Reachability Analysis of Pushdown Automata (extended abstract)

Ahmed Bouajjani   Oded Maler

VERIMAG
Miniparc-Zirst, Rue Lavoisier, 38330 Montbonnot St-Martin, France.
Email: Ahmed.Bouajjani@imag.fr, Oded.Maler@imag.fr

1 Introduction

Systems are commonly modeled by various types of transition systems, including finite automata, pushdown automata, Petri nets, timed or hybrid automata, etc. In this framework, most of the system analysis problems (model-checking, synthesis) reduce to (various kinds of) "*reachability problems*" on these models. Thus, a fundamental ingredient for system analysis algorithms is calculating the set of *all predecessors* of a given set of states, say S, i.e., calculating the set of states from which it is possible to reach S. This set is the smallest solution of the equation $X = S \cup pre(X)$ where pre is the function which gives, for any set of states, the set of its *immediate predecessors*. In other words, the set of predecessors of a set S is its image by the reflexive-transitive closure of the function pre, denoted by pre^*. By definition, $pre^*(S)$ is the limit of the *infinite* increasing sequence $\{X_i\}_{i \geq 0}$ such that $X_0 = S$ and $X_{i+1} = X_i \cup pre(X_i)$ for every $i \geq 0$. Then, the problem is how to calculate effectively $pre^*(S)$ and obtain a *finite* representation of this set. For finite-state systems this can be done by straightforward graph algorithms. In the case of infinite-state systems, the main problem to address is, given a class of systems, find a class of structures having the following properties:

1. It provides for finite representations of (certain) infinite sets of states
2. It is closed under boolean operations
3. Its emptiness problem is decidable
4. It is closed under the pre function

Then, we have to show that this class of structures is also closed under the pre^* function.

Closure under boolean operations and decidability of the emptiness problem imply that the inclusion and the membership problems are decidable. The decidability of the inclusion problem is useful for comparing the different X_i's and detect that the limit has been reached (if $X_i = X_{i+1}$), whereas the decidabililty of the membership problem allows to check that a given state (an initial state for instance) is in the set $pre^*(S)$.

Several instances of systems and corresponding representation structures have been considered in the litterature. For example, in the case of timed automata, special kinds of polyhedra (regions) are used to represent infinite sets of states (vectors of reals corresponding to clock valuations) [AD94]. Polyhedra are also used for linear hybrid systems. However, in this case, there is no algorithm for calculating a finite representation of the *exact* set of predecessors (the reachability problem is undecidable), but upper approximations of this set can be calculated [ACH$^+$95]. In [BG96], representation structures called QDD's are introduced for fifo-channel systems. These structures are finite-state automata representing sets of queue contents. But, as in the case of linear hybrid systems, the procedure for calculating the set of predecessors for these structures is not guaranteed to terminate. Finally, notice that the idea of using representations of sets of states is also used in the finite-state case to overcome the state-explosion problem [McM93]. The usual representation structures in this case are BDD's (binary decision diagrams) [Bry92].

In this work, we consider the case of pushdown systems, and actually, even a more general class of systems: *alternating pushdown systems*, i.e., systems with both existential and universal nondeterminism (see [Var95] for a survey on alternating automata). This general setting allows to reason in a uniform way about *analysis* problems where existential and universal path quantification must be considered, for instance in model-checking for branching-time temporal logics (see Section 5.2) as well as about *synthesis* problems, for instance finding winning strategies for 2-player games (see [AMP95]).

A state (we use rather the word "configuration") of a pushdown system $\mathcal{P}$ is a pair $\langle p, w \rangle \in P \times \Gamma^*$ where P is a finite set of control locations and Γ is a finite stack alphabet (w is the stack contents). Let $P = \{p^1, \ldots, p^m\}$. We can write any set of configurations (subset of $P \times \Gamma^*$) as $\{p^1\} \times L_1 \cup \ldots \cup \{p^m\} \times L_m$ where the L_i's are subsets of Γ^*. Then, we propose as a representation structure for such sets the *alternating multi-automaton* (AMA) which consists of an alternating finite-state automaton $\mathcal{A}$ with a set $\{s^1, \ldots, s^m\}$ of initial states such that L_i is the language accepted by $\mathcal{A}$ starting from s^i. It is important to remember that $\mathcal{A}$ is just a tool to represent sets of configurations, and not to confuse its behaviour with that of the system $\mathcal{P}$.

Using alternating automata as representation structures for sets of configurations makes the closure under boolean operations straightforward. As for the emptiness problem, it is well known that it is decidable for finite alternating automata. Then, our main result consists in proving that this representation admits a simple and natural procedure for calculating predecessors (we show the closure under the *pre* and pre^* functions). It can be observed that for some pushdown systems and sets of configurations, the sequence of the X_i's does not reach its limit (see Example 4.1). So, we propose an alternative procedure which constructs another sequence of sets of configurations converging to the *same* limit as the sequence of the X_i's, but in a *finite* number of steps.

As an application of our techniques, we show that they lead to comprehensible new model-checking algorithms for pushdown systems in both cases of linear and branching-time properties. We show that we can construct a representation of the set of *all* the configurations satisfying a property which is either an ω-regular property (expressible by a Büchi automaton or by a formula in the linear-time μ-calculus), or described by a formula in CTL [CES83]. It is worth noting that existing model-checking algorithms for pushdown systems (e.g., [BS94]) only solve the problem whether one given configuration satisfies a property.

2 Alternating Pushdown Systems

2.1 Definition

Given a set of atoms A, we denote by $\mathcal{B}_+(A)$ the set of positive boolean formulas over A, i.e., the set of formulas containing A and closed under the boolean connectives $\vee$ and $\wedge$.

Definition 2.1 *An alternating pushdown system (APDS for short) is a triplet $\mathcal{P} = (P, \Gamma, \Delta)$ where P is a finite set of control locations, Γ is a finite stack alphabet, and $\Delta : P \times \Gamma \rightarrow \mathcal{B}_+(P \times \Gamma^*)$ is a transition table.*

Notice that APDS's have no input alphabet. We do not use them as language acceptors but are rather interested in the behaviours (runs) they generate.

A *configuration* of $\mathcal{P}$ is a pair $\langle p, w \rangle$ where $p \in P$ is a control location and $w \in \Gamma^*$ is a stack contents. Let $\mathcal{C} = P \times \Gamma^*$ be the set of configurations of $\mathcal{P}$.

In order to define the *runs* of $\mathcal{P}$, let us suppose without loss of generality that for every $(p, \gamma) \in P \times \Gamma$, the expression $\Delta(p, \gamma)$ is in disjunctive normal form (DNF for short). Roughly speaking, the transition table Δ allows to develop, starting from any given configuration c, an OR/AND-tree whose nodes are labelled by configurations. The runs of $\mathcal{P}$ are the AND-subtrees obtained by choosing at each OR-branching only one successor, and taking all the successors at each AND-branching: Let $p \in P$ and $\gamma \in \Gamma$ with $\Delta(p, \gamma) = \bigvee_{i=1}^{n} \bigwedge_{j=1}^{m_i} \langle p_{i_j}, w_{i_j} \rangle$. Then, while being at control location p and observing the symbol γ at the *top* of the stack, the automaton chooses nondeterministically one of the disjuncts in the expression $\Delta(p, \gamma)$, say $\bigwedge_{j=1}^{m_k} \langle p_{k_j}, w_{k_j} \rangle$, and then forks into m_k copies, each in a configuration such that the control location is p_{k_j} and the stack contents is obtained by replacing γ (the top of the previous stack) with the sequence w_{k_j}.

Two special cases of APDS's can be considered, the $\exists$PDS's (or simply PDS's) and the $\forall$PDS's corresponding to APDS's where respectively only disjunctions or conjunctions appear in the definition of Δ.

2.2 Reachability

We introduce hereafter a reachability relation $\Rightarrow \subseteq \mathcal{C} \times 2^{\mathcal{C}}$ between configurations and sets of configurations. Informally, $c \Rightarrow C$ if and only if C is a finite frontier (finite maximal set of incomparable nodes) of a run tree of $\mathcal{P}$ starting from c. Note that in the case of an $\exists$PDS where a run is a sequence, each reachable frontier contains exactly one configuration. To give the formal definition of $\Rightarrow$ we need to introduce the notion of *transition rules*. Let

$p \in P$, $\gamma \in \Gamma$, and $\Delta(p,\gamma) = \bigvee_{i=1}^{n} \bigwedge_{j=1}^{m_i} \langle p_{i_j}, w_{i_j} \rangle$. Then, for every $i \in \{1,\ldots,n\}$, the pair $(\langle p,\gamma\rangle, \{\langle p_{i_1}, w_{i_1}\rangle, \ldots, \langle p_{i_{m_i}}, w_{i_{m_i}}\rangle\})$ is a transition rule of $\mathcal{P}$. Let $\mathcal{R}$ be the set of such rules. Then, the reachability relation $\Rightarrow$ is the smallest subset of $\mathcal{C} \times 2^{\mathcal{C}}$ such that:

1. $\forall c \in \mathcal{C}.\ c \Rightarrow \{c\}$,

2. if $(c, \{c_1, \ldots, c_n\}) \in \mathcal{R}$, then $c \Rightarrow \{c_1, \ldots, c_n\}$,

3. if $\langle p, u\rangle \Rightarrow \{\langle p_1, w_1\rangle, \ldots, \langle p_n, w_n\rangle\}$, then $\forall v \in \Gamma^*.\ \langle p, uv\rangle \Rightarrow \{\langle p_1, w_1 v\rangle, \ldots, \langle p_n, w_n v\rangle\}$,

4. if $\langle p, u\rangle \Rightarrow \{\langle p_1, w_1\rangle, \ldots, \langle p_n, w_n\rangle\}$ and $\forall i \in \{1,\ldots,n\}.\ \langle p_i, w_i\rangle \Rightarrow C_i$, then $\langle p, u\rangle \Rightarrow \bigcup_{i=1}^{n} C_i$.

2.3 Predecessors

We introduce the function $pre_{\mathcal{P}} : 2^{\mathcal{C}} \to 2^{\mathcal{C}}$ which associates with each set of configurations $C \subseteq \mathcal{C}$ the set of configurations that are its *immediate predecessors*. A configuration is an immediate predecessor of a set of configurations C if there is a transition rule r such that *all* the successors of c by this rule are in C. Formally, given a transition rule $r = (\langle p,\gamma\rangle, \{\langle p_1, w_1\rangle, \ldots, \langle p_n, w_n\rangle\}) \in \mathcal{R}$, let $pre^r_{\mathcal{P}}(C) = \{\langle p, \gamma u\rangle \in \mathcal{C} \ :\ \forall i \in \{1,\ldots,n\}.\ \langle p_i, w_i u\rangle \in C\}$. Then, $pre_{\mathcal{P}}(C) = \bigcup_{r \in \mathcal{R}} pre^r_{\mathcal{P}}(C)$.

We denote the reflexive-transitive closure of the function $pre_{\mathcal{P}}$ by $pre^*_{\mathcal{P}}$. It is easy to see that for every $C \subseteq \mathcal{C}$, $pre^*_{\mathcal{P}}(C) = \{c \in \mathcal{C} \ :\ \exists C' \subseteq C.\ c \Rightarrow C'\}$. Finally, let $pre^+_{\mathcal{P}} = pre_{\mathcal{P}} \circ pre^*_{\mathcal{P}}$.

We will omit the subscript $\mathcal{P}$ and write simply pre, pre^r, pre^*, and pre^+ when it is clear from the context which system is under consideration.

3 Alternating Multi-Automata

3.1 Definition

Definition 3.1 *Let $\mathcal{P} = (P, \Gamma, \Delta)$ be an APDS and let $P = \{p^1, \ldots p^m\}$. An alternating (finite-state) (P,Γ)-multi-automaton ($\mathcal{P}$-AMA for short) is a tuple $\mathcal{A} = (\Gamma, Q, \delta, I, F)$ where Q is a finite set of states, $\delta : Q \times \Gamma \to \mathcal{B}_+(Q)$ is the transition table, $I = \{s^1, \ldots s^m\} \subseteq Q$ is a set of initial states and $F \subseteq Q$ is a set of final states.*

If only disjunctions (resp. conjunctions) appear in the definition of the transition table δ, then we say that $\mathcal{A}$ is a $\mathcal{P}$-$\exists$MA (resp. $\mathcal{P}$-$\forall$MA).

The reflexive-transitive closure of δ is the function $\delta^* : Q \times \Gamma^* \to \mathcal{B}_+(Q)$ defined inductively by:

- $\delta^*(q, \varepsilon) = q$,

- $\delta^*(q, \gamma w) = \delta(q,\gamma)[q' \leftarrow \delta^*(q', w)]_{q' \in Q}$.

Using δ^*, we can define the language of $\mathcal{A}$ starting from a given state. For that, let $\mathcal{I}$ be the interpretation function of boolean formulas in $\mathcal{B}_+(Q)$ defined as usual by considering that for every atom (state) q, $\mathcal{I}(q) = true$ iff $q \in F$. Then, for every state $q \in Q$, the *language* accepted by $\mathcal{A}$ starting from q is $L(\mathcal{A}, q) = \{w \in \Gamma^* \ : \ \mathcal{I}(\delta^*(q, w)) = true\}$. The notion of accepted language can be generalized straightforwardly to expressions in $\mathcal{B}_+(Q)$: Given $e_1, e_2 \in \mathcal{B}_+(Q)$, $L(\mathcal{A}, e_1 \vee e_2) = L(\mathcal{A}, e_1) \cup L(\mathcal{A}, e_2)$, and $L(\mathcal{A}, e_1 \wedge e_2) = L(\mathcal{A}, e_1) \cap L(\mathcal{A}, e_2)$.

The set of configurations *recognized* by $\mathcal{A}$ is $Conf(\mathcal{A}) = \bigcup_{i=1}^{m} \{p_i\} \times L(\mathcal{A}, s^i)$. We say that a set of configurations $C \subseteq \mathcal{C}$ is *regular* iff it is recognized by some $\mathcal{P}$-AMA (i.e., there exists a $\mathcal{P}$-AMA $\mathcal{A}$ such that $Conf(\mathcal{A}) = C$).

3.2 Runs

We need in the next section to reason about *runs* of $\mathcal{P}$-AMA's and to relate them with runs in their corresponding APDS $\mathcal{P}$ (see Lemma 6.3). To define the runs of $\mathcal{A}$, we suppose that for every $(q, \gamma) \in Q \times \Gamma$, the expression $\delta(q, \gamma)$ is in DNF. Then, given a finite sequence $w \in \Gamma^*$ and a state $q \in Q$, a run of $\mathcal{A}$ over w starting from q is a finite tree whose nodes are labelled by states in Q and edges are labelled by symbols in Γ such that the root is labelled by q, and the labelling of the other nodes is consistent with δ (it corresponds to an AND-tree extracted from the OR/AND-tree induced by δ). Notice that in such a tree, each sequence of edges going from the root to the leaves is labelled by w, and hence, all the edges starting at the same level of the tree have the same label, and all the leaves of the tree are at the same height. Let us denote by $Run(q, w)$ the set of runs starting from q over the word w. Given a run $\rho \in Run(q, w)$, we denote by $Leaves(\rho)$ the set of leaves of ρ.

As in the case of APDA, we introduce a notion of *transition rules*. Let $q \in Q$, $\gamma \in \Gamma$, and $\delta(q, \gamma) = \bigvee_{i=1}^{n} \bigwedge_{j=1}^{m_i} q_{i_j}$. Then, for every $i \in \{1, \ldots, n\}$, the triplet $(q, \gamma, \{q_{i_1}, \ldots, q_{i_{m_i}}\})$ is a *transition rule* of $\mathcal{A}$. Let $\mathcal{T}$ be the set of these transition rules. We introduce a transition relation $\rightarrow \subseteq Q \times \Gamma^* \times 2^Q$ which is in some sense the reflexive-transitive closure of $\mathcal{T}$. This relation is inductively defined by:

1. $\forall q \in Q.\ q \xrightarrow{\varepsilon} \{q\}$,
2. if $(q, \gamma, \{q_1, \ldots, q_n\}) \in \mathcal{T}$, then $q \xrightarrow{\gamma} \{q_1, \ldots, q_n\}$,
3. if $q \xrightarrow{u} \{q_1, \ldots, q_n\}$ and $\forall i \in \{1, \ldots, n\}.\ q_i \xrightarrow{v} Q_i$, then $q \xrightarrow{uv} \bigcup_{i=1}^{n} Q_i$.

Clearly, for every $q \in Q$ and every $Q' \subseteq Q$, we have $q \xrightarrow{w} Q'$ iff there exists a run $\rho \in Run(q, w)$ such that $Leaves(\rho) = Q'$. It is also easy to see that, for every $q \in Q$, we have $L(\mathcal{A}, q) = \{w \in \Gamma^* \ : \ \exists Q' \subseteq F.\ q \xrightarrow{w} Q'\}$.

Given a run $\rho \in Run(q, w)$, we write $\rho : q' \xrightarrow{u} Q'$ if there exists a subrun ρ' of ρ over u which starts from q' and such that $Leaves(\rho') = Q'$. Notice that u must be a factor of w, i.e., $w = u_1 u u_2$ for some words u_1 and u_2.

3.3 Closure under boolean operations

Complementation: Let $\mathcal{A} = (\Gamma, Q, \delta, I, F)$ be a $\mathcal{P}$-AMA. We construct a $\mathcal{P}$-AMA $\overline{\mathcal{A}}$ recognizing $\mathcal{C} \setminus Conf(\mathcal{A})$. The automaton $\overline{\mathcal{A}}$ is defined by $(\Gamma, Q, \widetilde{\delta}, I, Q \setminus F)$ where $\widetilde{\delta}$ is the dual function of δ, i.e., for every $q \in Q$ and every $\gamma \in \Gamma$, $\widetilde{\delta}(q,\gamma)$ is obtained by replacing in $\delta(q,\gamma)$ disjunctions by conjunctions and vice-versa.

Union and Intersection: Let $\mathcal{A}_1 = (\Gamma, Q_1, \delta_1, I_1, F_1)$ and $\mathcal{A}_2 = (\Gamma, Q_2, \delta_2, I_2, F_2)$ be two $\mathcal{P}$-AMA's. We construct a $\mathcal{P}$-AMA $\mathcal{A}_{\#}$ recognizing $Conf(\mathcal{A}_1) \# Conf(\mathcal{A}_2)$, for $\# \in \{\cup, \cap\}$. Let $\mathcal{A}_{\#} = (\Gamma, \{t^1, \ldots, t^m\} \cup Q_1 \cup Q_2, \delta_{\#}, \{t^1, \ldots, t^m\}, F_1 \cup F_2)$, where $\delta_{\#}$ is given by: $\forall \gamma \in \Gamma$,

- $\forall i \in \{1,2\}.\ \forall q \in Q_i.\ \delta_{\#}(q,\gamma) = \delta_i(q,\gamma)$,
- $\forall j \in \{1, \ldots, m\}.\ \delta_{\cup}(t^j,\gamma) = \delta_1(s_1^j,\gamma) \vee \delta_2(s_2^j,\gamma)$ and $\delta_{\cap}(t^j,\gamma) = \delta_1(s_1^j,\gamma) \wedge \delta_2(s_2^j,\gamma)$.

3.4 Expressiveness and decidability

It is well known that alternating finite-state automata can be translated into ($\exists$-)finite-state automata. Hence, the emptiness problem of $\mathcal{P}$-AMA's, as well as their membership and inclusion problems, are decidable.

4 Calculating Predecessors

4.1 Calculating *pre*

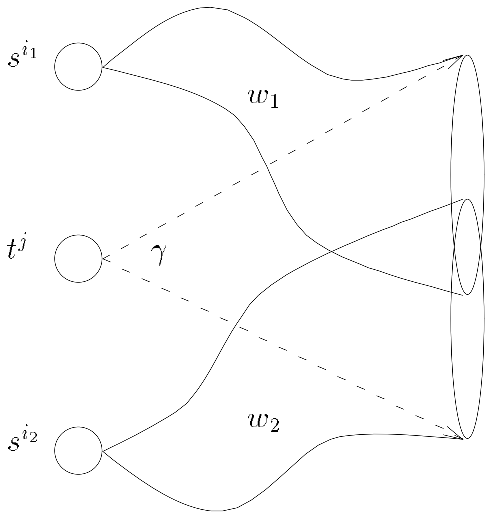
Given a $\mathcal{P}$-AMA $\mathcal{A}$, we define hereafter a $\mathcal{P}$-AMA $\mathcal{A}_{pre}$ recognizing $pre(Conf(\mathcal{A}))$. Since $pre(Conf(\mathcal{A})) = \bigcup_{r \in \mathcal{R}} pre^r(Conf(\mathcal{A}))$ and $\mathcal{P}$-AMA's are closed under union, it suffices to give, for every transition rule $r \in \mathcal{R}$, an automaton $\mathcal{A}_{pre}^r = (\Gamma, Q', \delta', I', F')$ such that $Conf(\mathcal{A}_{pre}^r) = pre^r(Conf(\mathcal{A}))$.

Let $r = (\langle p^j, \gamma \rangle, \{\langle p^{i_1}, w_1 \rangle, \ldots, \langle p^{i_n}, w_n \rangle\}) \in \mathcal{R}$. Then, the automaton $\mathcal{A}_{pre}^r$ is given by:

- $Q' = Q \cup \{t^1, \ldots, t^m\}$,
- $\delta' : Q' \times \Gamma \rightarrow \mathcal{B}_+(Q')$ is defined as follows:
  - $\forall \alpha \in \Gamma.\ \forall q \in Q.\ \delta'(q,\alpha) = \delta(q,\alpha)$,
  - $\delta'(t^j,\gamma) = \bigwedge_{k=1}^{n} \delta^*(s^{i_k}, w_k)$,
  - $\forall \alpha \in \Gamma.\ \forall i \in \{1, \ldots, m\}.\ i \neq j$ or $\alpha \neq \gamma.\ \delta'(t^i,\alpha) = false$,
- $I' = \{t^1, \ldots, t^m\}$,
- $F' = F$.

Let us justify briefly this construction. Immediate predecessors of $Conf(\mathcal{A})$ by the rule r must be of the form $\langle p^j, \gamma v\rangle$ such that $\forall k \in \{1, \ldots, n\}$, $\langle p^{i_k}, w_k v\rangle \in Conf(\mathcal{A})$. This means that, (i) $\forall i \in \{1, \ldots, m\}.\ i \neq j.\ L(\mathcal{A}^r_{pre}, t^i) = \emptyset$, (ii) the first symbol (the top of the stack) in every word accepted by $\mathcal{A}^r_{pre}$ starting from t^j must be γ, and (iii) $\forall v \in \Gamma^*.\ \gamma v \in L(\mathcal{A}^r_{pre}, t^j)$ iff $\forall k \in \{1, \ldots, n\}.\ w_k v \in L(\mathcal{A}, s^{i_k})$, which is equivalent to say that $v \in \bigcap_{k=1}^n L(\mathcal{A}, \delta^*(s^{i_k}, w_k))$.

Intuitively, the construction of $\mathcal{A}^r_{pre}$ consists in extending $\mathcal{A}$ by adding m new initial states $\{t^1, \ldots, t^m\}$ and new γ-transitions each of them is starting from the state t^j and going to all the states appearing in a same conjunction in the disjunctive normal form of the expression $\bigwedge_{k=1}^n \delta^*(s^{i_k}, w_k)$. The following picture illustrates the construction for $n = 2$. Given a run ρ_{i_1} over w_1 starting from s^{i_1} and a run ρ_{i_2} over w_2 starting from s^{i_2}, we add an AND-transition by γ from t^j to every state in $Leaves(\rho_{i_1}) \cup Leaves(\rho_{i_2})$.



The automaton $\mathcal{A}_{pre}$ is obtained by considering the union of all the automata $\mathcal{A}^r_{pre}$ (we extend $\mathcal{A}$ by the new initial states $\{t^1, \ldots, t^m\}$ and we add all the transitions starting from these states introduced in each $\mathcal{A}^r_{pre}$). Notice that if $\mathcal{A}$ is a $\mathcal{P}$-$\exists$MA, the automaton $\mathcal{A}_{pre}$ is also a $\mathcal{P}$-$\exists$MA (in this case, $\mathcal{A}_{pre}$ is obtained from $\mathcal{A}$ by adding a γ-transition from a new initial state t^i to a state q of $\mathcal{A}$ whenever there is a transition rule $(\langle p^i, \gamma\rangle, \{\langle p^k, w\rangle\}) \in \mathcal{R}$ such that $s^k \xrightarrow{w} \{q\}$).

Proposition 4.1 *For every APDS $\mathcal{P}$, and for every regular set of configurations C of $\mathcal{P}$, we can effectively construct a $\mathcal{P}$-AMA recognizing $pre(C)$.*

4.2 Calculating pre^*

We show in this section that given a regular set of configurations C represented by a $\mathcal{P}$-AMA $\mathcal{A}$, we can construct $\mathcal{P}$-AMA $\mathcal{A}_{pre^*}$ such that $Conf(\mathcal{A}_{pre^*}) = pre^*(C)$. By definition, $pre^*(C) = \bigcup_{i\geq 0} X_i$ with $X_0 = C$ and $\forall i \geq 0$, $X_{i+1} = X_i \cup pre(X_i)$. Then, since C is regular and we know how to calculate unions and immediate predecessors of regular sets of configurations, a way to calculate $pre^*(C)$ is to construct iteratively the increasing sequence $X_0, X_1, \ldots$. If for some index i, we have $X_{i+1} = X_i$, then it is clear that $X_i = pre^*(C)$. However, the existence of such a fixed point is not guaranteed in general, and we may never reach the limit of the X_i sequence. To overcome this problem, we calculate $pre^*(C)$ differently, as the limit of another increasing sequence of sets of configurations $Y_0, Y_1, \ldots$ for which we can prove that:

1. $\exists i \geq 0.\ Y_{i+1} = Y_i$,
2. $\forall i \geq 0.\ X_i \subseteq Y_i$,
3. $\forall i \geq 0.\ Y_i \subseteq \bigcup_{j \geq 0} X_j$.

The first point ensures the halting of the procedure calculating the sequence of the Y_i's. The second point means that, by calculating the limit of the Y_i's, we capture (at least) the whole set $pre^*(C)$, and the third point means that only elements of $pre^*(C)$ are captured.

Let us present intuitively how the Y_i's are defined. It can be observed that the $\mathcal{P}$-AMA we construct for X_{i+1} from the $\mathcal{P}$-AMA for X_i (using the construction of the previous subsection) has m additional states that are the new initial states. Then, the basic idea behind the definition of the sequence $\{Y_i\}_{i \geq 0}$ is to reuse at each step the *same initial* states in order to keep the number of states *constant* [1]. This guarantees termination but complicates the correctness proof.

Let us define formally the sequence $\{Y_i\}_{i \geq 0}$. We start with a $\mathcal{P}$-AMA $\mathcal{A}$ recognizing the regular set of configurations C. We assume without loss of generality that $\mathcal{A}$ has no transition leading to an initial state (every automaton can be converted to one having this property). Then, we define a sequence of $\mathcal{P}$-AMA's $\mathcal{A}_0, \mathcal{A}_1, \ldots$ such that $\mathcal{A}_0 = \mathcal{A}$ and for every $i \geq 0$, $\mathcal{A}_{i+1}$ is obtained from $\mathcal{A}_i$ by conserving the same states and transitions, and adding new transitions as follows: For every transition rule $r = (\langle p^j, \gamma \rangle, \{\langle p^{i_1}, w_1 \rangle, \ldots, \langle p^{i_n}, w_n \rangle\}) \in \mathcal{R}$, consider the $\mathcal{P}$-AMA $\mathcal{A}^r_{i+1}$ obtained from $\mathcal{A}_i$ by extending its transition table δ_i to the function δ^r_{i+1} defined by:

- $\forall \alpha \in \Gamma.\ \forall q \in Q.\ q \neq s^j$ or $\alpha \neq \gamma.\ \delta^r_{i+1}(q, \alpha) = \delta_i(q, \alpha)$,
- $\delta^r_{i+1}(s^j, \gamma) = \delta_i(s^j, \gamma) \vee \bigwedge_{k=1}^{n} \delta^*_i(s^{i_k}, w_k)$.

Now, the transition table δ_{i+1} of the automaton $\mathcal{A}_{i+1}$ is defined by: $\forall \alpha \in \Gamma.\ \forall q \in Q.\ \delta_{i+1}(q, \alpha) = \bigvee_{r \in \mathcal{R}} \delta^r_{i+1}(q, \alpha)$. Notice that if $\mathcal{A}$ is a $\mathcal{P}$-$\exists$MA, then all the $\mathcal{A}_i$'s are also $\mathcal{P}$-$\exists$MA's. We denote by $\rightarrow_i$ (resp. $\rightarrow^r_i$) for the transition relation of $\mathcal{A}_i$ (resp. $\mathcal{A}^r_i$).

Then, for every $i \geq 0$ and for every $r \in \mathcal{R}$, let $Y^r_i = Conf(\mathcal{A}^r_i)$ and $Y_i = \bigcup_{r \in \mathcal{R}} Y^r_i = Conf(\mathcal{A}_i)$. It is easy to see from the definition that the sequence $\{Y_i\}_{i \geq 0}$ is increasing.

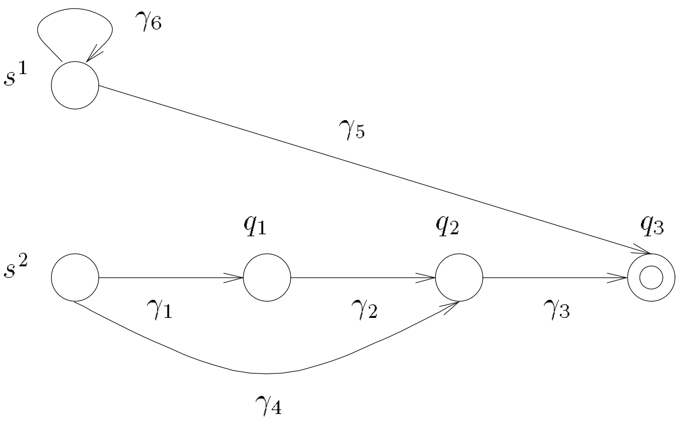
[1] The idea is inspired by the construction given in [BO93], pages 91-93, of a finite-state automaton recognizing the closure of a regular language under the rewriting relation induced by a *monadic string-rewriting system*.

Example 4.1 Let $\mathcal{P}$ be an $\exists$PDA such that $P = \{p^1, p^2\}$, $\Gamma = \{\gamma_1, \ldots, \gamma_6\}$, and Δ corresponds to the three transition rules

$$\begin{aligned} r_1 &= (\langle p^2, \gamma_4\rangle, \{\langle p^2, \gamma_1\gamma_2\rangle\}) \\ r_2 &= (\langle p^1, \gamma_5\rangle, \{\langle p^2, \gamma_4\gamma_3\rangle\}) \\ r_3 &= (\langle p^1, \gamma_6\rangle, \{\langle p^1, \varepsilon\rangle\}) \end{aligned}$$

Then, consider the set of configurations $C = \{\langle p^2, \gamma_1\gamma_2\gamma_3\rangle\}$. It can be represented by a $\mathcal{P}$-$\exists$MA $\mathcal{A}$ such that $Q = \{s^1, s^2, q_1, q_2, q_3\}$, $I = \{s^1, s^2\}$, $F = \{q_3\}$, and $\delta(s^2, \gamma_1) = q_1$, $\delta(q_1, \gamma_2) = q_2$, and $\delta(q_2, \gamma_3) = q_3$.

The picture below shows the automaton $\mathcal{A}_{pre^*}$ obtained by our construction:



This automaton is constructed in the following manner: Initially, we have $\mathcal{A}_0 = \mathcal{A}$. Then, we construct $\mathcal{A}_1$, which is equal to $\mathcal{A}_1^{r_1}$, by adding the transition from s^2 to q_2 by γ_4. The automaton $\mathcal{A}_2$ is equal to $\mathcal{A}_2^{r_2}$ and is obtained by adding the transition from s^1 to q_3 by γ_5. Finally, we construct the automaton $\mathcal{A}_3$, which is equal to $\mathcal{A}_3^{r_3}$, by adding the self-loop from s^1 to itself by γ_6. Then, it can be seen that the construction stops since no transition has to be added. So, we have $\mathcal{A}_{pre^*} = \mathcal{A}_3$, and $pre^*(C) = \{p^1\} \times \gamma_6^*\gamma_5 \cup \{p^2\} \times \{\gamma_1\gamma_2\gamma_3, \gamma_4\gamma_3\}$.

Observe that in this example, we have $X_1 = Y_1$ and $X_2 = Y_2$, but $X_3 \subset Y_3$. Indeed, in the third step of the construction, by adding a self-loop on s^1, Y_3 contains all the configurations of the form $\langle p^1, \gamma_6^k\gamma_5\rangle$ for every $k \geq 1$, whereas only $\langle p^1, \gamma_6\gamma_5\rangle \in X_3$. However, despite the fact that these configurations are not immediate predecessors of X_2 configurations, they are all in $pre^*(C)$ since it can be seen that, for every $k \geq 1$, $\langle p^1, \gamma_6^k\gamma_5\rangle \in X_{k+2}$. Notice that there is no index i such that X_i contains all these configurations (and hence, there is no index i such that $X_i = X_{i+1}$).

The correcteness of the construction of $\mathcal{A}_{pre^*}$ is proved in the appendix. Then, we have the following result:

Theorem 4.1 *For every APDS $\mathcal{P}$, and for every regular set of configurations C of $\mathcal{P}$, we can effectively construct a $\mathcal{P}$-AMA recognizing $pre^*(C)$.*

5 Application to Model-Checking

5.1 Model-Checking for ω-Regular Properties

Let $Prop$ be a finite set of atomic propositions and let $\Sigma = 2^{Prop}$. Given a PDS ($\exists$PDS) $\mathcal{P}$, a labelling function $\Lambda \in [P \to \Sigma]$, and an ω-regular set $\Pi \subseteq \Sigma^\omega$, we want to compute the set of configurations c of $\mathcal{P}$ such that every infinite path (run) starting from c is in Π (via the labelling function Λ). It is easy to see that this problem reduces to the following one which is expressed on Büchi pushdown automata: Given an $\exists$PDS and a subset of its control locations $G \in P$ (a set of repeating locations), compute the set of all configurations c such that there exists an accepting run starting from c (i.e., a run which visits infinitely often configurations with control locations in G). Notice that this problem generalizes the emptiness problem of Büchi pushdown automata (deciding whether there is an accepting run starting from one given initial configuration).

For every $p \in P$, and every $\gamma \in \Gamma$, let $C_p^\gamma = \{p\} \times \gamma\Gamma^*$. Then, the solution of our "generalized emptiness problem" is based on the following fact:

Lemma 5.1 *For every $c \in \mathcal{C}$, there is an accepting run starting from c if and only if there exists $p \in P$ and $\gamma \in \Gamma$ such that $c \in pre^*(C_p^\gamma)$ and $\langle p, \gamma \rangle \in pre^*(pre^+(C_p^\gamma) \cap (G \times \Gamma^*))$.*

It is obvious that the sets of configurations $G \times \Gamma^*$ and C_p^γ are regular. Hence, using Theorem 4.1 and Proposition 4.1, we can construct $\mathcal{P}$-AMA's recognizing the sets $pre^*(C_p^\gamma)$ and $pre^*(pre^+(C_p^\gamma) \cap (G \times \Gamma^*))$. Consequently, using Lemma 5.1, we are able to construct a $\mathcal{P}$-AMA which recognizes the set of all configurations having an accepting run: First, we determine all the pairs $(p, \gamma) \in P \times \Gamma$ (there are finitely many of them) such that $\langle p, \gamma \rangle \in pre^*(pre^+(C_p^\gamma) \cap (G \times \Gamma^*))$, and then, we construct a $\mathcal{P}$-AMA recognizing the union of the sets $pre^*(C_p^\gamma)$ for all such pairs.

5.2 Model-Checking for CTL

Let $Prop$ be a finite set of atomic propositions and let $\Sigma = 2^{Prop}$. Given a PDS $\mathcal{P}$, a labelling function $\Lambda \in [P \to \Sigma]$, and a CTL formula φ over the set of atomic propositions $Prop$, we want to compute the set of all configurations of $\mathcal{P}$ satisfying φ ; let us denote this set $[\![\varphi]\!]_\mathcal{P}$. We show that it is possible to construct a $\mathcal{P}$-AMA which recognizes $[\![\varphi]\!]_\mathcal{P}$ for every CTL formula φ. The construction of this automaton is described inductively on the structure of CTL formulas as follows:

- Given an atomic formula $\pi \in Prop$, we have $[\![\pi]\!]_\mathcal{P} = \{\langle p, w \rangle \ : \ \pi \in \Lambda(p) \text{ and } w \in \Gamma^*\}$ which is obviously recognizable by a $\mathcal{P}$-AMA.
- The case of boolean connectives is trivial since $\mathcal{P}$-AMA's are closed under boolean operations.
- The case of a formula of the form $\exists\bigcirc\varphi$ is solved using Proposition 4.1 since $[\![\exists\bigcirc\varphi]\!]_\mathcal{P} = pre_\mathcal{P}([\![\varphi]\!]_\mathcal{P})$.

- The case of a formula of the form $\forall\bigcirc\varphi$ is also solved using Proposition 4.1 because $[\![\forall\bigcirc\varphi]\!]_{\mathcal{P}}$ is equal to $pre_{\widetilde{\mathcal{P}}}([\![\varphi]\!]_{\mathcal{P}})$, where $\widetilde{\mathcal{P}}$ is obtained from $\mathcal{P}$ by replacing the transition table Δ with its dual function $\widetilde{\Delta}$ (notice that since $\mathcal{P}$ is an $\exists$PDS, $\widetilde{\mathcal{P}}$ is a $\forall$PDS).

- The cases of formulas of the form $\exists\Diamond\varphi$ and $\forall\Diamond\varphi$ are solved using Theorem 4.1 since $[\![\exists\Diamond\varphi]\!]_{\mathcal{P}} = pre^*_{\mathcal{P}}([\![\varphi]\!]_{\mathcal{P}})$ and $[\![\forall\Diamond\varphi]\!]_{\mathcal{P}} = pre^*_{\widetilde{\mathcal{P}}}([\![\varphi]\!]_{\mathcal{P}})$.

- To solve the cases of formulas of the form $\exists(\varphi_1\mathcal{U}\varphi_2)$ and $\forall(\varphi_1\mathcal{U}\varphi_2)$, we can extend the construction given in Section 4.2 and Theorem 4.1 to the *parametrized* predecessor function. Given a set of configurations $C \subseteq \mathcal{C}$, let $pre_{\mathcal{P}}[C] : 2^{\mathcal{C}} \to 2^{\mathcal{C}}$ be the function such that for every $C' \subseteq \mathcal{C}$, $pre_{\mathcal{P}}[C](C') = C \cap pre(C')$. Then, let $pre^*_{\mathcal{P}}[C]$ be the reflexive-transitive closure of $pre_{\mathcal{P}}[C]$. Clearly, $[\![\exists(\varphi_1\mathcal{U}\varphi_2)]\!]_{\mathcal{P}} = pre^*_{\mathcal{P}}[[\![\varphi_1]\!]_{\mathcal{P}}]([\![\varphi_2]\!]_{\mathcal{P}})$ and $[\![\forall(\varphi_1\mathcal{U}\varphi_2)]\!]_{\mathcal{P}} = pre^*_{\widetilde{\mathcal{P}}}[[\![\varphi_1]\!]_{\mathcal{P}}]([\![\varphi_2]\!]_{\mathcal{P}})$.

6 Conclusion

We have applied the "*symbolic*" analysis principle to a class of infinite state systems, namely pushdown systems. We have proposed a representation structure for their sets of configurations based on finite-state automata and given a procedure for calculating sets of predecessors. This procedure is used to define model-checking algorithms for both linear and branching-time properties. The question which other classes of systems with unbounded variables admit similar techniques is a subject for future investigations.

References

[ACH+95] R. Alur, C. Courcoubetis, N. Halbwachs, T. Henzinger, P. Ho, X. Nicollin, A. Olivero, J. Sifakis, and S. Yovine. The Algorithmic Analysis of Hybrid Systems. *TCS*, 138, 1995.

[AD94] R. Alur and D. Dill. A Theory of Timed Automata. *TCS*, 126, 1994.

[AMP95] E. Asarin, O. Maler, and A. Pnueli. Symbolic Controller Synthesis for Discrete and Timed Systems. In *Hybrid Systems II*. LNCS 999, 1995.

[BG96] B. Boigelot and P. Godefroid. Symbolic Verification of Communication Protocols with Infinite State Spaces using QDDs. In *CAV'96*. to appear, 1996.

[BO93] R.V. Book and F. Otto. *String-Rewriting Systems*. Springer-Verlag, 1993.

[Bry92] R. Bryant. Symbolic Boolean Manipulation with Ordered Binary-Decision Diagrams. *ACM Computing Surveys*, 24, 1992.

[BS94] O. Burkart and B. Steffen. Pushdown Processes: Parallel Composition and Model Checking. In *CONCUR'94*, 1994. LNCS 836.

[CES83] E.M. Clarke, E.A. Emerson, and E. Sistla. Automatic Verification of Finite State Concurrent Systems using Temporal Logic Specifications: A Practical Approach. In *POPL'83*. ACM, 1983.

[McM93] K.L. McMillan. *Symbolic Model-Checking: an Approach to the State-Explosion Problem.* Kluwer, 1993.

[Var95] M. Vardi. Alternating Automata and Program Verification. In *Computer Science Today*. LNCS 1000, 1995.

Appendix

We prove hereafter the correcteness of the construction of the automaton $\mathcal{A}_{pre^*}$. We start by proving that the sequence of the Y_i's eventually reaches its limit.

Lemma 6.1 $\exists i \geq 0.\ Y_{i+1} = Y_i$.

Proof: The usual implication ordering between boolean expressions induces straightforwardly an ordering $\preceq$ between transition tables: Given two transition tables $\delta, \delta' : Q \times \Gamma \rightarrow \mathcal{B}_+(Q)$, we write $\delta \preceq \delta'$ iff $\forall q \in Q$, $\forall \gamma \in \Gamma$, $\delta(q,\gamma)$ implies $\delta'(q,\gamma)$. Let $\equiv$ denote $\preceq \cap \preceq^{-1}$. By construction, the set of states of each of the $\mathcal{A}_i$'s is Q, and it is clear that $\forall i \geq 0$, $\delta_i \preceq \delta_{i+1}$ (δ_{i+1} is obtained by adding new transitions to δ_i). Since there are only finitely many transition tables on Q (modulo $\equiv$) there is necessarily some index i such that $\delta_i \equiv \delta_{i+1}$, and hence $Y_i = Y_{i+1}$. □

Remark 6.1 Notice that the proof of Lemma 6.1 depends on the fact that all the $\mathcal{A}_i$'s have the *same* set of states, and that their transition tables are growing. In order to keep the set of states constant, the use of *alternating* automata is crucial when $\mathcal{P}$ is not an $\exists$PDS.

Now, let us show the inclusion of the X_i's in the Y_i's.

Lemma 6.2 $\forall i \geq 0.\ X_i \subseteq Y_i$.

Proof: Let us denote by $\mathcal{B}_i$ the automaton recognizing the set X_i built using the construction given in Subsection 4.1. It is obvious from the definitions that for every run ρ of $\mathcal{B}_i$ starting from some initial state t^j, there is a run ρ' of $\mathcal{A}_i$ starting from the corresponding initial state s^j such that ρ and ρ' are identical except at their roots. □

It remains to prove that $\forall i \geq 0$, $Y_i \subseteq pre^*(C)$. For that, we need the following key lemma:

Lemma 6.3 *$\forall i \geq 0$, $\forall j \in \{1,\ldots,m\}$, $\forall w \in \Gamma^*$, $\forall U \subseteq Q$, if $s^j \xrightarrow{w}_i U$, then $\exists C_1, C_2 \subseteq \mathcal{C}$ such that $\langle p^j, w\rangle \Rightarrow C_1 \cup C_2$, $C_1 = \{\langle p^k, \varepsilon\rangle \ :\ k \in \{1,\ldots,m\}$ and $s^k \in U\}$, and $\forall c = \langle p^k, v\rangle \in C_2$ with $k \in \{1,\ldots,m\}$, $\exists V_c \subseteq U$. $s^k \xrightarrow{v}_0 V_c$ such that $\bigcup_{c \in C_2} V_c = U \setminus I$.*

Proof: The proof is by induction on i. Let us start by the case $i = 0$. If $w = \varepsilon$ then $U = \{s^j\}$ (thus $U \setminus I = \emptyset$), and in this case we can take $C_1 = \{\langle p^j, \varepsilon\rangle\}$ and $C_2 = \emptyset$ (recall that $\langle p^j, \varepsilon\rangle \Rightarrow \{\langle p^j, \varepsilon\rangle\}$). If $w \neq \varepsilon$ then, since we have supposed that $\mathcal{A}_0 = \mathcal{A}$ has no transitions leading to I, $s^j \xrightarrow{w}_i U$ implies that $U \setminus I = U$, which means that we must take $C_1 = \emptyset$, and then it suffices to take $C_2 = \{\langle p^j, w\rangle\}$.

Now, for $i \geq 1$, we have to prove that

> $\forall r \in \mathcal{R}$, $\forall j \in \{1, \ldots, m\}$, $\forall w \in \Gamma^*$, $\forall U \subseteq Q$, if $s^j \xrightarrow{w}{}_i^r U$, then $\exists C_1, C_2 \subseteq \mathcal{C}$ such that $\langle p^j, w\rangle \Rightarrow C_1 \cup C_2$, $C_1 = \{\langle p^k, \varepsilon\rangle \ : \ k \in \{1, \ldots, m\}$ and $s^k \in U\}$, and
> $\forall c = \langle p^k, v\rangle \in C_2$ with $k \in \{1, \ldots, m\}$, $\exists V_c \subseteq U$. $s^k \xrightarrow{v}_0 V_c$ such that $\bigcup_{c \in C_2} V_c = U \setminus I$.

Let $r = (\langle p^\ell, \gamma\rangle, \{\langle p^{i_1}, w_1\rangle, \ldots, \langle p^{i_n}, w_n\rangle\}) \in \mathcal{R}$. By definition, the transition table of Y_i^r is obtained by adding to δ_{i-1} new γ-transitions (each of them is starting from s^ℓ and going to all the states appearing in a same conjunction in the DNF of the expression $\bigwedge_{k=1}^n \delta_{i-1}^*(s^{i_k}, w_k)$). Then, the proof of the statement above is carried out by induction on the number of times runs starting from initial states (corresponding to transitions of the form $s^j \xrightarrow{w}{}_i^r U$) use one of these new γ-transitions.

Let us suppose that $s^j \xrightarrow{w}{}_i^r U$ and fix a run ρ of $\mathcal{A}_i^r$ over w starting from s^j and such that $Leaves(\rho) = U$. If ρ does not use new γ-transitions, this means that $s^j \xrightarrow{w}_{i-1} U$. Hence, the fact holds by induction hypothesis.

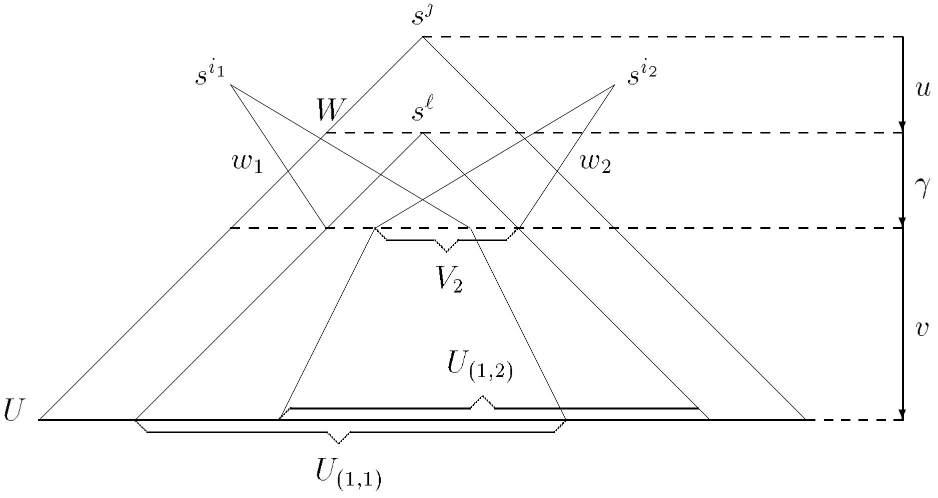
Now, suppose that the new γ-transitions are used at least once in the run ρ. Then, let $w = u\gamma v$, and let ρ_u be a subrun of ρ over u with $W = Leaves(\rho_u)$, such that the new γ-transitions are used at some state in W (notice that this state is necessarily s^ℓ).

Since ρ_u involves less new γ-transitions than the whole run ρ, by induction hypothesis, we have $\langle p^j, u\gamma v\rangle \Rightarrow C_1' \cup C_2'$ such that $C_1' = \{\langle p^\ell, \gamma v\rangle\} \cup \{\langle p^k, \gamma v\rangle \ : \ k \in \{1, \ldots, m\}, k \neq \ell$ and $s^k \in W\}$ and $\forall c = \langle p^k, v_k \gamma v\rangle \in C_2'$ with $k \in \{1, \ldots, m\}$, $\exists V_c \subseteq W$. $s^k \xrightarrow{v_k}_0 V_c$ such that $\bigcup_{c \in C_2'} V_c = W \setminus I$. Then, we have to show that $\exists C_1, C_2 \subseteq \mathcal{C}$ such that $C_1' \cup C_2' \subseteq pre^*(C_1 \cup C_2)$, $C_1 = \{\langle p^k, \varepsilon\rangle \ : \ k \in \{1, \ldots, m\}$ and $s^k \in U\}$, and $\forall c = \langle p^k, v\rangle \in C_2$ with $k \in \{1, \ldots, m\}$, $\exists V_c \subseteq U$. $s^k \xrightarrow{v}_0 V_c$ such that $\bigcup_{c \in C_2} V_c = U \setminus I$. For that, we decompose the proof into three parts according to the kind of configurations in $C_1' \cup C_2'$. First, we consider the particular configuration $\langle p^\ell, \gamma v\rangle \in C_1'$, then the C_1' configurations that are different from $\langle p^\ell, \gamma v\rangle$, and finally the C_2' configurations.

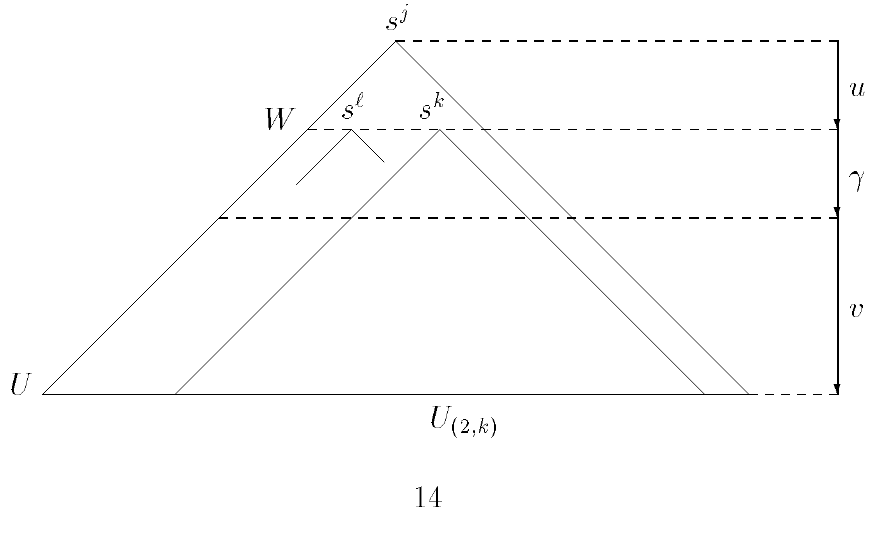
Case 1 Consider the configuration $\langle p^\ell, \gamma v\rangle \in C_1'$. Let $V \subseteq Q$ be the set of states such that $\rho : s^\ell \xrightarrow{\gamma}{}_i^r V$. For every $q \in V$, let ρ_q be the subrun of ρ over v starting from q, and let $U_q \subseteq U$ such that $U_q = Leaves(\rho_q)$. By construction of $\mathcal{A}_i^r$, we have necessarily $\forall k \in \{1, \ldots, n\}$, $\exists V_k \subseteq V$, $s^{i_k} \xrightarrow{w_k}_{i-1} V_k$, and $V = \bigcup_{k=1}^n V_k$. Let $U_{(1,k)} = \bigcup_{q \in V_k} U_q$ and $U_1 = \bigcup_{k=1}^n U_{(1,k)}$. Thus, $\rho : s^\ell \xrightarrow{\gamma v} U_1$, and $\forall k \in \{1, \ldots, n\}$, there exists a run ρ_k over $w_k v$ starting from s^{i_k} and such that $Leaves(\rho_k) = U_{(1,k)}$ ($\rho_k : s^{i_k} \xrightarrow{w_k v}{}_i^r U_{(1,k)}$), with $\rho_k : s^{i_k} \xrightarrow{w_k}{}_i^r V_k$ and $\forall q \in V_k$, the subrun of ρ_k over v starting from q is exactly ρ_q ($\rho_k : q \xrightarrow{v}{}_i^r U_q$). The following picture illustrate this case assuming for simplicity that $n = 2$.

It is easy to see that the number of applications of the new γ-transitions in each run ρ_k is equal to the one in all its subruns ρ_q's, and that this latter is strictly smaller than the one in the whole run ρ. Then, by induction hypothesis, we have $\forall k \in \{1,\ldots,n\}$, $\exists C_1^{(1,k)}, C_2^{(1,k)} \subseteq \mathcal{C}$, $\langle p^{i_k}, w_k v\rangle \Rightarrow C_1^{(1,k)} \cup C_2^{(1,k)}$, $C_1^{(1,k)} = \{\langle p^{k'},\varepsilon\rangle \ : \ k' \in \{1,\ldots,m\} \text{ and } s^{k'} \in U_{(1,k)}\}$, and $\forall c = \langle p^{k'}, v\rangle \in C_2^{(1,k)}$ with $k' \in \{1,\ldots,m\}$, $\exists V_c \subseteq U_{(1,k)}.\ s^{k'} \xrightarrow{v}_0 V_c$ such that $\bigcup_{c\in C_2^{(1,k)}} V_c = U_{(1,k)} \setminus I$.

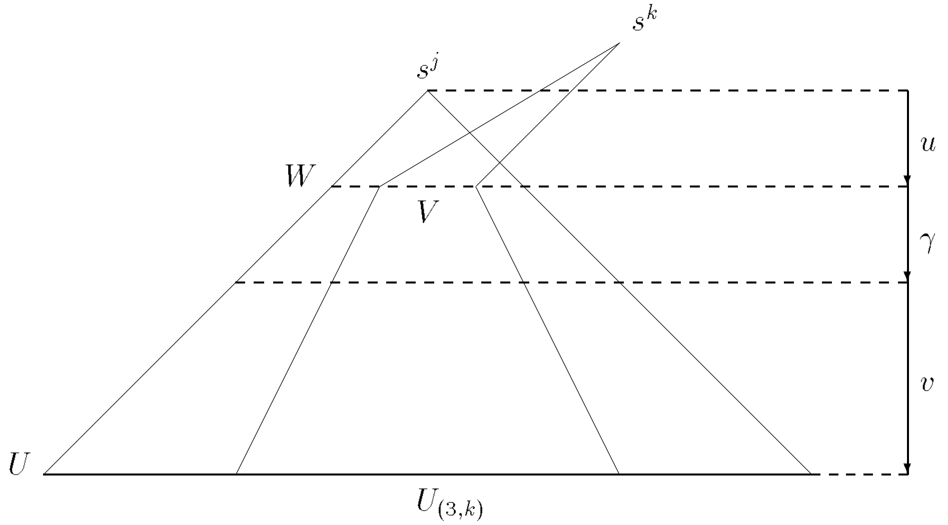
By applying the transition rule r, we have $\langle p^\ell, \gamma v\rangle \Rightarrow \{\langle p^{i_1}, w_1 v\rangle, \ldots, \langle p^{i_n}, w_n v\rangle\}$, and thus $\langle p^\ell,\gamma v\rangle \Rightarrow \bigcup_{k=1}^n (C_1^{(1,k)} \cup C_2^{(1,k)})$. Then, let $C_1^1 = \bigcup_{k=1}^n C_1^{(1,k)}$ and $C_2^{(1,k)} = \bigcup_{k=1}^n C_2^{(1,k)}$. Clearly, $C_1^1 = \{\langle p^k,\varepsilon\rangle \ : \ k\in\{1,\ldots,m\} \text{ and } s^k \in U_1\}$, and $\forall c = \langle p^k, v\rangle \in C_2^1$, $\exists V_c \subseteq U_1.\ s^k \xrightarrow{v}_0 V_c$ such that $\bigcup_{c\in C_2^1} V_c = U_1 \setminus I$.



Case 2 Let $\langle p^k,\gamma v\rangle$ be a configuration of C_1' such that $k\in\{1,\ldots,m\}$, $k\neq\ell$, and $s^k \in W$. Consider the subrun ρ_k of ρ such that $\rho_k : s^k \xrightarrow{\gamma v}{}_i^{r} U_{(2,k)}$ (for some $U_{(2,k)} \subseteq U$). It is clear that ρ_k contains strictly less new γ-transitions than ρ. Then, by induction hypothesis, we have immediately $\exists C_1^{(2,k)}, C_2^{(2,k)} \subseteq \mathcal{C}$, $\langle p^k,\gamma v\rangle \Rightarrow C_1^{(2,k)} \cup C_2^{(2,k)}$, $C_1^{(2,k)} = \{\langle p^{k'},\varepsilon\rangle \ : \ k' \in \{1,\ldots,m\} \text{ and } s^{k'} \in U_{(2,k)}\}$, and $\forall c = \langle p^{k'}, v\rangle \in C_2^{(2,k)}$ with $k' \in \{1,\ldots,m\}$, $\exists V_c \subseteq U_{(2,k)}.\ s^{k'} \xrightarrow{v}_0 V_c$ such that $\bigcup_{c\in C_2^{(2,k)}} V_c = U_{(2,k)}\setminus I$.



Case 3 Let $\langle p^k, v_k\gamma v\rangle$ be a configuration in C_2'. Consider the set of states $V \subseteq W \setminus I$ such that $s^k \xrightarrow{v_k}_0 V$. Recall that we have supposed that the original automaton $\mathcal{A}_0$ has no transitions leading to I. Then, it is easy to see that, by construction, for every $i \geq 1$, the automaton $\mathcal{A}_i$ has no transitions from $Q \setminus I$ to I (because all the new transitions start from a state in I). Consequently, for every $q \in V$, there exists $U_q \subseteq U$ such that $q \xrightarrow{\gamma v}_0 U_q$. Then, we obtain $s^k \xrightarrow{v_k\gamma v}_0 U_{(3,k)} = \bigcup_{q\in V} U_q$.



By induction hypothesis (since the property holds for $i = 0$), we have $\exists C_1^{(3,k)}, C_2^{(3,k)} \subseteq \mathcal{C}$, $\langle p^k, v_k\gamma v\rangle \Rightarrow C_1^{(3,k)} \cup C_2^{(3,k)}$, $C_1^{(3,k)} = \{\langle p^{k'}, \varepsilon\rangle \ : \ k' \in \{1,\ldots,m\}$ and $s^{k'} \in U_{(3,k)}\}$, and $\forall c = \langle p^{k'}, v\rangle \in C_2^{(3,k)}$ with $k' \in \{1,\ldots,m\}$, $\exists V_c \subseteq U_{(3,k)}$. $s^{k'} \xrightarrow{v}_0 V_c$ such that $\bigcup_{c\in C_2^{(3,k)}} V_c = U_{(3,k)} \setminus I$.

Let $U_2 = \bigcup\{U_{(2,k)} \ : \ \langle p^k, \gamma v\rangle \in C_1', k \in \{1,\ldots,m\}, k \neq \ell,$ and $s^k \in W\}$ and $U_3 = \bigcup\{U_{(3,k)} \ : \ \langle p^k, v_k\gamma v\rangle \in C_2'\}$. It is easy to see that $U = \bigcup_{i=1}^3 U_i$. We define also, $i = 1, 2$, $C_i^2 = \bigcup\{C_i^{(2,k)} \ : \ \langle p^k, \gamma v\rangle \in C_1', k \in \{1,\ldots,m\}, k \neq \ell,$ and $s^k \in W\}$ and $C_i^3 = \bigcup\{C_i^{(3,k)} \ : \ \langle p^k, v_k\gamma v\rangle \in C_2'\}$. Then, for $i = 1, 2$, let $C_i = \bigcup_{j=1}^3 C_i^j$. We have shown by considering the three cases above that $C_1' \cup C_2' \subseteq pre^*(C_1 \cup C_2)$, and that $C_1 = \{\langle p^k, \varepsilon\rangle \ : \ k \in \{1,\ldots,m\}$ and $s^k \in U\}$, and $\forall c = \langle p^k, v\rangle \in C_2$ with $k \in \{1,\ldots,m\}$, $\exists V_c \subseteq U$. $s^k \xrightarrow{v}_0 V_c$ such that $\bigcup_{c\in C_2} V_c = U \setminus I$. □

Lemma 6.4 *$\forall i \geq 0$, $Y_i \subseteq pre^*(C)$.*

Proof: The proof follows the same scheme as the one of Lemma 6.3. It is carried out by induction on i. The case $i = 0$ is trivial since $Y_0 = C$. For $i \geq 1$, we have to show that

> $\forall j \in \{1,\ldots,m\}$, $\forall w \in \Gamma^*$, $\forall r \in \mathcal{R}$, $\forall W \subseteq F$, if $s^j \xrightarrow{w}{}_i^r W$ then $\langle p^j, w\rangle \in pre^*(C)$.

Let $r = (\langle p^\ell, \gamma\rangle, \{\langle p^{i_1}, w_1\rangle, \ldots, \langle p^{i_n}, w_n\rangle\}) \in \mathcal{R}$. As for Lemma 6.3, we prove the fact above by induction on the number of times runs starting from initial states use one of the new γ-transitions introduced (due to the rule r) in the construction of $\mathcal{A}_i^r$ from $\mathcal{A}_{i-1}$.

Let us suppose that $s^j \xrightarrow{w}{}^{r}_{i} W$ and fix a run ρ of $\mathcal{A}_i^r$ over w starting from s^j and such that $Leaves(\rho) = W$. If ρ does not use new γ-transitions, we have $s^j \xrightarrow{w}_{i-1} W$, or equivalently $\langle p^j, w\rangle \in Y_{i-1}$. Then, since by induction hypothesis $Y_{i-1} \subseteq pre^*(C)$, we obtain $\langle p^j, w\rangle \in pre^*(C)$.

Consider now the case when the new γ-transitions are used at least once in ρ. Let $w = u\gamma v$, and let ρ_u be a subrun of ρ over u with $U = Leaves(\rho_u)$, such that the new γ-transitions are used for the first time at some state in U (this state is s^ℓ). By Lemma 6.3, $\langle p^j, u\gamma v\rangle \Rightarrow C_1' \cup C_2'$ such that $C_1' = \{\langle p^\ell, \gamma v\rangle\} \cup \{\langle p^k, \gamma v\rangle \;:\; k \in \{1, \ldots, m\}, k \neq \ell \text{ and } s^k \in U\}$ and $\forall c = \langle p^k, v_k\gamma v\rangle \in C_2'$ with $k \in \{1, \ldots, m\}$, $\exists V_c \subseteq U$. $s^k \xrightarrow{v_k}_0 V_c$ such that $\bigcup_{c \in C_2} V_c = U \setminus I$. So, let us show that $C_1' \cup C_2' \subseteq pre^*(C)$. We decompose the proof into three cases:

Case 1 Consider the configuration $\langle p^\ell, \gamma v\rangle \in C_1'$. Let $V \subseteq Q$ be the set of states such that $\rho : s^\ell \xrightarrow{\gamma}{}^{r}_{i} V$. For every $q \in V$, let ρ_q be the subrun of ρ over v starting from q, and let $W_q \subseteq W$ such that $W_q = Leaves(\rho_q)$. By construction of $\mathcal{A}_i^r$, we have $\forall k \in \{1, \ldots, n\}$, $\exists V_k \subseteq V$, $s^{i_k} \xrightarrow{w_k}_{i-1} V_k$, and $V = \bigcup_{k=1}^n V_k$. Let $W_k = \bigcup_{q \in V_k} W_q$ and $W' = \bigcup_{k=1}^n W_k$. Thus, we have $\rho : s^\ell \xrightarrow{\gamma v}{}^{r}_{i} W'$, and $\forall k \in \{1, \ldots, n\}$, there exists a run ρ_k over $w_k v$ starting from s^{i_k} such that $Leaves(\rho_k) = V_k$ ($\rho_k : s^{i_k} \xrightarrow{w_k v}{}^{r}_{i} W_k$), with $\rho_k : s^{i_k} \xrightarrow{w_k}{}^{r}_{i} V_k$ and $\forall q \in V$, the subrun of ρ_k over v starting from q is exactly ρ_q ($\rho_k : q \xrightarrow{v}{}^{r}_{i} W_q$). The number of applications of the new γ-transitions in ρ_k is equal to the one in all the ρ_q's, which is itself strictly smaller than the one in the whole run ρ. Then, by induction hypothesis, we have $\forall k \in \{1, \ldots, n\}$, $\langle p^{i_k}, w_k v\rangle \in pre^*(C)$. Finally, since by applying the transition rule r we have $\langle p^\ell, \gamma v\rangle \Rightarrow \{\langle p^{i_1}, w_1 v\rangle, \ldots, \langle p^{i_k}, w_k v\rangle\}$, we obtain $\langle p^\ell, \gamma v\rangle \in pre^*(C)$.

Case 2 Let $\langle p^k, \gamma v\rangle$ be a configuration of C_1' such that $k \in \{1, \ldots, m\}$, $k \neq \ell$, and $s^k \in U$. Consider the subrun ρ_k of ρ such that $\rho_k : s^k \xrightarrow{\gamma v}{}^{r}_{i} W'$ (for some $W' \subseteq W$). Clearly, ρ_k contains strictly less applications of new γ-transition rules than ρ. Then, by induction hypothesis, we have immediately $\langle p^k, \gamma v\rangle \in pre^*(C)$.

Case 3 Let $\langle p^k, v_k\gamma v\rangle$ be a configuration in C_2'. Consider the set of states $V \subseteq U \setminus I$ such that $s^k \xrightarrow{v_k}_0 V$. As we have said in the proof of Lemma 6.3 (Case 3), all the automata $\mathcal{A}_i$ have no transitions from $Q \setminus I$ to I. Thus, for every $q \in V$, there exists $W_q \subseteq W$ such that $q \xrightarrow{\gamma v}_0 W_q$. Then, $s^k \xrightarrow{v_k\gamma v}_0 \bigcup_{q \in V} W_q$, and by induction hypothesis, we obtain $\langle p^k, v_k\gamma v\rangle \in pre^*(C)$. □

Then, from Lemmas 6.1, 6.2, and 6.4, we deduce Theorem 4.1.