linear constraints.

\[
\begin{align*}
\text{pred}_3(c'((\vec{y}))) &= \bigvee \{\text{proj}_V(c'(\vec{y}) \land c(\vec{x}, \vec{y})) | \vec{x} \xrightarrow{c(\vec{x}, \vec{y})} \vec{y} \in T\} \\
\text{pred}_\forall(c'((\vec{y}))) &= \text{pred}_3(c'((\vec{y}))) \land \neg(\text{pred}_3(\neg c'((\vec{y})))) \\
\text{states}(p) &= p
\end{align*}
\]

In the definition of states we use \(p\) both as the proposition (the argument of states) and as a set of points (the result).

4.3 Abstract Domains Based on a Disjoint State-Space Partition

Suppose we have a transition system with \(n\) state variables; we take as the concrete domain the complete lattice \((2^C, \subseteq)\) where \(C \subseteq 2^{R^n}\) is some nonempty, possibly infinite set of \(n\)-tuples including all the reachable states of the system.

We build an abstraction of the state space based on a disjoint partition of \(C\) say \(A = \{d_1, \ldots, d_k\}\) such that \(\bigcup A = C\). Such a partition can itself be constructed by an abstract interpretation of the transition relation [2]. Define a representation function \(\beta : C \rightarrow 2^A\), such that \(\beta(\vec{x}) = \{d \in A | x \in d\}\). We extend the representation function to sets of points, obtaining the abstraction function \(\alpha : 2^C \rightarrow 2^A\) given by \(\alpha(S) = \bigcup\{\beta(x) | x \in S\}\). Define the concretisation function \(\gamma : 2^A \rightarrow 2^C\), as \(\gamma(V) = \{x \in C | \beta(x) \subseteq V\}\). As shown in [24, 7], \((C, \subseteq) \xrightarrow{\gamma} (A, \subseteq)\) is a Galois connection. Because the partition \(A\) is disjoint the value of \(\beta(x)\) is a singleton for all \(x\), and the \(\gamma\) function can be written as \(\gamma(V) = \bigcup\{\gamma(\{d\}) | d \in V\}\).
4.4 Representation of Abstraction Using Constraints

A constraint can be identified with the set of points that satisfies it. Suppose that each element \( d \) of the partition \( A \) is representable as a linear constraint \( c_d \) over the variables \( x_1, \ldots, x_n \). The \( \bar{\beta} \) function can be rewritten as \( \bar{\beta}(x) = \{ d \mid x \text{ satisfies } c_d \} \). Assuming that we apply \( \alpha \) to sets of points represented by a linear constraint over \( x_1, \ldots, x_n \), we can rewrite the \( \alpha \) and \( \gamma \) functions as follows.

\[
\alpha(c) = \{ d \mid \text{SAT}(c_d \land c) \} \quad \gamma(V) = \bigvee \{ c_d \mid d \in V \}
\]

4.5 Computation of \( \alpha \) and \( \gamma \) functions using constraint solvers

The constraint formulations of the \( \alpha \) and \( \gamma \) functions allows them to be effectively computed. The expression \( \text{SAT}(c_d \land c) \) occurring in the \( \alpha \) function means “\( (c_d \land c) \) is satisfiable” and can be checked by an SMT solver. In our experiments we use the SMT solver Yices [13]. The \( \gamma \) function simply collects a disjunction of the constraints associated with the given set of partitions; no solver is required.

4.6 Implementation of constraint-based abstract semantics

Combining the constraint-based evaluation of the functions \( \text{pred}_3 \) and \( \text{pred}_4 \) with the constraint-based evaluation of the \( \alpha \) and \( \gamma \) functions gives us (in principle) a method of computing the abstract semantic counterparts of \( \text{pred}_3 \) and \( \text{pred}_4 \), namely \( \{ \alpha \circ \text{pred}_3 \circ \gamma \} \) and \( \{ \alpha \circ \text{pred}_4 \circ \gamma \} \). This gives us a sound abstract semantics for CTL as discussed previously. The question we now address is the feasibility of this approach. Taken naively, the evaluation of these constraint-based functions (in particular \( \text{pred}_4 \)) does not scale up. We now show how we can transform these definitions to a form which can be computed much more efficiently, with the help of an SMT solver.

Consider the evaluation of \( (\alpha \circ \text{pred}_4 \circ \gamma)(V) \) where \( V \in 2^A \) is a set of disjoint partitions represented by constraints.

\[
(\alpha \circ \text{pred}_4 \circ \gamma)(V) = \alpha(\text{pred}_4((\bigvee \{ c_d \mid d \in V \}) \land \neg(\text{pred}_3(\neg(\bigvee \{ c_d \mid d \in V \}))))
\]

In the last step, we use the equivalence \( \neg(\bigvee \{ c_d \mid d \in V \}) \leftrightarrow \bigvee \{ c_d \mid d \in A \setminus V \} \), which is justified since the abstract domain \( A \) is a disjoint partition of the concrete domain; thus \( A \setminus V \) represents the negation of \( V \) restricted to the state space of the system. The computation of \( \text{pred}_3(\bigvee \{ c_d \mid d \in A \setminus V \}) \) is much easier to compute (with available tools) than \( \text{pred}_3(\neg(\bigvee \{ c_d \mid d \in V \})) \). The latter requires the projection operations \( \text{proj} \) to be applied to complex expressions of the form \( \text{proj}_k(\neg(c_1(y)) \lor \cdots \lor c_k(y)) \land \neg(c(x,y)) \), which involves expanding the expression (to d.n.f. for example); by contrast the former requires evaluation of simpler expressions of the form \( \text{proj}_k(c_d(y)) \land \neg(c(x,y)) \).
4.7 Further Optimisation by Pre-Computing Predecessor Constraints

We now show that we can improve the computation of the abstract function \((\alpha \circ \text{pred}_3 \circ \gamma)\). Let \(\{c_i\}\) be a set of constraints, each of which represents a set of points. It can easily be seen that \(\text{pred}_3(V\{c_i\}) = V\{\text{pred}_3(c_i)\}\). Consider the evaluation of \((\alpha \circ \text{pred}_3 \circ \gamma)(V)\) where \(V \in 2^A\) is a set of disjoint partitions represented by constraints.

\[
(\alpha \circ \text{pred}_3 \circ \gamma)(V) = (\alpha \circ \text{pred}_3)(\bigvee\{c_d \mid d \in V\})
= \alpha(\bigvee\{\text{pred}_3(c_d) \mid d \in V\})
\]

Give a finite partition \(A\), we pre-compute the constraint \(\text{pred}_3(c_d)\) for all \(d \in A\). Let \(\text{Pre}(d)\) be the stored predecessor constraint for partition element \(d\). The results can be stored as a table, and whenever it is required to compute \((\alpha \circ \text{pred}_3 \circ \gamma)(V)\) where \(V \in 2^A\), we simply evaluate \(\alpha(\bigvee\{\text{Pre}(d) \mid d \in V\})\). The abstraction function \(\alpha\) is evaluated efficiently using the SMT solver, as already discussed.

Note that expressions of the form \(\alpha(\text{pred}_3(V\{\cdots\}))\) occur in the transformed expression for \((\alpha \circ \text{pred}_3 \circ \gamma)(V)\) above. The same optimisation can be applied here too. Our experiments show that this usually yields a considerable speedup (2-3 times faster) compared to dynamically computing the \(\text{pred}_3\) function during model checking.

5 Implementation

The abstract CTL semantic function was implemented directly in Prolog without serious attempt at optimisation of the algorithm. The function \(\llbracket \phi \rrbracket^a\) yielding a set \(S\) is represented by the predicate \(\text{absCtl}(\Phi, S)\), where \(\Phi\) is a suitable representation of a CTL formula. Thus for example, the rule for evaluating a formula \(\text{AG} \phi\), namely

\[
\llbracket \text{AG} \phi \rrbracket^a = \nu Z.(\llbracket \phi \rrbracket^a \cap \text{apred}_v(Z))
\]

is rendered in Prolog by the clauses

\[
\text{absCtl}(\text{ag}(F), \text{States}) :-
\text{absCtl}(F, \text{FStates}),
\text{gfp}_\text{ag}(\text{FStates}, \text{States}).
\]

\[
\text{gfp}_\text{ag}(F, S) :-
\text{gfp}('\$VAR\'( 'Z' ), \text{intersect}(F, \text{predForall}('\$VAR\'( 'Z' ))), S).
\]

The predicate \(\text{gfp}(Z, F, S)\) computes the greatest fixed point of the function \(\lambda Z. (F(Z))\), and is implemented naively as shown below by computing successive iterations \(A, F(A), F(F(A)), \ldots\) until \(F^j(A) = F^{j+1}(A)\) for some \(j\). Here, \(A\) is the set of all abstract regions for the system under consideration. There are improved fixpoint algorithms in the literature which could be applied, e.g. [4].
The most relevant aspect of the prototype implementation is the interface to external libraries to perform constraint-solving functions. In implementing the \( \text{pred}_3 \) operation we make use of a Ciao-Prolog interface to the PPL library [1]. In particular, this is used to compute the \( \text{proj} \) function. The \( \alpha \) function is implemented using the SMT solver Yices [13]. We implemented an interface predicate \( \text{yices_sat}(C,Xs) \), where \( C \) is a constraint and \( Xs \) is the set of variables in \( C \). This predicate simply translates \( C \) to the syntax of Yices, and succeeds if and only if Yices finds that the constraint is satisfiable. Using this predicate the definition of \( \alpha \), that is \( \alpha(c) = \{ d \mid \text{SAT}(c_d \land c) \} \) can be implemented directly as defined.

6 Experiments Using an SMT Constraint Solver

Figure 1 shows the transitions of a water-level controller taken from [18]. The transitions are represented as constraint logic program clauses generated automatically from a Linear Hybrid Automaton specification of the controller, as explained in detail in [2]. The state variables in an atomic formula of form \( \text{rState}(X,W,T,L) \) represent the rate of flow \( (X) \), the water-level \( (W) \), the elapsed time \( (T) \) and the location identifier \( (L) \). The meaning of a clause of form

\[
\text{rState}(X,W,T,L) :- \text{c}(X,W,T,L,X1,W1,T1,L1), \text{rState}(X1,W1,T1,L1)
\]

is a transition rule \( (X,W,T,L) \xrightarrow{c(X,W,T,L,X1,W1,T1,L1)} (X1,W1,T1,L1) \). The initial state is given by the clause \( \text{rState}(0,1,\_1) \). Note that there are transitions both from one location to another, and also from a location to itself, since the controller can remain in a location so long as an invariant is satisfied.

Figure 2 shows the result of an analysis of the reachable states of the system, based on computing an approximation of the minimal model of the constraint program in Figure 1. There are 8 regions, which cover the reachable states of the controller starting in the initial state (which is region 1). The term \( v(N, \text{rState}(A,B,C,D),[\_]) \) means
rState(0,1,_,1).
rState(A,B,C,1) :- D < A, B = E + A - D, C = F + A - D, B + 10,
    rState(D,E,F,1).

rState(0,_,B,1) :- C < D, E = F - 2*D - D, G = H + D - C, G + 2, B = G, A = E,
    rState(C,F,H,4).

rState(A,_,C,2) :- D < A, B = E + A - D, C = F + A - D, C + 2,
    rState(D,E,F,2).

rState(0,_,B,2) :- C < D, E = F + D - C, _ = G + D - C, E + 10, B = 0, A = E,
    rState(C,F,G,1).

rState(A,B,_,1) :- C < D, E = F - 2*D - D, C = F + A - D, B + 5,
    rState(D,E,F,3).

rState(0,_,B,3) :- C < D, E = F + D - C, G = H + D - C, G + 2, B = G, A = E,
    rState(C,F,H,2).

rState(A,_,C,4) :- D < A, B = E - 2*A - D, C = F + A - D, C + 2,
    rState(D,E,F,4).

rState(0,_,B,4) :- C < D, E = F - 2*D - D, _ = G + D - C, E + 5, B = 0, A = E,
    rState(C,F,G,3).

**Fig. 1.** The Water-Level Controller

v(1, rState(A,B,C,D), [1*A=0,1*B=1,D=1]).

v(2, rState(A,B,C,D), [1*B> -10,1*B>1,1*A> -1,B= -1,D=1]).

v(3, rState(A,B,C,D), [1*A=10,1*A=0,1*C=0,D=2]).

v(4, rState(A,B,C,D), [1*C> -2,1*C=0,1*A> -1,C=0,B=0,D=1]).

v(5, rState(A,B,C,D), [1*B=12,1*A=0,1*C=2,D=3]).

v(6, rState(A,B,C,D), [-2*C> -11,1*C>2,1*A> -1,C= -2,B=0,C=16,D=3]).

v(7, rState(A,B,C,D), [1*B=5,1*A=0,C=0,D=4]).

v(8, rState(A,B,C,D), [-1*C> -2,1*C=0,1*A=0,B=0,C=5,D=4]).

**Fig. 2.** Disjoint Regions of the Water-Level Controller

that the region labelled \( N \) is defined by the constraint in the third argument, with constraint variables \( A, B, C, D \) corresponding to the given state variables. The 8 regions are disjoint. We use this partition to construct the abstract domain as described in Section 4.3.

Our implementation of the abstract semantics function is in Ciao-Prolog with external interfaces to the Parma Polyhedra Library [1] and the Yices SMT solver [13]. Our prototype implementation of the fixpoint computations is very naive. Nonetheless we successfully checked many CTL formulas including those with CTL operators nested in various ways, which in general is not allowed in either UPPAAL [3] or HYTECH [19].

Table 1 gives the results of abstract model checking two systems, namely, a water level monitor and a task scheduler. Both of these systems are taken from [18]. In the table: (i) the columns System and Formula indicate the system and the formula being checked; (ii) the columns \( A \) and \( \Delta \), respectively, indicate the number of abstract regions and original transitions in a system and (iii) the column time indicates the computation...
time to check a formula on the computer with an Intel XEON CPU running at 2.66GHz and with 4GB RAM.

6.1 Water level controller

For the water level system, which has 4 state variables, no formula that we have tried to evaluate takes longer than 0.2 seconds to check. Since verifying certain properties requires finer abstractions (as discussed shortly in Section 6.3), we consider two variants, a coarse one with eight and a more refined abstraction with twelve abstract regions.

The formula $EF(W = 10)$ means “there exists a path along which eventually the water level ($W$) reaches 10”, while $AG(W = 10 \rightarrow EF(W \neq 10))$ means “on every path it is possible for the water level not to get stuck at 10”. The formula $EF(AG(min \leq W \leq max))$ where $min, max \in R$ states “possibly the water level stabilises and fluctuates between a minimum $min$ and maximum $max$”. Using this formula, we can check whether the system reaches a stable region with the given bounds on the water level.

6.2 Scheduler

For the scheduler system that has 8 state variables, 42 abstract regions and 12 transitions, the checking time increases. Here, formulas can take up to 5-6 seconds to check in our prototype implementation. We proved a number of safety and liveness properties, again successfully checking properties of a form beyond the capability of other model checkers. For example the formula $AG(K^2 > 0 \rightarrow EF(K^2 = 0))$, containing an $EF$ nested within an $AG$, means that the tasks of high priority (whose presence is indicated by a strictly positive value of $K^2$) do not get starved (that is, the value of $K^2$ eventually returns to zero).

6.3 Increasing precision by property-specific refinements

The property $EF(W = 3)$ in the water level controller should hold on the system. But this formula cannot be verified when the state space is abstracted with 8 abstract regions. Because of the coarse abstraction, we cannot distinguish $W = 3$ from $W \neq 3$. The negation of the formula, namely $AG(W > 3 \lor W < 3)$, holds in the abstract initial state since there are infinite paths from region 1 which always stay in regions that are consistent with $W \neq 3$.

One approach to solving such cases is to make a property-specific refinement to the abstraction. Each region is split into three regions by adding $W = 3$, $W > 3$ and $W < 3$ respectively to each region. Consequently, since there are 8 regions in the current abstraction (shown in 2), we get a new abstraction with 24 abstract regions, of which only 12 are satisfiable. Only the satisfiable regions need to be retained, giving a total of 12 regions in this example. With this refined abstraction, the property $EF(W = 3)$ can then successfully be checked.
### Table 1. Experimental Results

<table>
<thead>
<tr>
<th>System</th>
<th>Formula</th>
<th>A</th>
<th>Δ</th>
<th>Time (secs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waterlevel Monitor</td>
<td>$EF(W = 10)$</td>
<td>8</td>
<td>8</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>$AG(W = 10 \rightarrow EF(W \neq 10))$</td>
<td>8</td>
<td>8</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>$EF(AG(1 \leq W \leq 12))$</td>
<td>8</td>
<td>8</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>$AG(W \geq 1))$</td>
<td>8</td>
<td>8</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>$EF(W = 3)$</td>
<td>12</td>
<td>8</td>
<td>0.16</td>
</tr>
<tr>
<td>Task Scheduler</td>
<td>$EF(K2 = 1)$</td>
<td>42</td>
<td>14</td>
<td>5.51</td>
</tr>
<tr>
<td></td>
<td>$AG(K2 &gt; 0 \rightarrow EF(K2 = 0))$</td>
<td>42</td>
<td>14</td>
<td>3.49</td>
</tr>
<tr>
<td></td>
<td>$AG(K2 \leq 1)$</td>
<td>42</td>
<td>14</td>
<td>3.49</td>
</tr>
</tbody>
</table>

#### 6.4 Limitations implied by our modelling technique

We cannot always successfully check formulas of the form $AF \phi$, due to an abstraction introduced into our model of continuous behaviour (rather than the abstraction induced by the Galois connection). The reason for this is that the transitions of the system in general include additional self-transitions that were not intended in the original system. Such transitions with the same location for the successor state as well as predecessor state are those which do not respect the continuity of the physical system. For example, the transition rules of the water level controller allow a transition within location 1 directly from $W = 1$ to $W = 5$ without passing through $W = 3$. When trying to prove a formula of the form $AF \phi$, we need to refute the formula $\neg(AF \phi)$ i.e. $EG \neg \phi$. Because of the extra self transitions, there might exist a path from the initial state on which $\neg \phi$ holds forever. Thus refutation might not be possible. Other modelling techniques are needed to capture continuity.

#### 7 Related Work

The topic of model-checking infinite state systems using some form of abstraction has been already widely studied. Abstract model checking is described by Clarke et al. [6]. In this approach a state-based abstraction is defined where an abstract state is a set of concrete states. A state abstraction together with a concrete transition relation $\Delta$ induces an abstract transition relation $\Delta_{abs}$. Specifically, if $X_1, X_2$ are abstract states, $(X_1, X_2) \in \Delta_{abs}$ iff $\exists x_1 \in X_1, x_2 \in X_2$ such that $(x_1, x_2) \in \Delta$. From this basis an abstract Kripke structure can be built; the initial states of the abstract Kripke structure are the abstract states that contain a concrete initial state, and the property labelling function of the abstract Kripke structure is induced straightforwardly as well. Model checking over the abstract Kripke structure is correct for universal temporal formulas (ACTL), that is, formulas that do not contain operators $EX, EF, EG$ or $EU$. Intuitively, the set of paths in the abstract Kripke structure represents a superset of the paths of the concrete Kripke structure. Hence, any property that holds for all paths of the abstract Kripke structure also holds in the concrete structure. If there is a finite number of abstract states, then the abstract transition relation is also finite and thus a standard (finite-state) model checker can be used to perform model-checking of ACTL properties. However,
if an ACTL property does not hold in the abstract structure, nothing can be concluded
about the concrete structure, and furthermore checking properties containing existential
path quantifiers is not sound in such an approach.

This technique for abstract model checking can be reproduced in our approach,
although we do not explicitly use an abstract Kripke structure. Checking an ACTL for-
mula is done by negating the formula and transforming it to negation normal form,
yielding an existential temporal formula (ECTL formula). Checking such a formula
using our semantic function makes use of the pred function but not the pred function.
It can be shown that the composition \((\alpha \circ \text{pred} \circ \gamma)\) gives the pred function
for the abstract transition relation defined by Clarke et al. Note that whereas abstract
model checking the ACTL formula with an abstract Kripke structure yields an under-
approximation of the set of states where the formula holds, our approach yields the
complement, namely an over-approximation of the set of states where the negation of
the formula holds.

There have been different techniques proposed in order to overcome the restriction
to ACTL formulas. Dams et al. [10] present a framework for constructing abstract inter-
pretations for transition systems. This involves constructing a mixed transition system
containing two kinds of transition relations, the so-called free and constrained transi-
tions. Godefroid et al. [16] proposed the use of modal transition systems [22] which
consist of two components, namely must-transitions and may-transitions. In both [10]
and [16], given a state abstraction together with a concrete transition system, a mixed
transition system, or an (abstract) modal transition system respectively, is automatically
generated. Following this, a modified model-checking algorithm is defined in which
any formula can be checked with respect to the dual transition relations. There are cer-
tainly similarities between these approaches and ours, though more study of the precise
relationship is needed. The may-transitions are captured by the abstract transitions de-
fined by Clarke et al. [6] and hence by our abstract function \((\alpha \circ \text{pred} \circ \gamma)\), as discussed
above. We conjecture that the must-transitions are closely related to the abstract function
\((\alpha \circ \text{pred} \circ \gamma)\). We argue that the construction of abstract transition systems, and
the consequent need to define different transitions preserving universal and existential
properties, is an avoidable complication, and that our approach is conceptually simpler.
Probably the main motivation for the definition of abstract transition systems is to re-use
existing model checkers, as remarked by Cousot and Cousot [9].

The application of the theory of abstract interpretation to temporal logic, includ-
ing abstract model checking, is thoroughly discussed by Cousot and Cousot [8, 9]. Our
abstract semantics is inspired by these works, in that we also proceed by direct abstrac-
tion of a concrete semantic function using a Galois connection, without constructing
any abstract transition relations. The technique of constructing abstract functions based
on the pattern \((\alpha \circ f \circ \gamma)\), while completely standard in abstract interpretation [7], is not
discussed explicitly in the temporal logic context. We focus only on state-based abstrac-
tions (Section 9 of [9]) and we ignore abstraction of traces. Our contribution compared
to these works is to work out the abstract semantics for a specific class of constraint-
based abstractions, and point the way to effective abstract model checking implemen-
tations using SMT solvers. Kelb [21] develops a related abstract model checking algo-
rithm based on abstraction of universal and existential predecessor functions which are essentially the same as our \( \text{pred}_\forall \) and \( \text{pred}_\exists \) functions.

Giacobazzi and Quintarelli [15] discuss abstraction of temporal logic and their refinement, but deal only with checking universal properties.

Our technique for modelling and verifying real time and concurrent systems using constraint logic programs builds on the work of a number of other authors, including Gupta and Pontelli [17], Jaffar et al. [20] and Delzanno and Podelski [11]. However we take a different direction from them in our approach to abstraction and checking of CTL formulas, in that we use abstract CLP program semantics when abstracting the state space (only briefly covered in the present work), but then apply this abstraction in a temporal logic framework, which is the topic of this work. Other authors have encoded both the transition systems and CTL semantics as constraint logic programs [5, 23, 25, 12, 14, 26, 27]. However none of these develops a comprehensive approach to abstract semantics when dealing with infinite-state systems. Perhaps a unified CLP-based approach to abstract CTL semantics could be constructed based on these works.

8 Conclusion

We have demonstrated a practical approach to abstract model checking, by constructing an abstract semantic function for CTL based on a Galois connection. Most previous work on abstract model checking is restricted to verifying universal properties and requires the construction of an abstract transition system. In other approaches in which arbitrary properties can be checked [16, 10], a dual abstract transition system is constructed. Like Cousot and Cousot [9] we do not find it necessary to construct any abstract transition system, but abstract the concrete semantic function systematically. Using abstract domains based on constraints we are able to implement the semantics directly. The use of an SMT solver adds greatly to the effectiveness of the approach.

Acknowledgements. We gratefully acknowledge discussions with Dennis Dams and César Sánchez.

References


