Ensuring Completeness of Symbolic Verification Methods for Infinite-State Systems *

Parosh Aziz Abdulla
Dept. Computer Systems
Uppsala University
P.O.Box 325
751 05 Uppsala, Sweden
parosh@docs.uu.se

Bengt Jonsson
Dept. Computer Systems
Uppsala University
P.O.Box 325,
751 05 Uppsala, Sweden
bengt@docs.uu.se

Abstract

Over the last few years there has been an increasing research effort directed towards the automatic verification of infinite state systems. For different classes of such systems, e.g., hybrid automata, data-independent systems, relational automata, Petri nets, and lossy channel systems, this research has resulted in numerous highly nontrivial algorithms. As the interest in this area increases, it will be important to extract common principles that underly these and related results. In this paper, we will present a general model of infinite-state systems, and describe a standard algorithm for reachability analysis of such systems. Our contribution consists in finding conditions under which the algorithm can be fully automated. We perform backward reachability analysis. Using an iterative procedure, we generate successively larger approximations of the set of all states from which a given final state is reachable. We consider classes of systems where these approximations are *well quasi-ordered*, implying that the iterative procedure

*Supported in part by the Swedish Board for Industrial and Technical Development (NUTEK) and by the Swedish Research Council for Engineering Sciences (TFR).

always terminates. Starting from these general termination conditions, we derive several computations models for which reachability is decidable. Many of these models are extensions of those existing in the literature. Using a well-known reduction from safety properties to reachability properties, we can use our algorithm to decide large classes of safety properties for infinite-state systems. A motivation for our approach is the long-term desire to build general tools for verification of infinite-state systems, which implies that we should employ principles applicable across a rather wide range of such systems.

1 Introduction

In recent years, several approaches to mechanized program verification have been developed. Substantial progress has been made in the development and use of interactive theorem provers, such as PVS [Sha96]. Fully automated techniques have now been developed to the extent that they can routinely handle systems with several millions states (for some applications even several magnitudes more). Partial order techniques [GW93, Pel96, Val93] and Binary Decision Diagrams (BDDs) [BCM^{+}90] have extended the power of these techniques dramatically.

Automated verification of infinite-state systems is becoming practical for significant special cases. Nontrivial verification algorithms have been developed, e.g., for timed automata [ACD90, AH89, Čer92a], hybrid automata [Hen95], data-independent systems [JP93, Wol86], relational automata [BBK77, Čer92b, Čer94], Petri nets [Jan90, JM95], pushdown processes [BS95, Sti96], systems with unbounded communication channels [Fin94, AJ93, AK95], and systems consisting of an unbounded number of identical finite-state processes [GS92, KMM^{+}97].

A large portion of existing work on verification algorithms for infinite-state systems consider a particular model of an infinite-state system, and exploit its particular properties to develop a special-purpose verification algorithm. However, in order to be able to verify a general class of infinite-state systems, it is necessary to find and exploit common principles that can be uniformly applied for the entire class. In this paper, we will present an approach to verification of infinite-state systems which does not start from a particular model of systems. Rather, we will present a general model together with some general well-known methods for verification. Our contribution consists in finding conditions under which these general methods can be fully automated. In this paper, the criterion for "fully automated" will be decidability, i.e., that the method is guaranteed to terminate. A motivation for our approach is the long-term desire to build general tools for verification of infinite-state systems. In order to be able build a verification tool which can successfully handle a reasonably large class of systems, we must employ principles applicable across a rather wide range of such systems.

We will concentrate our treatment on the verification of reachability properties. For a particular system and two sets of states (a set of initial states and a set of final states) of the system, the reachability problem asks whether some final state can be reached in an execution which starts from an initial state. A typical application of the problem is to check that some undesired situation, such as deadlock, cannot occur in an execution of a system. More general classes of safety properties can be reduced to the reachability problem [VW86, GW93]: the property is represented by a test or observer process, and then one checks whether or not the test process can reach a state where violation of the safety property has been observed.

There are several standard methods for verifying reachability. An approach which is suitable for automation is to systematically search for an execution path from some initial state to some final state. Such a search can be conducted, starting from the set of initial states (forward search) or from the final states (backward search). In this paper, we will consider the backward search method.

A standard technique for backward search, is to attempt to generate the set of all states from which a state in F is reachable, using an iterative procedure. For successively larger j, we compute the set of states from which a state in F can be reached by a sequence of transitions of length less than or equal to j. We obtain the $(j+1)$-st approximation from the jth approximation by adding the *pre-image* of the jth approximation, i.e., the set of states from which a state in the jth approximation can be reached by some transition. If this procedure converges (i.e., the jth and $j+1$-st approximations coincide for some j), one checks whether the result intersects the set of initial states of the model. The method just mentioned will become a decision procedure for a specific class of systems, provided that we find a suitable representation of infinite sets of states such that we can

- invent a method for computing pre-images, and
- prove that the iteration always converges.

Note that, of course, the method will not converge for arbitrary classes of infinite-state systems.

A typical approach is to represent infinite sets of states using assertions, which we call *constraints*, in some language. The iterative procedure will then generate a sequence of successively weaker constraints. We will try to enforce convergence, by looking for systems of constraints that disallow infinite sequences of successively weaker constraints.

To present the key ideas of this paper, we have chosen to present a simple example in the next section. In this example, we show how to check a simple reachability property for a simple cache protocol, and at the same time introduce the main concepts of the paper.

Related Work Unifying work for verification of infinite-state systems appears e.g., in the framework of the modal mu-calculus, where Bradfield and Stirling [BS92] have presented general techniques for verifying temporal properties of infinite-state systems. In [Fin90], Finkel introduces a class of infinite-state systems called *well structured systems* and describes algorithms to check, e.g. , coverability and eventuality properties for this class. In our earlier work [AČJYK96], we presented similar results for a larger class of infinite-state systems, and showed how they could be applied to particular classes of infinite-state systems. This paper contains part of our earlier results, but presents a more thorough treatment on the verification of reachability properties. A distinguishing feature of this work is, that rather than simply noting that our framework can be extended to some existing models of infinite-state systems from the literature, we try to generate new classes of infinite-state systems in the most general way. In many cases, we derive models that are generalizations of existing models in the literature.

Outline In the next section we present the main ideas of the paper through a simple example. In Section 3 we introduce the basic notions of *transitions systems* and *well quasi-orderings*. Well quasi-orderedness is a property which we assume to hold of our constraint systems. We use this property both to find finite representations of constraint systems, and to prove termination of our verification algorithms. In Section 4 we describe a standard algorithm for reachability analysis of finite-state systems. The basic concepts of this algorithm are generalized in Section 5 to enable us to deal with infinite-state systems. In Section 6 we give sufficient conditions to guarantee that a constraint system is closed under a given transition relation. In Section 7 we describe methods to generate new well quasi-orderings from existing ones. In Section 8 we introduce UNITY-like programs and use them for syntactical representation of transition systems. In Section 9 and Section 10 we consider examples of constraint systems. We use these examples to give a general scheme for applying the methods developed of Section 5 and Section 6 to programs operating on several different data structures, such as natural numbers, sequences, real-time clocks, etc.

2 A Simple Example

In order to illustrate the ideas more concretely, let us consider a simplified version of the IEEE Futurebus Protocol, which is a transaction-based cache coherence protocol. We use a model of the protocol taken from a tutorial by Rushby [Rus97]. The following description is from [Rus97].

> The protocol consists of a memory and several processors, each with a local cache, attached to a bus. The caches maintain coherence by snooping all bus transactions and updating their local states appropriately. Memory is organized as *lines*. Each cache can store copies of values of several lines; each copy can be either **shared** or **exclusive**. When a processor needs to read a value from a memory location, it checks whether its line is present in its cache. If it is, the processor obtains the value from the cache; if not (in which case the line is said to be **invalid** for that cache), it issues a **read-shared** transaction on the bus. Any cache that has an **exclusive** copy of that line can respond by sending its value over the bus (and changing the status of its own copy to **shared** or **invalid**), otherwise the memory supplies the value. In either case, the requesting cache will load the line as a **shared** copy when its value comes over the bus (other caches may choose, opportunistically, to do the same). When a processor needs to

write a value to a memory location, it checks whether it has an **exclusive** copy of the relevant line in its cache. If not, it first obtains a **shared** copy of the line as described above (unless it already has one), then issues a **read-modified** transaction and changes the status of its copy to **exclusive**. Other caches that have a copy of the line change the status of their copies to **invalid** when they see a **read-modified** transaction. A processor that has an **exclusive** copy of a line can relinquish it by issuing a **write-back** transaction. This causes the cache to write its value back to memory and to mark the line as **invalid**.

A model of the behavior of a single cache line can be formulated as the below program

Program *IEEE Futurebus*
  declare s, e, i : natural numbers
  initially $s = e = 0$
  assign
    read-shared :
      $\langle [\!] n : 0 \leq n \leq i :: i \geq 1 \longrightarrow s, e, i := s + \mathit{sign}(e) + n \ , \ e \ominus 1 \ , \ i - n \rangle$
    $[\!]$
    read-modified :
      $s \geq 1 \longrightarrow s, e, i := 0 \ , \ e + 1 \ , \ s + i - 1$
    $[\!]$
    write-back :
      $e \geq 1 \longrightarrow s, e, i := s \ , \ e - 1 \ , \ i + 1$
end

written in UNITY notation. In this program, the variables s, e, and i denote the number of processors that have a **shared** copy, an **exclusive** copy, and an **invalid** copy of this cache line, respectively. Initially, all copies are invalid, implying that the variables s and e are initially 0. The allowed changes to the program variables are determined by three guarded multiple-assignment statements, one for each type of transaction. The statement corresponding to a read-shared transaction has a parameter n, which may range over integers from 0 to i. The parameter represents the number of processors that change state from invalid to shared. If there is at least one exclusive copy, then the number of exclusive copies decreases by 1. If that copy changes its state to shared, then there will be an extra shared copy. The notation *sign*(e) denotes

1 if $e \geq 1$, and 0 otherwise. The notation $e \ominus 1$ denotes $e - \mathit{sign}(e)$. Thus the action covers the case that n copies change from invalid to shared, and that an eclusive copy, if one exists, changes to shared. In the case that the exclusive copy changes to invalid, then the statement models a transaction in which $n+1$ copies change from invalid to shared, where $n < i$. The read-modified statement assumes that some cache is in shared; it changes that cache to exclusive, and all other shared to invalid. The write-back statement makes a copy change from exclusive to invalid.

A state of the program is given by the values of s, e and i. Let us therefore represent a state by a triple $\langle \mathtt{s}, \mathtt{e}, \mathtt{i} \rangle$ of natural numbers. Consider the problem of checking the invariant that no two cache copies are simultaneously in state **exclusive**. This amounts to checking that some state $\langle \mathtt{s}, \mathtt{e}, \mathtt{i} \rangle$ with $\mathtt{e} \geq 2$ is not reachable from any initial state of the program (recall that an initial state satisfies $s = e = 0$). This problem can be solved, using backward reachability analysis, as follows:

Let F be the set $\{\langle \mathtt{s}, \mathtt{e}, \mathtt{i} \rangle \ : \ \mathtt{e} \geq 2\}$. We use predicates which we call *constraints* to characterize sets of states. In this example we use constraints of the form $\phi_{a,b,c}$ denoting the set of states $[\![\phi_{a,b,c}]\!] = \{\langle \mathtt{s}, \mathtt{e}, \mathtt{i} \rangle \ : \ \mathtt{s} \geq a \wedge \mathtt{e} \geq b \wedge \mathtt{i} \geq c\}$. We also use finite sets of constraints, where a set Φ of constraints denotes the set $[\![\Phi]\!] = \cup_{\phi \in \Phi} [\![\phi]\!]$ of states. We perform a backward reachability analysis from F. For successively larger j, we compute Φ_j, such that $[\![\Phi_j]\!]$ is the set of states from which a state in F can be reached by a sequence of at most j executions of statements in the program. We start by $\Phi_0 = \{\phi_{0,2,0}\}$, i.e. $[\![\Phi_0]\!] = F$.

A state in $[\![\Phi_0]\!]$ can be reached either by executing **read-shared** from a state in $[\![\phi_{0,3,1}]\!]$, by executing **read-modified** from a state in $[\![\phi_{1,1,0}]\!]$, or by executing **write-back** from a state in $[\![\phi_{0,3,0}]\!]$. Thus $\Phi_1 = \Phi_0 \cup \{\phi_{0,3,1}, \phi_{1,1,0}, \phi_{0,3,0}\} = \{\phi_{0,2,0}, \phi_{0,3,1}, \phi_{1,1,0}, \phi_{0,3,0}\}$. Observe that the constraints in Φ_1 are either included in Φ_0, or are obtained by taking the preimage with respect to a program statement. At this point, we note that $\phi_{0,3,1}$ and $\phi_{0,3,0}$ are both *covered* by $\phi_{0,2,0}$, i.e. $[\![\phi_{0,3,1}]\!] \subseteq [\![\phi_{0,2,0}]\!]$ and $[\![\phi_{0,3,0}]\!] \subseteq [\![\phi_{0,2,0}]\!]$. Therefore it is meaningless to add $\phi_{0,3,1}$ and $\phi_{0,3,0}$ to Φ_1, since $\phi_{0,2,0}$ is already in Φ_0. Thus $\Phi_1 = \{\phi_{0,2,0}, \phi_{1,1,0}\}$.

A state in Φ_1 can be reached either by executing **read-shared** from a state in $[\![\{\phi_{0,3,1}, \phi_{0,2,1}\}]\!]$, by executing **read-modified** from a state in $[\![\phi_{1,1,0}]\!]$, or by executing **write-back** from a state in $[\![\{\phi_{0,3,0}, \phi_{1,2,0}\}]\!]$. Thus $\Phi_2 =$

$\{\phi_{0,2,0}, \phi_{1,1,0}, \phi_{0,3,1}, \phi_{0,2,1}, \phi_{1,1,0}, \phi_{0,3,0}, \phi_{1,2,0}\}$, which by removing redundant constraints can be written as $\{\phi_{0,2,0}, \phi_{1,1,0}\}$. As a result, we obtain $\Phi_2 = \{\phi_{0,2,0}, \phi_{1,1,0}\} = \Phi_1$.

A simulation of the verification is described in Figure 1 as follows. For each j, we describe the constraints in Φ_j. We also show the constraints generated by applying the statements **r-s** (read-shared), **r-s** (read-modified), and **r-s** (write-back) backwards to the constraints in Φ_j. We describe for each constraint, whether it is covered by an already generated constraint, and hence discarded, or added to Φ_{j+1}.

j	Φ_j	rule	new constraints generated	covered / added
0	$\phi_{0,2,0}$	**r-s**	$\phi_{0,3,1}$	covered by $\phi_{0,2,0}$
		r-m	$\phi_{1,1,0}$	added to Φ_1
		w-b	$\phi_{0,3,0}$	covered by $\phi_{0,2,0}$
1	$\phi_{0,2,0}$, $\phi_{1,1,0}$	**r-s**	$\phi_{0,3,1}$, $\phi_{0,2,1}$	covered by $\phi_{0,2,0}$
		r-m	$\phi_{1,1,0}$	already in Φ_1
		w-b	$\phi_{0,3,0}$, $\phi_{1,2,0}$	covered by $\phi_{0,2,0}$

Figure 1: Table representing verification of the simplified example

Notice that $\Phi_2 = \Phi_1$, implying that $\Phi_j = \Phi_1$ for all $j \geq 1$. We conclude that F is reachable only from states in $[\![\Phi_1]\!]$. Since $\langle 0, 0, k\rangle \notin [\![\Phi_1]\!]$ for any $k \geq 0$ (an easy check), we conclude that F is not reachable from any initial state of the program.

Let us examine the above procedure in order to find out what general principles emerge. We are computing a fixpoint by an iterative procedure. Each approximation Φ_j can be described as a finite set of constraints. For a statement s, let $pre(s, \Phi)$ denote the set of states from which a state in Φ is reachable through one application of s. The following properties are essential for our algorithm.

1. For each finite set Φ of constraints we should be able to compute $pre(s, \Phi)$ as a finite set of constraints, i.e. , the constraint system should be closed under applications of the three program statements. In Section 6, we show that this property is equivalent to the statements of the

program being *monotone*, in the sense that if a statement transforms the state $\langle \mathtt{s}, \mathtt{e}, \mathtt{i} \rangle$ into $\langle \mathtt{s}', \mathtt{e}', \mathtt{i}' \rangle$ then any state $\langle \mathtt{a}, \mathtt{b}, \mathtt{c} \rangle$ with $\mathtt{s} \leq \mathtt{a}$, $\mathtt{e} \leq \mathtt{b}$, and $\mathtt{i} \leq \mathtt{c}$ is transformed into a state $\langle \mathtt{a}', \mathtt{b}', \mathtt{c}' \rangle$ with $\mathtt{s}' \leq \mathtt{a}'$, $\mathtt{e}' \leq \mathtt{b}'$, and $\mathtt{i}' \leq \mathtt{c}'$. We can then conclude that if F can be represented by a finite set of constraints, then all Φ_j can be computed as finite sets of constraints.

2. Several times, we were able to discard constraints $\phi_{\mathsf{s},\mathsf{e},\mathtt{i}}$ since they *entailed* (were covered by) some other constraint $\phi_{\mathsf{a},\mathsf{b},\mathsf{c}}$. More precisely, $\phi_{\mathsf{s},\mathsf{e},\mathtt{i}}$ *entails* $\phi_{\mathsf{a},\mathsf{b},\mathsf{c}}$ if $\mathtt{a} \leq \mathtt{s}$, $\mathtt{b} \leq \mathtt{e}$, and $\mathtt{i} \leq \mathtt{c}$, (implying $[\![\phi_{\mathsf{s},\mathsf{e},\mathtt{i}}]\!] \subseteq [\![\phi_{\mathsf{a},\mathsf{b},\mathsf{c}}]\!]$). In this way, the growth of the Φ_j's was slowed down, so that eventually convergence was reached. In this example, it turns out that entailment among constraints of form $\phi_{\mathsf{s},\mathsf{e},\mathtt{i}}$ has an important property, which guarantees convergence; namely, entailment is a *well quasi-ordering* (Section 3.2 and Section 7). Roughly, this means that in any sequence of added constraints of form $\phi_{\mathsf{s},\mathsf{e},\mathtt{i}}$, we will eventually reach a situation where any subsequent constraint will be redundant, and hence discarded. Intuitively, we could note this effect in the example: "fresh" constraints $\phi_{\mathsf{s},\mathsf{e},\mathtt{i}}$ had to introduce a component (a or b) which was less than the already produced constraints. Clearly, we cannot produce an infinite sequence of constraints that are "fresh" in this sense.

This general observation can be used to conclude that, e.g., for programs over natural numbers, containing only monotone statements, the problem of checking reachability of a constraint of the above form is decidable. We observe that Petri nets are a special case of such programs, and conclude that the coverability problem for Petri nets is decidable. These ideas can be generalized (Section 7, Section 9, and Section 10) to much richer classes of programs, e.g., to programs that operate over compound data structures such as sequences, sets, and multi-sets. The only thing that must be done is to generalize the ordering to the data domain in question, and to consider programs whose statements are monotone with respect to this ordering. If the ordering is still a well quasi-ordering, this guarantees decidability of the reachability problem.

3 Preliminaries

We introduce the notions of transition systems and well quasi-orderings.

3.1 Transition Systems

We present the basic definitions for transition systems.

Definition 3.1 A *transition system* $\mathcal{T}$ is a pair $\langle \Sigma, \longrightarrow \rangle$, where Σ is a set of *states*, and $\longrightarrow \subseteq \Sigma \times \Sigma$ is a *transition relation*. $\square$

According to Definition 3.1, a transition relation denotes a set of pairs of states. Each such pair is called a *transition*. We use $\sigma_1 \longrightarrow \sigma_2$ to denote that $(\sigma_1, \sigma_2) \in \longrightarrow$. Let $\longrightarrow^*$ denote the transitive and reflexive closure of $\longrightarrow$. We say that a state σ_2 is *reachable* from a state σ_1 if $\sigma_1 \longrightarrow^* \sigma_2$. For two sets *Init* and F of states, we say that F is *reachable* from *Init* if some state σ_F in F is reachable from a state σ_I in *Init*.

For a set $S \subseteq \Sigma$ of states and a binary relation $\longrightarrow$ on Σ, we use $pre(\longrightarrow, S)$ to denote the set $\{\sigma \ : \ \exists \sigma' \in S : \sigma \longrightarrow \sigma'\}$ of states from which a state in S can be reached via a transition in $\longrightarrow$.

The reachability problem is defined as follows.

Instance: A transition system $\langle \Sigma, \longrightarrow \rangle$ and two sets of states *Init* and F.

Question: Is F reachable from *Init*?

The reachability problem is central to automated verification, since it can be used to verify a large class of safety properties of transition systems [VW86, GW93].

3.2 Well Quasi-Orderings

We introduce the notion of *well quasi-orderings*.

A *quasi-order*[1] $\preceq$ on A is a binary relation over A which is reflexive and transitive. A set $I \subseteq A$ is said to be an *ideal (with respect to $\preceq$)* if it is

[1] Frequently, the term *preorder* is used instead.

the case that $a \in I$ and $a \preceq b$ imply $b \in I$. The *ideal generated by* $a \in A$, denoted $id(\preceq, a)$, is defined to be the set $\{b \ : \ a \preceq b\}$. We say that $B \subseteq A$ is a *minor set* of A, if (i) for all $a \in A$ there exists $b \in B$ such that $b \preceq a$, and (ii) $a, b \in B$ and $a \preceq b$ imply $a = b$.

Definition 3.2 A quasi-order $\preceq$ on a set A is a *well quasi-ordering* (wqo) if in each infinite sequence $a_0 \ a_1 \ a_2 \ a_3 \cdots$ of elements in A, there are indices $i < j$ such that $a_i \preceq a_j$. □

Intuively, a well quasi-ordering has the property that each infinite sequence of elements contains a pair of ordered elements. We observe that each well quasi-ordering is well-founded. However the converse is not true. For example the prefix relation among strings over a given alphabet is well-founded but not a well quasi-ordering.

Example 3.3 An example of a well quasi-ordering is the identity relation on any finite set. Another example is the "less-than-or-equal" relation $\leq$ on the set $\mathcal{N}$ of natural numbers. However, the relation $\leq$ is *not* a well quasi-ordering on the set of integers, nor on the set of nonnegative rational numbers.

Proposition 3.4 *For a set A and a well quasi-ordering $\preceq$ on A, there exists at least one finite minor set of A.*

Proof: Suppose that no finite minor set of A exists. We show that $\preceq$ is not a well quasi-ordering. We define the infinite sequence $a_0, a_1, a_2, \ldots$ of elements in A as follows. Let a_0 be any arbitrary element in A. We choose a_{i+1} such that $a_j \not\preceq a_{i+1}$ for each $j : 1 \leq j \leq i$. The element a_{i+1} exists, since otherwise $\{a_0, a_1, \ldots, a_i\}$ would be a minor set of A, contradicting the assumption that no such sets exist. It is clear that the sequence $a_0, a_1, a_2, \ldots$ violates the well quasi-orderedness property. □

4 Finite-State Verification

In this section we review a standard approach (*Reachable*1, Figure 2) to solving the reachability problem in the case that the set Σ of states of the transition system $\langle \Sigma, \longrightarrow \rangle$ is finite. The reachability problem can be solved by systematically enumerating all states σ from which some state $\sigma_F \in F$ is reachable (i.e., computing $pre(\longrightarrow^*, F)$), and then checking whether some state in *Init* is contained in this set. In Figure 2 we give a naive presentation of this idea.

Procedure *Reachable*1
Input
    $\mathcal{T} = \langle \Sigma, \longrightarrow \rangle$: transition system
    Init: set of initial states
    F: set of final states finite set of constraints
Output are there $\sigma \in Init$ and $\sigma' \in [\![\Phi_F]\!]$ such that $\sigma \longrightarrow^* \sigma'$?

var V : set of states
begin
    $V := F$;
    while $pre(\longrightarrow, V) \not\subseteq V$ **do** $V := V \cup pre(\longrightarrow, V)$ **od** ;
      if $Init \cap V = \emptyset$ **then return** unreachable **else return** reachable
end

Figure 2: Algorithm for deciding reachability

The above procedure is guaranteed to terminate since the value of the variable V increases at each iteration of the while-loop, but is bounded by the finite set Σ.

Let us refine *Reachable*1 into a more pragmatic version (*Reachable*2, Figure 3) which does not manipulate entire sets, but rather computes preimages for individual states. This is the way that reachability analysis is normally implemented. In this refinement, the set V (in *Reachable*1) is represented (in *Reachable*2) as the union of two sets: a set V of states whose predecessors have already been generated, and a set W of states whose predecessors have not yet been generated. The analysis works by repeatedly choosing from the already generated states in W one which is so far "unexplored" and adding its predecessors to the set of generated states. If during the exploration some state in *Init* is generated, then F is reachable from *Init*, otherwise F is not reachable from *Init*.

A description of the algorithm is given in Figure 3. In the algorithm, the set

Procedure *Reachable*2
Input
  $\mathcal{T} = \langle \Sigma, \longrightarrow \rangle$: transition system
  Init: set of initial states
  F: set of final states finite set of constraints
Output are there $\sigma \in Init$ and $\sigma' \in [\![\Phi_F]\!]$ such that $\sigma \longrightarrow^* \sigma'$?

var W , V : sets of states
begin
  $W, V := F, \emptyset$;
  while $W \neq \emptyset$ **do**
    choose $\sigma \in W$;
    $W := W \setminus \{\sigma\}$;
    if $\sigma \in Init$ **then return** *reachable*
    else
      if $\sigma \notin V$
      then
        $V := V \cup \{\sigma\}$;
        $W := W \cup pre(\longrightarrow, \{\sigma\})$
  od ;
  return *unreachable*
end

Figure 3: Algorithm for deciding reachability

W is the set of visited but still unexplored states, whereas V is the set of visited and explored states. One step in the exploration consists of choosing an unexplored state σ. If $\sigma \in Init$ then we have found a path "backwards" from F to *Init*. Otherwise we check whether σ is already explored. If not, we add the predecessors of σ to the set of unexplored states, and move σ to the set of explored states. The algorithm in Figure 3 trivially terminates because there are only a finite number of states that can be added to V.

5 Symbolic Verification

In this section drop the assumption that the program is finite-state. We investigate a "symbolic" generalization of the the algorithm in Figure 3. Instead of letting the reachability algorithm manipulate individual states, we let it manipulate sets of states, represented by predicates which we call *constraints*. Each constraint may characterize a finite or an infinite set of states. In our generalization, we will require that F is represented as a finite union Φ_F of constraints, while *Init* is represented by the negation of a finite union Φ_{Init} of constraints. The sets V and W in the program will also be finite unions of constraints. In each step of the procedure, we consider an "unexplored" constraint and generate its predecessors, again represented by a finite union of constraints. Note that we here make the nontrivial assumption that we can indeed represent the predecessors of a constraint as a finite union of constraints. If the search generates some constraint representing a set of states that has a nonempty intersection with *Init*, then F is reachable from *Init*, otherwise F is not reachable from *Init*.

The choice of constraints in the reachability analysis may depend on the program, so let us define the notion of constraint system relative to a given transition system.

Definition 5.1 Let $\mathcal{T} = \langle \Sigma, \longrightarrow \rangle$ be a transition system. A *constraint system* $\mathcal{C}$ for $\langle \Sigma, \longrightarrow \rangle$ is a set of objects, called *constraints*, where to each ϕ is assigned a denotation $[\![\phi]\!] \subseteq \Sigma$. We define a quasi-order $\sqsubseteq$ on $\mathcal{C}$, where $\phi \sqsubseteq \phi'$ if $[\![\phi']\!] \subseteq [\![\phi]\!]$. Intuitively, $\phi \sqsubseteq \phi'$ denotes that ϕ' "entails" ϕ, or that ϕ' is "stronger than" ϕ. A set Φ of constraints denotes the union of the denotations of its elements, i.e., $[\![\Phi]\!] = \cup_{\phi \in \Phi} [\![\phi]\!]$. The definition of $\sqsubseteq$ is extended to sets of constraints, so that $\Phi \sqsubseteq \Phi'$ if $[\![\Phi']\!] \subseteq [\![\Phi]\!]$. In the sequel we assume that $\sqsubseteq$ is computable among finite sets of constraints.

Example 5.2 In the example of Section 2, we have e.g. $\phi_{0,2,0} \sqsubseteq \phi_{0,3,1}$.

In the context of finite transition systems, one of the most well-known constraint systems is *Binary Decision Diagrams (BDDs)* [Bry86]. Each BDD is a constraint representing a propositional logical formula, and thus characterizes a subset of Σ. In this paper, we consider constraint systems which allow us to analyze infinite state spaces.

Definition 5.3 For a transition system $\mathcal{T} = \langle \Sigma, \longrightarrow \rangle$ and a constraint system $\mathcal{C}$, we say that $\mathcal{C}$ is *(effectively) closed* with respect to $\mathcal{T}$, if for each $\phi \in \mathcal{C}$, we can compute a finite set $\Phi \subseteq \mathcal{C}$ such that $pre(\longrightarrow, [\![\phi]\!]) = [\![\Phi]\!]$. □

Definition 5.4 We say that a constraint system $\mathcal{C}$ is *well quasi-ordered* if the relation $\sqsubseteq$ is a well quasi-ordering on the elements of $\mathcal{C}$.

We can now generalize the procedure of Figure 3 in a straight-forward manner, obtaining the procedure in Figure 4. The procedure assumes that we have selected a constraint system which is effectively closed with respect to the transition system. The procedure *Reachable*3 in Figure 4 works in the same way as *Reachable*2 in Figure 3, with individual states being replaced by constraints. A more natural version of the insertion test is the condition $[\![\phi]\!] \not\subseteq [\![V]\!]$. This test may be expensive to perform, in the case where V contains a large number of constraints. In many practical situations it is cheaper to have the weaker insertion test shown in Figure 4.

The following theorem gives sufficient conditions for decidability of the reachability algorithm.

Theorem 5.5 *For a well quasi-ordered constraint system $\mathcal{C}$, and a transition system $\mathcal{T}$, if $\mathcal{C}$ is effectively closed with respect $\mathcal{T}$ then the reachability problem is decidable.*

Proof: Since $\mathcal{C}$ is closed with respect to $\mathcal{T}$, and $\sqsubseteq$ is computable it follows that each iteration of the loop can be performed effectively. Observe that the test $[\![\phi]\!] \cap \mathit{Init} \neq \emptyset$ is equivalent to the negation of $\Phi_{\mathit{Init}} \sqsubseteq \{\phi\}$, which can be checked since Φ_{Init} is finite and $\sqsubseteq$ is assumed (Definition 5.1) to be computable among finite sets of constraints.

Termination follows from the following argument. Consider the sequence $\phi_1\phi_2\phi_3\ldots$ of constraints that are added to V. Due to the test $\not\exists \phi' \in V : \phi' \sqsubseteq \phi$, we have $\phi_i \not\sqsubseteq \phi_j$ for all $i < j$. Since $\sqsubseteq$ is a well quasi-ordering, the sequence cannot be infinite, and hence each execution the algorithm will terminate after a finite number of iterations of the main loop. □

From Theorem 5.5 we conclude that two main challenges in designing Reachability3 for a certain class of systems, are

Procedure $Reachable3(\langle \Sigma, \longrightarrow \rangle, Init, F)$
Input
  $\mathcal{T} = \langle \Sigma, \longrightarrow \rangle$: transition system
  $Init$: set of states, $Init = \neg [\![\Phi_{Init}]\!]$ for some finite set Φ_{Init} of constraints
  Φ_F: finite set of constraints
Output are there $\sigma \in Init$ and $\sigma' \in [\![\Phi_F]\!]$ such that $\sigma \longrightarrow^* \sigma'$?

```
var W , V : sets of constraints
begin
   W, V := Φ_F, ∅ ;
   while W ≠ ∅ do
      choose φ ∈ W ;
      W := W \ {φ} ;
      if ⟦φ⟧ ∩ Init ≠ ∅ then return reachable
      else
         if ∄φ' ∈ V : φ' ⊑ φ                    (* insertion test *)
         then
            V := V ∪ {φ} ;
            W := W ∪ Φ where ⟦Φ⟧ = pre(⟶, ⟦φ⟧) ;
   od
   return unreachable
end
```

Figure 4: Symbolic Reachability Algorithm

1. *closedness*: to prove the closedness of the constraint system with respect the transition system. In Section 6 we describe sufficient conditions for achieving closedness.

2. *termination:* to investigate conditions under which termination of the algorithm can be guaranteed. In this paper we achieve termination through the assumption that our constraint systems are well quasi-ordered. In Section 7 we present a method to generate new well quasi-orderings from existing ones.

6 Closedness

We introduce the notion of *monotonicity* for transition systems. We show (Theorem 6.2 and Theorem 6.8) that the closedness of a constraint system ϕ with respect to a transition system $\mathcal{T}$, can be characterized as whether $\mathcal{T}$ is *monotone* with respect to a preorder derived from $\mathcal{C}$ in a natural way.

Definition 6.1 Let $\mathcal{T} = \langle \Sigma, \longrightarrow \rangle$ be a transition system, with a preorder $\preceq$ defined on Σ. We say that $\mathcal{T}$ is *monotone with respect to* $\preceq$ if, for any states σ_1, σ_2, and σ_3, with $\sigma_1 \preceq \sigma_2$ and $\sigma_1 \longrightarrow \sigma_3$, there exists a state σ_4 such that $\sigma_3 \preceq \sigma_4$ and $\sigma_2 \longrightarrow \sigma_4$.

Theorem 6.2 *For a transition system $\mathcal{T} = \langle \Sigma, \longrightarrow \rangle$, which is monotone with respect to a preorder $\preceq$ on $\mathcal{T}$, if $I \subseteq \Sigma$ is an ideal, then $pre(\longrightarrow, I)$ is also an ideal.*

Proof: Suppose that $\sigma_1 \in pre(\longrightarrow, I)$ and $\sigma_1 \preceq \sigma_2$. We show that $\sigma_2 \in pre(\longrightarrow, I)$. We know that there is $\sigma_3 \in I$ such that $\sigma_1 \longrightarrow \sigma_3$. By monotonicity it follows that there is a state σ_4 such that $\sigma_3 \preceq \sigma_4$ and $\sigma_2 \longrightarrow \sigma_4$. Since I is an ideal, we get $\sigma_4 \in I$, and hence $\sigma_2 \in pre(\longrightarrow, I)$. □

Definition 6.3 Let $\mathcal{C}$ be a constraint system. We define a preorder $\preceq_{\mathcal{C}}$ on Σ, such that $\sigma_1 \preceq_{\mathcal{C}} \sigma_2$ if and only $\sigma_1 \in [\![\phi]\!]$ implies $\sigma_2 \in [\![\phi]\!]$, for each constraint $\phi \in \mathcal{C}$.

Proposition 6.4 *For a constraint system $\mathcal{C}$, each constraint $\phi \in \mathcal{C}$ is an ideal with respect to $\preceq_{\mathcal{C}}$.*

Proof: The proof follows immediately from the definitions. $\square$

Definition 6.5 A constraint system $\mathcal{C}$ is said to be *disjunctive* if for each $\phi_1, \phi_2 \in \mathcal{C}$, there is a (possibly infinite) set $\Psi \subseteq \mathcal{C}$ such that $[\![\phi_1]\!] \cap [\![\phi_2]\!] = \bigcup_{\phi \in \Psi} [\![\phi]\!]$.

Proposition 6.6 *For a disjunctive constraint system $\mathcal{C}$, and a transition system $\mathcal{T} = \langle \Sigma, \longrightarrow \rangle$, if $I \subseteq \Sigma$ is an ideal with respect to $\preceq_{\mathcal{C}}$, then there is a (possibly infinite) set $\Psi \subseteq \mathcal{C}$ such that $I = \bigcup_{\phi \in \Psi} [\![\phi]\!]$.*

Proof: We define $\Psi = \bigcup_{[\![\phi]\!] \subseteq I} \phi$. It is obvious that $[\![\Psi]\!] \subseteq I$. We show that $I \subseteq [\![\Psi]\!]$. Suppose that $\sigma \in I$. We prove that $\sigma \in [\![\Psi]\!]$.

From the definitions we know that $id(\preceq_{\mathcal{C}}, \sigma) = \bigcap_{\sigma \in [\![\phi]\!]} [\![\phi]\!]$. Since $\sigma \in I$ and I is an ideal, it follows that $id(\preceq_{\mathcal{C}}, \sigma) \subseteq I$, i.e. $\bigcap_{\sigma \in [\![\phi]\!]} [\![\phi]\!] \subseteq I$. We know that $\mathcal{C}$ is disjunctive, and hence there exists a set $\Psi' \subseteq \mathcal{C}$ such that $\bigcap_{\sigma \in \phi} [\![\phi]\!] = \bigcup_{\phi \in \Psi'} [\![\phi]\!]$. From the above it follows that $\sigma \in id(\preceq_{\mathcal{C}}, \sigma) = \bigcup_{\phi \in \Psi'} [\![\phi]\!]$ and hence there is a $\phi \in \Psi'$ such that $\sigma \in [\![\phi]\!]$. It also follows that $\bigcup_{\phi \in \Psi'} [\![\phi]\!] \subseteq I$ and hence $[\![\phi]\!] \subseteq I$ for each $\phi \in \Psi'$. This implies that there is a constraint ϕ such that $[\![\phi]\!] \subseteq I$ and $\sigma \in [\![\phi]\!]$. By definition of Ψ it follows that $\sigma \in [\![\Psi]\!]$. $\square$

Corollary 6.7 *For a disjunctive and well quasi-ordered constraint system $\mathcal{C}$, and a transition system $\mathcal{T} = \langle \Sigma, \longrightarrow \rangle$, if $I \subseteq \Sigma$ is an ideal with respect to $\preceq_{\mathcal{C}}$, then there is a finite set $\Phi \subseteq \mathcal{C}$ such that $I = \bigcup_{\phi \in \Phi} [\![\phi]\!]$.*

Proof: From Proposition 6.6, we know that there is a set $\Psi \subseteq \mathcal{C}$ such that $I = \bigcup_{\phi \in \Psi} [\![\phi]\!]$. We take Φ to be any minor set of Ψ. The existence of Φ follows from Proposition 3.4 and the well quasi-orderedness of $\mathcal{C}$. $\square$

Observe that the properties described in Proposition 6.6 and Corollary 6.7 are not dependent on the transition relation $\longrightarrow$.

Theorem 6.8 *A disjunctive and well quasi-ordered constraint system $\mathcal{C}$ is closed with respect to a transition system $\mathcal{T}$, if and only if $\mathcal{T}$ is monotone with respect to $\preceq_{\mathcal{C}}$.*

Proof: *(if):* Suppose that $\mathcal{T}$ is monotone with respect to $\preceq_{\mathcal{C}}$. Let $\phi \in \mathcal{C}$. We show that there is a finite set $\Phi \subseteq \mathcal{C}$ such that $pre(\longrightarrow, [\![\phi]\!]) = [\![\Phi]\!]$.

By Proposition 6.4 it follows that ϕ is an ideal with respect to $\preceq_{\mathcal{C}}$. By Theorem 6.2 we know that $pre(\longrightarrow, [\![\phi]\!])$ is an ideal. The result follows Corollary 6.7.

(only if): Suppose that $\sigma_1 \preceq \sigma_2$ and $\sigma_1 \longrightarrow \sigma_3$. We show that there exists a state σ_4 such that $\sigma_3 \preceq \sigma_4$ and $\sigma_2 \longrightarrow \sigma_4$. By Corollary 6.7 it follows that there is a finite set $\Phi \subseteq \mathcal{C}$ such that $[\![\Phi]\!] = id(\preceq_{\mathcal{C}}, \sigma_3)$. Since $\mathcal{C}$ is closed with respect to $\mathcal{T}$, it follows that there is a finite $\Phi' \subseteq \mathcal{C}$ such that $[\![\Phi']\!] = pre(\longrightarrow, [\![\Phi]\!])$. We know that there is a $\phi \in \Phi'$ such that $\sigma_1 \in \phi$. From the fact that $\sigma_1 \preceq_{\mathcal{C}} \sigma_2$ we get $\sigma_2 \in [\![\phi]\!]$, and hence $\sigma_2 \in pre(\longrightarrow, [\![\Phi]\!])$. This implies that there is a state σ_4 such that $\sigma_2 \longrightarrow \sigma_4$ and $\sigma_4 \in [\![\Phi]\!] = id(\preceq_{\mathcal{C}}, \sigma_3)$. Consequently $\sigma_3 \preceq_{\mathcal{C}} \sigma_4$. □

7 Building Well Quasi-Orders

We will describe techniques for generating well-quasi ordered constraint systems on compound data structures such as sequences, bags, sets, etc. To do that we restate two standard lemmas, which allow us to lift well quasi-orderings from elements to bags and to sequences. Let A^* denote the set of finite strings over A, and let A^B denote the set of finite bags over A. For a natural number n, let $\widehat{n}$ denote the set $\{1, \ldots, n\}$. An element of A^* and of A^B can be represented as a mapping $w\colon \widehat{|w|} \mapsto A$ where $|w|$ is the size of the bag or the length of the sequence. Given a quasi-order $\preceq$ on a set A, define the quasi-order $\preceq^*$ on A^* by letting $w \preceq^* w'$ if and only if there is a monotone[2] injection $h\colon \widehat{|w|} \mapsto \widehat{|w'|}$ such that $w(j) \preceq w'(h(j))$ for $1 \leq j \leq |w|$. Define the quasi-order $\preceq^B$ on bags of A by $w \preceq^* w'$ if and only if there is a (not necessarily monotone) injection $h\colon \widehat{|w|} \mapsto \widehat{|w'|}$ such that $w(j) \preceq w'(h(j))$ for $1 \leq j \leq |w|$.

Lemma 7.1 *If $\preceq$ is a wqo on A, then $\preceq^*$ is a wqo on A^* and $\preceq^B$ is a wqo on A^B.*

Proof: The proof can be found in [Hig52]. □

In the following, we give some examples of applications of Lemma 7.1.

[2] meaning that $h(j_1) \leq h(j_2)$ if and only if $j_1 \leq j_2$

Example 7.2 1. The equality relation on a finite set A is trivially a well quasi-ordering.

2. The *less-than-equal* relation $\leq$ on the $\mathcal{N}$ of natural numbers is a well quasi-ordering.

3. Consider the set A^* of finite strings over a finite set A. Let $\preceq$ denote the *substring relation* on A^*, i.e. $w_1 \preceq w_2$ if w_1 is a (not necessarily contiguous) substring of w_2. According to Lemma 7.1 and 1, the relation is a well quasi-ordering (Higman's lemma [Hig52]).

4. Consider the *subbag relation* $\preceq$ on the set $\mathcal{N}^B$ of bags of natural numbers. The relation is defined by $b_1 \preceq b_2$ if there is an injection from the elements of b_1 to the elements of b_2, such that $x \leq h(x)$, for each element x in b_1. According Lemma 7.1 and 2, the relation is a well quasi-ordering.

5. Consider the set of vectors of natural numbers of length n for some $n \geq 1$. Define a preorder $\preceq$ on the set such that $\langle x_1, \ldots, x_n \rangle \preceq \langle y_1, \ldots, y_n \rangle$ if $x_i \leq y_i$ for each $i : 1 \leq i \leq n$. According Lemma 7.1 and 2, the relation is a well quasi-ordering (Dickson's lemma [Dic13]).

8 Programs

In this section, we present *programs* and use them as a simple notation for syntactical representation of transition systems. The notation is very similar to the action systems by Back and Kurki-Suonio [BKS88] and to UNITY by Chandy and Misra [CM88]

We assume a set of *domains*. We assume an *assertion language* containing typed variables, functions, constants and predicates, for forming expressions and assertions. We use $\mathrm{dom}(x)$ to denote the domain of the variable x, and more generally we use $\mathrm{dom}(e)$ to denote the domain of the expression e. An assertion ψ with free variables $x_1, \ldots, x_n$ denotes a subset $[\![\psi]\!]$ of $\mathcal{D}_1 \times \cdots \times \mathcal{D}_n$, where $\mathcal{D}_i = \mathrm{dom}(x_i)$, in the usual way. We will use an overbar notation for tuples and write, e.g., $\overline{x}$ instead of $x_1, \ldots, x_n$.

Definition 8.1 A *program* P consists of

- a set V of (typed) *program variables*,
- a finite set of *actions*. Each action is a guarded command of form

  $$g \rightarrow v_1, \ldots, v_n := e_1, \ldots, e_n$$

  where the guard g is an assertion, $v_1, \ldots, v_n$ are program variables, and $e_1, \ldots, e_n$ are expressions[3]. The guard g and the expressions $e_1, \ldots, e_n$ may contain free variables, some of which (but not necessarily all) are program variables. The free variables in the action which are not in V are called the *parameters* of the action. $\square$

A *variant* of α is an action of form $g[\overline{d}/\overline{x}] \rightarrow \overline{v} := \overline{e}[\overline{d}/\overline{x}]$ obtained by replacing the parameters $x_1, \ldots, x_m$ by values $d_1, \ldots, d_m$ such that $d_i \in \mathrm{dom}(x_i)$ for $i = 1, \ldots, m$.

Definition 8.2 A program P denotes a transition system $[\![P]\!] = \langle \Sigma, \longrightarrow \rangle$, where

Σ is the set of mappings from program variables to values, such each $v \in V$ is mapped to a value $\sigma(v) \in \mathrm{dom}(v)$. A state[4]can be extended to a mapping from expressions and assertions in the natural way.

$\longrightarrow$ is the set of pairs (σ, σ') such that for some variant $g \rightarrow \overline{v} := \overline{e}$ of some action of P, we have

- $\sigma(g) = \mathit{true}$, i.e., g is true in σ, and
- $\sigma' = \sigma[\sigma(\overline{e})/\overline{v}]$ i.e., σ' is obtained from σ by performing the assignment $\overline{v} := \overline{e}$.

Let us now assume that each domain $\mathcal{D}_i$ is equipped with a preorder $\preceq$.

Definition 8.3 Let $g(\overline{v}, \overline{x}) \rightarrow x_1 := e_1(\overline{v}, \overline{x}), \ldots, x_n := e_n(\overline{v}, \overline{x})$ be a guarded command, where $\overline{v}$ and $\overline{x}$ denote the sets of program variables and parameters

[3]We assume that the assignment is "type-correct" in the sense that $\mathrm{dom}(v_i) = \mathrm{dom}(e_i)$, i.e., that the types of v_i and e_i coincide for $i = 1, \ldots, n$.

[4]Observe that each tuple $\langle d_1, \ldots, d_n \rangle \in \langle \mathcal{D}_1 \times \cdots \times \mathcal{D}_n \rangle$ defines a unique state $\sigma \in \Sigma$, where $\sigma(x_i) = d_i$ for $i : 1 \leq i \leq n$.

respectively. We say that the guarded command is *monotone* if $g(\overline{d}_1, \overline{a}_1)$ and $\overline{d}_1 \preceq \overline{d}_2$ imply that there is a $\overline{a}_2$ such that $g(\overline{d}_2, \overline{a}_2)$ and $e_i(\overline{d}_1, \overline{a}_1) \preceq e_i(\overline{d}_2, \overline{a}_2)$, for $i : 1 \leq i \leq n$. A program P is said to be *monotone* if all the guarded commands in P are monotone.

Lemma 8.4 *If a program P is monotone with respect to a preorder $\preceq$, then $[\![P]\!]$ will also be monotone with respect to $\preceq$.*

Proof: The proof follows from the definitions. □

9 Unary Constraint Systems

In this section, we introduce a class of constraint systems called *unary constraint systems*. Assume a program P with the set $V = \{v_1, \ldots, v_n\}$ of program variables. Assume a set of domains each with a decidable[5] preorder defined on it. Let[6] $\langle d_1, \ldots, d_m \rangle \preceq \langle d'_1, \ldots, d'_m \rangle$ denote $d_i \preceq d'_i$ for $i : 1 \leq i \leq m$. A *unary constraint* is of the form form $\phi_{\bar{d}}$, where $\bar{d} \in \mathrm{dom}(v_1) \times \cdots \times \mathrm{dom}(v_n)$, and $[\![\phi_{\bar{d}}]\!] = \{\bar{v} \ : \ \bar{d} \preceq \bar{v}\}$. Observe $\phi_{\bar{d}_1} \sqsubseteq \phi_{\bar{d}_2}$ if and only if $\bar{d}_1 \preceq \bar{d}_2$, and hence it follows by decidability of $\preceq$ that $\sqsubseteq$ is also decidable. Notice that in the case of unary constraint systems the relation $\preceq_{\mathcal{C}}$, derived from the constraint system, coincides with the relation $\preceq$ on the set of states.

A function F is said to be *monotone* if $\overline{y}_1 \preceq \overline{y}_2$ implies $F(\overline{y}_1) \preceq F(\overline{y}_2)$.

In Section 9.1 and Section 9.2 we give examples of unary constraint systems. Furthermore all the programs in these sections are of a certain form (described in the following proposition) which guarantees monotonicity.

Proposition 9.1 *Let P be a program. If each action of P is of form*

$$\overline{d}(\overline{x}) \preceq \overline{f}(\overline{v}, \overline{x}) \to \overline{v} := \overline{e}(\overline{v}, \overline{x})$$

where all functions occurring in the expressions $\overline{f}$ and $\overline{e}$ are monotone, then P is monotone.

Proof: The proof follows from the definitions. □

[5] By decidability of $\preceq$ we mean that $d_1 \preceq d_2$ can be computed for any d_1 and d_2.
[6] We assume that d_i and d'_i belong to the same domain for $i : 1 \leq i \leq m$.

9.1 Programs over Natural Numbers

Let $\mathcal{N}$ be the set of natural numbers equipped with the standard less-than or equal ordering $\leq$, which is also a well quasi-ordering (Section 7). The relation is extended in the usual way to tuples of natural numbers. We consider programs that operate on a finite set $\{v_1, \ldots, v_n\}$ of variables with domain $\mathcal{N}$. Each state σ is defined by a tuple $\langle d_1, \ldots, d_n \rangle$ of natural numbers. We consider a constraint system $\mathcal{C}_\mathcal{N}$, where each constraint in $\mathcal{C}_\mathcal{N}$ is of the form $\phi_{\overline{d}}$, denoting $\{\overline{v} \; : \; \overline{d} \leq \overline{v}\}$. Observe that for finite sets $\Phi_1, \Phi_2 \subseteq \Phi$, we have $\Phi_1 \sqsubseteq \Phi_2$ if and only if for each $\phi_{\overline{d}_2} \in \Phi_2$ there is a constraint $\phi_{\overline{d}_1} \in \Phi_1$ such that $\overline{d}_1 \leq \overline{d}_2$. Observe also that the relations $\preceq_{\mathcal{C}_\mathcal{N}}$ and $\leq$ coincide. It follows that $\mathcal{C}_\mathcal{N}$ is well quasi-ordered and is trivially disjunctive.

From Proposition 9.1 we get the monotonicity (with respect to $\preceq_{\mathcal{C}_\mathcal{N}}$) for programs, whose actions are of form

$$\overline{d} \leq \overline{f}(\overline{x}) \rightarrow v_1, \ldots, v_n := e_1(\overline{x}), \ldots, e_n(\overline{x})$$

where all occurring functions are monotone. Examples of monotone functions are addition, multiplication, exponentiation, maximum, minimum, and addition/subtraction of a constant.

A particular case of this class of programs are Petri nets, for which it follows that the coverability problem is decidable. But the class is much wider since it allows statements of the form

$$x_1 := x_1 + x_2$$

or even

$$x_1 \geq 2 \;\wedge\; x_2 \geq 5 \quad \longrightarrow \quad x_1 := x_1 * x_2 - 4$$

9.2 Sequences

Let us now consider programs that employ as domain the set $\mathcal{D}^*$ of sequences of elements in some domain $\mathcal{D}$ with a well quasi-ordering $\preceq$. In Section 7 we showed how a well quasi-ordering $\preceq^*$ on $\mathcal{D}^*$ can be generated in a natural way form $\preceq$.

As a concrete application, let $\mathcal{D}$ be a finite set with $\preceq$ taken to be the identity relation. Assume that a program contains one variable v with $\mathcal{D}^*$ as domain.

We consider a constraint system $\mathcal{C}_\mathcal{S}$, where a constraint in $\mathcal{C}_\mathcal{S}$ is of the form ϕ_x with the interpretation $[\![\phi_x]\!] = \{y \ : \ x \preceq^* y\}$. Notice that $\mathcal{C}_\mathcal{S}$ is well quasi-ordered and disjunctive.

We can then define actions of the following forms. We let m be a parameter which ranges over $\mathcal{D}$, and x, y be parameters which range over $\mathcal{D}^*$.

- *Send*(m): $v := v \bullet m$
  This action adds an element m at the end of v.
- *Recieve*(m): $x \bullet m \bullet y \preceq v \to v := y$
  This action is enabled if v contains the element m, in which case it removes all elements in v before and including the element m, and also (nondeterministically) some of the elements after the occurrence of m.
- *Lose*: $x \preceq v \to v := x$.
  This operation arbitrarily loses some of the elements in v.

The above three types of operations form the basis for so-called *lossy channel systems* studied in [AJ96b]. All three operations can easily be checked to be monotone with respect to $\preceq_{\mathcal{C}_\mathcal{S}}$.

10 Binary Constraint Systems

In this section, we will investigate another form of constraints, called *binary constraints*, in which two program variables are compared. Recall that a unary constraint essentially relates the value of a program variable to a constant in its domain. Our intention is that a binary constraint should be a predicate which compares the values of two variables in some way.

Assume a program P. Let *dist* be a binary function whose range $\mathcal{D}$ is equipped with a decidable preorder $\preceq$. A *binary constraint* is a formula of form $d \preceq dist(u, u')$, where $d \in \mathcal{D}$ and where u and u' are program variables or parameters. A *polytope* is a conjunction of binary constraints.

Assume a set of binary constraints over the program variables and parameters of P. Let $\mathcal{C}$ be the set of polytopes. According to Definition 6.3, $\mathcal{C}$ gives rise to a preorder $\preceq_\mathcal{C}$ which for each tuple $\overline{u}$ of variables (each of which is either a program variable of parameter), compares tuples $\overline{d}$ of values of $\overline{u}$ according to $\overline{d} \preceq_{\mathcal{C}} \overline{d}'$ iff $d \preceq dist(d_i, d_j) \implies d \preceq dist(d_i', d_j')$ whenever $d \preceq dist(v_i, v_j)$ is a binary constraint in $\mathcal{C}$. As a special case, the restriction of $\preceq_{\mathcal{C}}$ to values of the tuple $\overline{v}$ of program variables is a constraint system for P.

Let us now give a sufficient criterion for monotonicity of a program with respect to $\preceq_{\mathcal{C}}$.

Proposition 10.1 *Let P be a program. If each action of P is of form*

$$g(\overline{v}, \overline{x}) \longrightarrow \overline{v} := \overline{e}(\overline{v}, \overline{x})$$

where

- *the guard is a polytope in $\mathcal{C}$, and*
- *each e_i is either a program variable or a parameter, i.e., e_i is either of form v_j or of form x_j for some j, and*
- *for each pair $\overline{d}, \overline{d}'$ of values of the program variables $\overline{v}$ with $\overline{d} \preceq_{\mathcal{C}'} \overline{d}'$, and each possible value $\overline{a}$ of the parameters $\overline{x}$, there is a value $\overline{a}'$ of $\overline{x}$ with $\langle \overline{d}, \overline{a} \rangle \preceq_{\mathcal{C}} \langle \overline{d}', \overline{a}' \rangle$,*

then P is monotone.

Proof: The proof follows from Lemma 8.4. □

We shall now see how Proposition 10.1 can be applied to two specific models of infinite-state systems.

10.1 Relational Automata

An example of the use of binary constraints is the class of relational automata, studied by Cerans [Čer94], either with rational numbers or integers as the domain of values. Let us here consider the case of integers.

Let the domain of program variables and parameters be the set $\mathcal{Z}$ of integers. Define a distance function by $dist(v, w) = w - v$. Define $\mathcal{C}$ to be the set of polytopes built from binary constraints of form $d \leq dist(u, u')$ where $d \in \mathcal{N}$ is a nonnegative integer and where u, u' are either program variables or parameters. The ordering on states intuitively makes $\overline{d} \preceq_C \overline{d}'$ be true if the values of all variables appear in the same relative order, and if the distance between adjacent variables is at least the same in $\overline{d}'$ as in $\overline{d}$. This ordering is called "sparser than" in [Čer94].

Let Integral Relational Automata be the class of programs that operate on a set of integer-valued variables, and whose actions are of form $g \to \overline{v} := \overline{e}$ where

- g is a conjunction of statements of form $u + k \leq u'$ where u, u' are either program variables, parameters, or integer constants, and where k is a nonnegative integer,
- Each e_i in $\overline{e}$ is either a program variable, a parameter, or an integer constant.

The class IRA essentially corresponds to integral relational automata, studied by Cerans in [Čer94]. By Proposition 10.1, each program in IRA is monotone. Furthermore, we note that the constraint system is well quasi-ordered, whence the reachablility problem is decidable.

We note that we can use the same presentation to obtain the class of Rational Relational automata, simply by changing the domains of variables, but retaining the constraint system. Proposition 10.1 holds also for this model.

10.2 Programs over Real Valued Clocks

As another application of Proposition 10.1, we will present another class of program that operates on nonnegative real-valued clocks. The class essentially corresponds to the class of timed automata studied in e.g., [ACD90, AH89, Čer92a], with some added capability for random assignment. In this presentation, we omit the finite-state control component for simplicity. It can be added without difficulty.

Let the domain of program variables and parameter be the set $\mathcal{R}^{\geq 0}$ of nonnegative reals. Let K be a program-dependent nonnegative integer which denotes the largest constant that syntactically appears in the program. Intuitively, our constraints will compare clocks to each other and to 0, and record the integer part of the difference if the difference is at most K. Differences larger than K need not be distinguished from K. Define a distance function on $\mathcal{R}^{\geq 0}$ by

$$dist(v,w) = \begin{array}{l} \text{if } v > K \text{ then } -\infty \\ \qquad\qquad \text{else if } w > K \text{ then } K - v \\ \qquad\qquad\qquad\qquad \text{else } w - v \end{array}$$

Intuitively, $dist(v,w)$ records the difference $w - v$ if both v and w are between 0 and K. If v or w is larger than K, then we make sure that it does not matter how much larger than K this argument actually is.

Consider a set of variables and parameters. We shall actually consider two constraint systems.

A *binary constraint* is either of form

1. $d \leq dist(v,v')$ where v,v' are program variables and d is an integer with $-K \leq d \leq K$, or
2. $u \leq d$ where u is a variable (either a program variable or a parameter), and d is an integer with $0 \leq d \leq K$, or
3. $d \leq u$ where u is a variable (either a program variable or a parameter), and d is an integer with $0 \leq d \leq K$.

A binary constraint is *downward closed* if it is of one of the first two types. A polytope is *downward closed* if it is the conjunction of downward closed binary constraints. To motivate our distance function, we note that we can write all constraints as binary constraints in the following way: $u \leq d$ is equivalent to $-d \leq dist(u,0)$, and $d \leq u$ is equivalent to $d \leq dist(0,u)$.

Let $\mathcal{C}$ be the set of polytopes, and let let $\mathcal{C}'$ be the set of downward closed polytopes. For a tuple $\overline{v} = \langle v_1, \ldots, v_n \rangle$ of program variables and a parameter δ with values in $\mathcal{R}^{\geq 0}$, define define $\overline{v} + \delta$ as $\langle v_1 + \delta, \ldots, v_n + \delta \rangle$.

Consider the class CP (for clock programs), whose actions are of form $g(\overline{v} + \delta, \overline{x}) \rightarrow \overline{v} := \overline{e}(\overline{v} + \delta, \overline{x})$, where

- $g(\overline{v}, \overline{x})$ is a polytope, and

- Each e_i in $\overline{e}(\overline{v} + \delta, \overline{x})$ is either of form $v_j + \delta$ (a program variable with an addition of δ), a parameter, or an integer constant.

We are going to show that the program is monotone with respect to the ordering $\preceq_{\mathcal{C}'}$ on program states. However, we cannot use Proposition 10.1 directly, due to the use of addition in the actions. We therefore separate each action into two parts: first an action which allows "time to pass" by the amount δ, and then an action which performs the assignment of possibly new values to the variables. It turns out that we have to use the stronger preorder $\preceq_{\mathcal{C}}$ to make the second part monotone. The structure of the argument is therefore the following.

1. We first show that the action $g(\overline{v} + \delta, \overline{x}) \to \overline{v} := \overline{v} + \delta$ is monotone from $\preceq_{\mathcal{C}'}$ to $\preceq_{\mathcal{C}}$. This means that for any states $\overline{d}_1$, $\overline{d}_2$, and $\overline{d}_3$, with $\overline{d}_1 \preceq_{\mathcal{C}'} \overline{d}_2$ and $\overline{d}_1 \longrightarrow \overline{d}_3$, there exists a state $\overline{d}_4$ such that $\overline{d}_3 \preceq_{\mathcal{C}} \overline{d}_4$ and $\overline{d}_2 \longrightarrow \overline{d}_4$, where we consider transitions derived from this action.
2. By Proposition 10.1, we see that the action $g(\overline{v}, \overline{x}) \to \overline{v} := \overline{e}(\overline{v}, \overline{x})$ is monotone with respect to $\preceq_{\mathcal{C}}$

It follows that the action $g(\overline{v} + \delta, \overline{x}) \to \overline{v} := \overline{e}(\overline{v} + \delta, \overline{x})$ is monotone from $\preceq_{\mathcal{C}'}$ to $\preceq_{\mathcal{C}}$. Since the preorder $\preceq_{\mathcal{C}}$ is stronger than $\preceq_{\mathcal{C}'}$, it follows that $g(\overline{v}+\delta, \overline{x}) \to \overline{v} := \overline{e}(\overline{v}+\delta, \overline{x})$ is monotone with respect to $\preceq_{\mathcal{C}'}$. Since $\mathcal{C}'$ is well quasi-ordered, we can use this constraint system to verify reachability for clock programs.

11 Conclusion

In this paper, we have considered symbolic reachability analysis for infinite-state systems, and presented conditions under which this analysis is guaranteed to terminate. From these general conditions, we have derived several computation models in which the reachability problem is decidable. These models have state variables that range over many different domains, such as natural numbers, multisets, unbounded finite sequences, integers, and real numbers. Systems for which reachability can be analyzed using our methods include Petri Nets [Jan90, JM95], Lossy Channel Systems [AJ96a], relational automata [Čer94], timed automata [ACD90], and unbounded networks of process with clock variables [AJ98]. The general conditions have been presented in our earlier work [AČJYK96], and in this paper we have shown in more detail how to find new computation models that satisfy the conditions. We hope to have conveyed the conclusion that the conditions are very natural, and cover many situations which are still to be investigated.

Acknowledgments

This work ows much to our Karlis Čerāns and Yih-Kuen Tsay, who collaborated in developing the general framework of [AČJYK96]. Thanks also to Tatiana Rybina for contributions to the modeling and analysis of the Futurebus protocol.

References

[ACD90] R. Alur, C. Courcoubetis, and D. Dill. Model-checking for real-time systems. In *Proc. 5^{th} IEEE Int. Symp. on Logic in Computer Science*, pages 414–425, Philadelphia, 1990.

[AČJYK96] Parosh Aziz Abdulla, Karlis Čerāns, Bengt Jonsson, and Tsay Yih-Kuen. General decidability theorems for infinite-state systems. In *Proc. 11^{th} IEEE Int. Symp. on Logic in Computer Science*, pages 313–321, 1996.

[AH89] R. Alur and T. Henzinger. A really temporal logic. In *Proc. 30^{th} Annual Symp. Foundations of Computer Science*, pages 164–169, 1989.

[AJ93] Parosh Aziz Abdulla and Bengt Jonsson. Verifying programs with unreliable channels. In *Proc. 8^{th} IEEE Int. Symp. on Logic in Computer Science*, pages 160–170, 1993.

[AJ96a] Parosh Aziz Abdulla and Bengt Jonsson. Undecidable verification problems for programs with unreliable channels. *Information and Computation*, 130(1):71–90, 1996.

[AJ96b] Parosh Aziz Abdulla and Bengt Jonsson. Verifying programs with unreliable channels. *Information and Computation*, 127(2):91–101, 1996.

[AJ98] Parosh Aziz Abdulla and Bengt Jonsson. Verifying networks of timed processes. In Bernhard Steffen, editor, *Proc. TACAS '98, 7^{th} Int. Conf. on Tools and Algorithms for the Construction and Analysis of Systems*, volume 1384 of *Lecture Notes in Computer Science*, pages 298–312, 1998.

[AK95] Parosh Aziz Abdulla and Mats Kindahl. Decidability of simulation and bisimulation between lossy channel systems and finite state systems. In Lee and Smolka, editors, *Proc. CONCUR '95, 6^{th} Int. Conf. on Concurrency Theory*, volume 962 of *Lecture Notes in Computer Science*, pages 333 – 347. Springer Verlag, 1995.

[BBK77] J. M. Barzdin, J. J. Bicevskis, and A. A. Kalninsh. Automatic construction of complete sample systems for program testing. In *IFIP Congress, 1977*, 1977.

[BCM+90] J.R. Burch, E.M. Clarke, K.L. McMillan, D.L. Dill, and L.J. Hwang. Symbolic model checking: 10^{20} states and beyond. In *Proc. 5^{th} IEEE Int. Symp. on Logic in Computer Science*, 1990.

[BKS88] R. J. R. Back and R. Kurki-Suonio. Distributed cooperation with action systems. *ACM Trans. on Programming Languages and Systems*, 10(4):513–554, Oct. 1988.

[Bry86] R.E. Bryant. Graph-based algorithms for boolean function manipulation. *IEEE Trans. on Computers*, C-35(8):677–691, Aug. 1986.

[BS92] J. Bradfield and C. Stirling. Local model checking for infinite state spaces. *Theoretical Computer Science*, 96:157–174, 1992.

[BS95] O. Burkart and B. Steffen. Composition, decomposition, and model checking of pushdown processes. *Nordic Journal of Computing*, 2(2):89–125, 1995.

[Čer92a] K. Čerāns. Decidability of bisimulation equivalence for parallel timer processes. In *Proc. Workshop on Computer Aided Verification*, volume 663 of *Lecture Notes in Computer Science*, pages 302–315, 1992.

[Čer92b] K. Čerāns. Feasibility of finite and infinite paths in data dependent programs. In *LFCS'92*, volume 620 of *Lecture Notes in Computer Science*, pages 69–80, 1992.

[Čer94] K. Čerāns. Deciding properties of integral relational automata. In Abiteboul and Shamir, editors, *Proc. ICALP '94*, volume 820 of *Lecture Notes in Computer Science*, pages 35–46. Springer Verlag, 1994.

[CM88] K.M. Chandy and J. Misra. *Parallel Program Design: A Foundation.* Addison-Wesley, 1988.

[Dic13] L. E. Dickson. Finiteness of the odd perfect and primitive abundant numbers with n distinct prime factors. *Amer. J. Math.*, 35:413–422, 1913.

[Fin90] A. Finkel. Reduction and covering of infinite reachability trees. *Information and Computation*, 89:144–179, 1990.

[Fin94] A. Finkel. Decidability of the termination problem for completely specified protocols. *Distributed Computing*, 7(3), 1994.

[GS92] S. M. German and A. P. Sistla. Reasoning about systems with many processes. *Journal of the ACM*, 39(3):675–735, 1992.

[GW93] P. Godefroid and P. Wolper. Using partial orders for the efficient verification of deadlock freedom and safety properties. *Formal Methods in System Design*, 2(2):149–164, 1993.

[Hen95] T.A. Henzinger. Hybrid automata with finite bisimulations. In *Proc. ICALP '95*, 1995.

[Hig52] G. Higman. Ordering by divisibility in abstract algebras. *Proc. London Math. Soc.*, 2:326–336, 1952.

[Jan90] P. Jančar. Decidability of a temporal logic problem for Petri nets. *Theoretical Computer Science*, 74:71–93, 1990.

[JM95] P. Jančar and F. Moller. Checking regular properties of Petri nets. In *Proc. CONCUR '95, 6^{th} Int. Conf. on Concurrency Theory*, pages 348–362, 1995.

[JP93] B. Jonsson and J. Parrow. Deciding bisimulation equivalences for a class of non-finite-state programs. *Information and Computation*, 107(2):272–302, Dec. 1993.

[KMM$^+$97] Y. Kesten, O. Maler, M. Marcus, A. Pnueli, and E. Shahar. Symbolic model checking with rich assertional languages. In O. Grumberg, editor, *Proc. 9^{th} Int. Conf. on Computer Aided Verification*, volume 1254, pages 424–435, Haifa, Israel, 1997. Springer Verlag.

[Pel96] D. Peled. Combining partial order reductions with on-the-fly model checking. *Formal Aspects of Computing*, 8:39–64, 1996.

[Rus97] J.M. Rushby. Specification, proof checking, and model checking for protocols and distributed systems with PVS. Tutorial notes for FORTE/PSTV '97, Nov. 1997.

[Sha96] N. Shankar. PVS: Combining specification, proof checking, and model checking. In *Proc. FMCAD '96*, volume 1166 of *Lecture Notes in Computer Science*, 1996.

[Sti96] C. Stirling. Decidability of bisimulation equivalence for normed pushdown processes. In *Proc. CONCUR '96, 7^{th} Int. Conf. on Concurrency Theory*, volume 1119 of *Lecture Notes in Computer Science*, pages 217–232. Springer Verlag, 1996.

[Val93] A. Valmari. On-the-fly verification with stubborn sets. In Courcoubetis, editor, *Proc. 5^{th} Int. Conf. on Computer Aided Verification*, number 697 in Lecture Notes in Computer Science, pages 59–70, 1993.

[VW86] M. Y. Vardi and P. Wolper. An automata-theoretic approach to automatic program verification. In *Proc. 1^{st} IEEE Int. Symp. on Logic in Computer Science*, pages 332–344, June 1986.

[Wol86] Pierre Wolper. Expressing interesting properties of programs in propositional temporal logic (extended abstract). In *Proc. 13^{th} ACM Symp. on Principles of Programming Languages*, pages 184–193, Jan. 1986.