Characterizing Difference of Abstract Probabilistic Automata

Benoît Delahaye[1], Uli Fahrenberg[2], Kim G. Larsen[1], and Axel Legay[2]

1 **Aalborg University, Denmark**
`{benoit,kgl}@cs.aau.dk`
2 **INRIA/IRISA, Rennes, France**
`{ulrich.fahrenberg,axel.legay}@inria.fr`

Abstract

This paper studies counterexample generation for stochastic systems whose specifications are represented by Abstract Probabilistic Automata (APAs). Our contribution is an algorithm that allows to over/under approximate the difference of two APAs with an arbitrary precision. Our technique relies on new notion of distances between APAs used to assess convergence of the approximations as well as on an indeep inspection of the refinement relation defined on APAs. The procedure is effective and not more complex to implement than the refinement checking.

1998 ACM Subject Classification F.1.1 Models of Computation

Keywords and phrases Abstract Probabilistic Automata, Counter-example, Difference

1 Introduction

To establish good specification theories has been the subject of intensive studies, among which one finds classical (temporal) logical specifications such as LTL, MITL, and PCTL [23, 2, 15], and various process algebrae such as CSP [16]. Other specification theories rely on an automata-based representations rather than logical formulae. This is for example the case for Input/Output automata [22], interface automata [8], and modal specifications [20, 24, 5].

In recent work [12, 6], we have proposed Constraint Markov Chains (CMCs), a complete automata-based specification theory for pure stochastic systems, namely Markov Chains (MCs). Contrary to Interval Markov Chains, where sets of distributions are represented by intervals, CMCs use constraints on the next-state probabilities, which makes them closed under both composition and conjunction. Later [9], the CMC approach was extended to handle systems with both stochastic and non-deterministic behaviors, i.e., Probabilistic Automata (PA). Abstract Probabilistic Automata (APAs), whose theory is implemented in the APAC toolset [11], combine Modal Automata and CMCs – the abstractions for labelled transition systems and Markov Chains, respectively. Our theory, which should be viewed as an alternative to classical PCTL and stochastic extensions of CSP, is equipped with a series of operations, including refinement to compare sets of implementations represented by APAs, and a satisfaction relation to decide whether a PA satisfies an APA.

In such context, refinement is crucial for the stepwise implementation of specifications. However, in practice, a successful refinement step may have been preceded by a number of erroneous refinement attempts. To assist the developer, it is of utmost importance that informative debugging information is given in order that she may correct the error as fast as possible.

For model-checking of finite-state systems, debugging information may be given in terms of a single run of the system, thus providing a counterexample to the property of interest. For model-checking of probabilistic systems, where properties have quantitative constraints, a single run does in general not suffice as counterexample, rather a *set* of runs whose total probability violates the given constraint is needed. Here the challenge is how to effectively represent such a set of runs. The

LIPICS Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

problem has recently been studied for a wide range of probabilistic model checking problems [1, 14, 29, 3, 18, 26]. The objective of this paper is to study this problem for APAs; more precisely, we aim at computing effective representations for counterexamples representing differences between APAs. As APAs act both as specifications and abstractions, our work has potential applications in a wide range of areas that includes CEGAR model checking and component-based design.

Our first contribution is to propose a methodology that computes a counterexample PA in case refinement between two APAs does not hold. A key point in our approach is that our algorithm exploits an in-depth inspection of the refinement relation, which makes it easy to generate the reason for which refinement does not hold. In fact, generating the counterexample is not more complex than checking refinement itself. Those advantages are lost when working with different formalisms at the specification and implementation level.

Given a counterexample produced by the methodology above, the refinement attempt may now be modified in order to take account of the provided counterexample. However, as observed for the case of modal specifications [25], there might be several other reasons for the failing refinement which have not been taken into account. Thus, also the modification may lead to failure of the refinement check. It thus seems that it could be beneficial to provide *all* counterexamples rather than just a single one. In particular, in the setting of refinement checking, this calls for a *difference operator*, which given two specifications A and B provides *all* implementations satisfied by A but not B.

We hence propose to represent and synthesize such a difference between two APAs again as an APA. Unfortunately, because APAs are finite-state structures, this is not always possible. To overcome this difficulty, we propose two algorithms used to generate APA under- and over-approximations of the difference. Our algorithms are parametrized by a variable that can be used to tune the precision of the approximation. Using new notions of distances adapted from [28, 4], we show that both approximations converge to the exact difference.

Related work. There has been a lot of recent work on characterizing counterexamples for stochastic systems with respect to (fragments) of PCTL. Some of those works characterize the counterexample by sets of executions [1, 3, 29], by sets of words [26], by probabilistic automata formalisms [18, 14, 17], or even by regular expressions [14]. In [7], Chadha and Viswanathan suggested to use a pair of simulation and properties, and in [19], Komuravelli et.al. employ learning algorithms and counterexamples for PA abstraction refinement. Those approaches can compute the smallest counterexample and are embedded in a CEGAR procedure, which is not the case for our approach. On the other hand, they focus on one single implementation with respect to a given specification. Other approaches such as [13] use may/must abstractions based on under/over approximation. However, this shall not be confused with our approach as they do not intend to compute counterexamples or difference between specifications. In such context, our approach is clearly original as it is the first capable of characterizing the whole set of counterexamples when comparing two specifications, and the only one that is defined on automata-based specifications. A major advantage of our approach is indeed that both specifications and implementations are represented by automata. This makes it easier and more efficient to reason on satisfaction than other approaches where implementations and specifications are represented by different formalisms. As a future work, we should relax the assumption of determinism and embed our work in a CEGAR procedure using APA as an abstraction for implementations; this would certainly require to define a notion of minimality for APA counterexamples. We also plan to explore learning algorithms for more efficient difference generation.

2 Background

This section introduces the background material that will be used throughout the paper. Let *Dist*(S) denote the set of all discrete probability distributions over a finite set S and $\mathbb{B}_2 = \{\top, \bot\}$. We assume that implementations are represented by Probabilistic Automata.

▶ **Definition 1.** A PA [27] is a tuple (S, A, L, AP, V, s_0), where S is a finite set of states with the initial state $s_0 \in S$, A is a finite set of actions, $L\colon S \times A \times \mathit{Dist}(S) \to \mathbb{B}_2$ is a (two-valued) transition function, AP is a finite set of atomic propositions and $V\colon S \to 2^{AP}$ is a state-labeling function.

Consider a state s, an action a, and a probability distribution μ. The value of $L(s, a, \mu)$ is set to $\top$ in case there exists a transition from s under action a to a distribution μ on successor states. In other cases, we have $L(s, a, \mu) = \bot$. We now introduce Abstract Probabilistic Automata (APA) [9], that is a specification theory for PAs. For a finite set S, we let $C(S)$ denote the set of constraints over discrete probability distributions on S. Each element $\varphi \in C(S)$ describes a set of distributions: $\mathit{Sat}(\varphi) \subseteq \mathit{Dist}(S)$. Let $\mathbb{B}_3 = \{\top, ?, \bot\}$. APAs are formally defined as follows.

▶ **Definition 2.** An APA [9] is a tuple (S, A, L, AP, V, S_0), where S is a finite set of states, $S_0 \subseteq S$ is a set of initial states, A is a finite set of actions, and AP is a finite set of atomic propositions. $L : S \times A \times C(S) \to \mathbb{B}_3$ is a *three*-valued distribution-constraint function, and $V : S \to 2^{2^{AP}}$ maps each state in S to a set of admissible labelings.

APAs play the role of specifications in our framework. An APA transition abstracts transitions of a certain unknown PA, called its implementation. Given a state s, an action a, and a constraint φ, the value of $L(s, a, \varphi)$ gives the modality of the transition. More precisely, the value $\top$ means that transitions under a must exist in the PA to some distribution in $\mathit{Sat}(\varphi)$; ? means that these transitions are allowed to exist; $\bot$ means that such transitions must not exist. We will sometimes view L as a *partial* function, with the convention that a lack of value for a given argument is equivalent to the $\bot$ value. The function V labels each state with a subset of the powerset of AP, which models a disjunctive choice of possible combinations of atomic propositions. We say that an APA $N = (S, A, L, AP, V, S_0)$ is in *Single Valuation Normal Form* (SVNF) if the valuation function V assigns at most one valuation to all states, i.e. $\forall s \in S, |V(s)| \leq 1$. From [9], we know that every APA can be turned into an APA in SVNF with the same set of implementations. An APA is *deterministic* [10] if (1) there is at most one outgoing transition for each action in all states, (2) two states with overlapping atomic propositions can never be reached with the same transition, and (3) there is only one initial state.

Let $N = (S, A, L, AP, V, \{s_0\})$ be an APA in SVNF and let $v \subseteq AP$. Given a state $s \in S$ and an action $a \in A$, we will use the notation $\mathsf{succ}_{s,a}(v)$ to represent the set of potential a-successors of s that have v as their valuation. Formally, $\mathsf{succ}_{s,a}(v) = \{s' \in S \mid V(s') = \{v\}, \exists \varphi \in C(S), \mu \in \mathit{Sat}(\varphi) : L(s, a, \varphi) \neq \bot, \mu(s') > 0\}$. When clear from the context, we may use $\mathsf{succ}_{s,a}(s')$ instead of $\mathsf{succ}_{s,a}(V(s'))$. Remark that when N is deterministic, we have $|\mathsf{succ}_{s,a}(v)| \leq 1$ for all s, a, v.

3 Refinement and Distances between APAs

We introduce the notion of refinement between APAs. Roughly speaking, refinement guarantees that if A_1 refines A_2, then the set of implementations of A_1 is included in the one of A_2. We first recall the notion of simulation $\Subset_{\mathcal{R}}$ between two given distributions.

▶ **Definition 3.** Let S and S' be non-empty sets, and μ, μ' be distributions; $\mu \in \mathit{Dist}(S)$ and $\mu' \in \mathit{Dist}(S')$. We say that μ is *simulated* by μ' with respect to a relation $\mathcal{R} \subseteq S \times S'$ and a *correspondence function* $\delta : S \to (S' \to [0, 1])$ iff

1. for all $s \in S$ with $\mu(s) > 0$, $\delta(s)$ is a distribution on S',
2. for all $s' \in S'$, $\sum_{s \in S} \mu(s) \cdot \delta(s)(s') = \mu'(s')$, and
3. whenever $\delta(s)(s') > 0$, then $(s, s') \in \mathcal{R}$.

We write $\mu \Subset_{\mathcal{R}}^{\delta} \mu'$ if μ is simulated by μ' w.r.t $\mathcal{R}$ and δ, and $\mu \Subset_{\mathcal{R}} \mu'$ if there exists δ with $\mu \Subset_{\mathcal{R}}^{\delta} \mu'$.

We will also need distribution simulations without the requirement of a relation $\mathcal{R} \subseteq S \times S'$ (hence also without claim 3 above); these we denote by $\mu \Subset^{\delta} \mu'$. We are ready to define the notion of refinement between APAs.

▶ **Definition 4** (Refinement [10]). Let $N_1 = (S_1, A, L_1, AP, V_1, S_0^1)$ and $N_2 = (S_2, A, L_2, AP, V_2, S_0^2)$ be APAs. A relation $\mathcal{R} \subseteq S_1 \times S_2$ is a *refinement* relation if and only if, for all $(s_1, s_2) \in \mathcal{R}$, we have $V_1(s_1) \subseteq V_2(s_2)$ and

1. $\forall a \in A$, $\forall \varphi_2 \in C(S_2)$, if $L_2(s_2, a, \varphi_2) = \top$, then $\exists \varphi_1 \in C(S_1) : L_1(s_1, a, \varphi_1) = \top$ and $\forall \mu_1 \in Sat(\varphi_1)$, $\exists \mu_2 \in Sat(\varphi_2)$ such that $\mu_1 \Subset_{\mathcal{R}} \mu_2$,
2. $\forall a \in A$, $\forall \varphi_1 \in C(S_1)$, if $L_1(s_1, a, \varphi_1) \neq \bot$, then $\exists \varphi_2 \in C(S_2)$ such that $L_2(s_2, a, \varphi_2) \neq \bot$ and $\forall \mu_1 \in Sat(\varphi_1)$, $\exists \mu_2 \in Sat(\varphi_2)$ such that $\mu_1 \Subset_{\mathcal{R}} \mu_2$.

We say that N_1 refines N_2, denoted $N_1 \preceq N_2$, if and only if there exists a refinement relation such that $\forall s_0^1 \in S_0^1, \exists s_0^2 \in S_0^2$ such that $(s_0^1, s_0^2) \in \mathcal{R}$. Since any PA P is also an APA, we say that P *satisfies* the APA N iff it is a refinement of N. In [10], it is shown that for deterministic APAs N_1, N_2, we have $N_1 \preceq N_2 \iff [\![N_1]\!] \subseteq [\![N_2]\!]$, where $[\![N_i]\!]$ denotes the set of implementations of APA N_i.

To show a convergence theorem about our difference construction in Sect. 5.2 below, we need a relaxed notion of refinement which takes into account that APAs are a *quantitative* formalism. Indeed, refinement as of Def. 4 is a purely qualitative relation; if both $N_2 \not\preceq N_1$ and $N_3 \not\preceq N_1$, then there are no criteria to compare N_2 and N_3 with respect to N_1, saying which one is the closest to N_1. We provide a solution to this question by generalizing refinement to a *distance* which provides precisely such criteria. In Sect. 5.2, we will show how those distances can be used to show that increasingly precise difference approximations between APAs converge to the real difference. The next definition shows how a distance between states is lifted to a distance between constraints.

▶ **Definition 5.** Let $d : S_1 \times S_2 \to \mathbb{R}^+$ and $\varphi_1 \in C(S_1)$, $\varphi_2 \in C(S_2)$ be constraints in N_1 and N_2. Define the distance D_{N_1,N_2} between φ_1 and φ_2 as follows:

$$D_{N_1,N_2}(\varphi_1, \varphi_2, d) = \sup_{\mu_1 \in Sat(\varphi_1)} \inf_{\mu_2 \in Sat(\varphi_2)} \inf_{\delta : \mu_1 \Subset^{\delta} \mu_2} \sum_{(s_1, s_2) \in S_1 \times S_2} \mu_1(s_1)\delta(s_1, s_2)d(s_1, s_2)$$

For the definition of d below, we say that states $s_1 \in S_1$, $s_2 \in S_2$ are *not compatible* if either (1) $V_1(s_1) \neq V_2(s_2)$, (2) there exists $a \in A$ and $\varphi_1 \in C(S_1)$ such that $L_1(s_1, a, \varphi_1) \neq \bot$ and for all $\varphi_2 \in C(S_2), L_2(s_2, a, \varphi_2) = \bot$, or (3) there exists $a \in A$ and $\varphi_2 \in C(S_2)$ such that $L_2(s_2, a, \varphi_2) = \top$ and for all $\varphi_1 \in C(S_1), L_1(s_1, a, \varphi_1) \neq \top$. For compatible states, their distance is similar to the accumulating branching distance on modal transition systems as introduced in [4, 28], adapted to our formalism. In the rest of the paper, the real constant $0 < \lambda < 1$ represents a discount factor. Formally, $d : S_1 \times S_2 \to [0, 1]$ is the least fixpoint to the following system of equations:

$$d(s_1, s_2) = \begin{cases} 1 \text{ if } s_1 \text{ is not compatible with } s_2 & \\ \max \begin{cases} \max\limits_{a, \varphi_1 : L_1(s_1, a, \varphi_1) \neq \bot} \min\limits_{\varphi_2 : L_2(s_2, a, \varphi_2) \neq \bot} \lambda D_{N_1,N_2}(\varphi_1, \varphi_2, d) \\ \max\limits_{a, \varphi_2 : L_2(s_2, a, \varphi_2) = \top} \min\limits_{\varphi_1 : L_1(s_1, a, \varphi_1) = \top} \lambda D_{N_1,N_2}(\varphi_1, \varphi_2, d) \end{cases} & \text{otherwise} \end{cases} \tag{1}$$

Since the above system of linear equations defines a *contraction*, the existence and uniqueness of its least fixpoint is ensured, cf. [21]. This definition intuitively extends to PA implementations, which allows us to propose the two following notions of distance:

▶ **Definition 6.** Let $N_1 = (S_1, A, L_1, AP, V_1, S_0^1)$ and $N_2 = (S_2, A, L_2, AP, V_2, S_0^2)$ be APAs in SVNF. The *syntactic* distance and *thorough* distances between N_1 and N_2 are defined as follows:

- **Syntactic distance.** $d(N_1, N_2) = \max_{s_0^1 \in S_0^1} \big(\min_{s_0^2 \in S_0^2} d(s_0^1, s_0^2) \big)$.
- **Thorough distance.** $d_t(N_1, N_2) = \sup_{P_1 \in [\![N_1]\!]} \big(\inf_{P_2 \in [\![N_2]\!]} d(P_1, P_2) \big)$.

The intuition behind this distance is that $d(s_1, s_2)$ compares not only the probability distributions at s_1 and s_2, but also (recursively) the distributions at all states reachable from s_1 and s_2, weighted by their probability. Each step is discounted by λ, hence steps further in the future contribute less and less to the distance. We also note that $N_1 \preceq N_2$ implies $d(N_1, N_2) = 0$. It can easily be shown, cf. [28], that both d and d_t are *asymmetric pseudometrics*, i.e. that they satisfy $d(N_1, N_1) = 0$ and $d(N_1, N_2) + d(N_2, N_3) \geq d(N_1, N_3)$ (the triangle inequality) for all APAs N_1, N_2, N_3 (and similarly for d_t). The fact that they are only pseudometrics, i.e. that $d(N_1, N_2) = 0$ does not imply $N_1 = N_2$, will play a role in our convergence arguments later. The following theorem shows that the thorough distance is bounded above by the syntactic distance. Hence we can bound distances between (sets of) implementations by the syntactic distance between their specifications.

▶ **Theorem 7.** *For all APAs N_1 and N_2 in SVNF, it holds that $d_t(N_1, N_2) \leq d(N_1, N_2)$.*

4 Counterexample Generation

In this section, we propose a technique that computes a witness PA in case refinement does not hold between two APAs in SVNF.

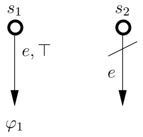
4.1 On Justifying Counterexamples

First, remark that Definition 4 can be trivially turned into an algorithm for checking refinement. Let $N_1 = (S_1, A, L_1, AP, V_1, \{s_0^1\})$ and $N_2 = (S_2, A, L_2, AP, V_2, \{s_0^2\})$ be two deterministic APAs in SVNF. Consider the initial relation $\mathcal{R}_0 = S_1 \times S_2$. Compute $\mathcal{R}_{k+1}$ by removing all pairs of states not satisfying Definition 4 for $\mathcal{R}_k$. The sequence $(\mathcal{R}_n)_{n\in\mathbb{N}}$ is then strictly decreasing and converges to a fixpoint within a finite number of steps $K \leq |S_1 \times S_2|$. This fixpoint $\mathcal{R}_K$ coincides with the maximal refinement relation $\mathcal{R}$ between N_1 and N_2. Let the index of this fixpoint be denoted with $\mathsf{Ind}(\mathcal{R}) = K$. In the following, $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2)$ denotes the maximal index $k \leq K$ for which $(s_1, s_2) \in \mathcal{R}_k$. Formally, let $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) = \min(\max(\{k \mid (s_1, s_2) \in \mathcal{R}_k\}), K)$.

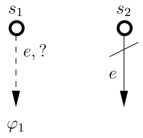
Let us assume that $N_1 \not\preceq N_2$, and let $\mathcal{R}$ be a maximal refinement relation between N_1 and N_2. Since $N_1 \not\preceq N_2$, we know that $(s_0^1, s_0^2) \notin \mathcal{R}$. Given $(s_1, s_2) \in S_1 \times S_2$, we can distinguish between the following cases:

1. $(s_1, s_2) \in \mathcal{R}$
2. $V_1(s_1) \neq V_2(s_2)$,
3. $(s_1, s_2) \notin \mathcal{R}$ and $V_1(s_1) = V_2(s_2)$, and

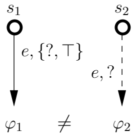
   a. there exists $e \in A$ and $\varphi_1 \in C(S_1)$ such that $L_1(s_1, e, \varphi_1) = \top$ and $\forall \varphi_2 \in C(S_2) : L_2(s_2, e, \varphi_2) = \bot$,



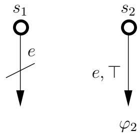
   b. there exists $e \in A$ and $\varphi_1 \in C(S_1)$ such that $L_1(s_1, e, \varphi_1) =?$ and $\forall \varphi_2 \in C(S_2) : L_2(s_2, e, \varphi_2) = \bot$,



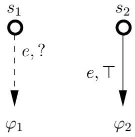
c. there exists $e \in A$ and $\varphi_1 \in C(S_1)$ such that $L_1(s_1, e, \varphi_1) \geq ?$ and $\exists \varphi_2 \in C(S_2) : L_2(s_2, e, \varphi_2) = ?, \exists \mu \in Sat(\varphi_1)$ such that $\forall \mu' \in Sat(\varphi_2) : \mu \notin_{\mathcal{R}} \mu'$,



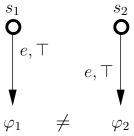
d. there exists $e \in A$ and $\varphi_2 \in C(S_2)$ such that $L_2(s_2, e, \varphi_2) = \top$ and $\forall \varphi_1 \in C(S_1) : L_1(s_1, e, \varphi_1) = \bot$,



e. there exists $e \in A$ and $\varphi_2 \in C(S_2)$ such that $L_2(s_2, e, \varphi_2) = \top$ and $\exists \varphi_1 \in C(S_1) : L_1(s_1, e, \varphi_1) = ?$,



f. there exists $e \in A$ and $\varphi_2 \in C(S_2)$ such that $L_2(s_2, e, \varphi_2) = \top$, $\exists \varphi_1 \in C(S_1) : L_1(s_1, e, \varphi_1) = \top$ and $\exists \mu \in Sat(\varphi_1)$ such that $\forall \mu' \in Sat(\varphi_2) : \mu \notin_{\mathcal{R}} \mu'$.



Remark that because of the determinism and SVNF of APAs N_1 and N_2, cases 1, 2 and 3 cannot happen at the same time. Moreover, although the cases in 3 can happen simultaneously, they cannot be "triggered" by the same action. In order to keep track of these "concurrent" situations, we define the following sets.

Given a pair of states (s_1, s_2), let us define $B_a(s_1, s_2)$ to be the set of actions in A such that case $3.a$ above holds. If there is no such action, then $B_a(s_1, s_2) = \emptyset$. Similarly, we define $B_b(s_1, s_2), B_c(s_1, s_2), B_d(s_1, s_2), B_e(s_1, s_2)$ and $B_f(s_1, s_2)$ to be the sets of actions such that cases $3.b, c, d, e$ and $3.f$ holds respectively. Given a set $X \subseteq \{a, b, c, d, e, f\}$, let $B_X(s_1, s_2) = \cup_{x \in X} B_x(s_1, s_2)$. In addition, let $B(s_1, s_2) = B_{\{a,b,c,d,e,f\}}(s_1, s_2)$.

4.2 Computing Counterexamples

We first observe that if a pair of states (s_1, s_2) is removed from the relation $\mathcal{R}$ by case 3, then we need to keep track of the actions that lead to this removal in order to use them in our counterexample. Whenever a pair of states is in cases 3.a, 3.b, 3.d or 3.e, we have that $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) = 0$ and the counterexample can be easily produced by allowing or disallowing the corresponding transitions from N_1 and N_2. Cases 3.c and 3.f play a different role: due to the fact that they exploit distributions, they are the only cases in which refinement can be broken by using its *recursive* axiom. In these cases, producing a counterexample can be done in two ways: either by using a distribution that does not satisfy the constraints in N_2 (if such a distribution exists, then $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) = 0$), or by using a distribution that reaches a pair of states $(s_1', s_2') \notin \mathcal{R}$. When $0 < \mathsf{Ind}_{\mathcal{R}}(s_1, s_2) < \mathsf{Ind}(\mathcal{R})$, only the latter is possible. This recursive construction has disadvantages: it allows us to produce loops that may lead to incorrect counterexamples. In order to prevent these loops, we propose to use only those distributions that decrease the value of Ind in this particular case. The set $\mathsf{Break}(s_1, s_2)$ defined hereafter allows us to distinguish the actions for which the value of Ind decreases, hence ensuring (by Lemma 8 below) the correctness of our counterexample construction. Let $(s_1, s_2) \in S_1 \times S_2$ be such that $V_1(s_1) \subseteq V_2(s_2)$ and $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) = k < \mathsf{Ind}(\mathcal{R})$. We define $\mathsf{Break}(s_1, s_2)$ to be the set $\{a \in A \mid$ either $a \in B_{a,b,d,e}(s_1, s_2)$ or there exists $\varphi_1 \in C(S_1)$ such that $L_1(s_1, a, \varphi_1) \neq \bot$, $\varphi_2 \in C(S_2)$ such that $L_2(s_2, a, \varphi_2) \neq \bot$ and $\mu_1 \in Sat(\varphi_1)$ such that $\forall \mu_2 \in Sat(\varphi_2), \mu_1 \notin_{\mathcal{R}_k} \mu_2\}$.

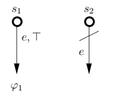
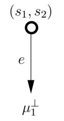
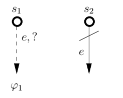
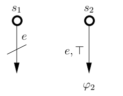
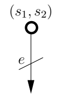
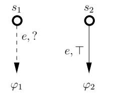
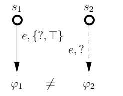
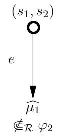
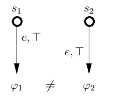
$e \in$	N_1, N_2	P	Formal Definition of L
$B_a(s_1, s_2)$			Let $\varphi_1 \in C(S_1)$ such that $L_1(s_1, e, \varphi_1) \neq \bot$ and let μ_1 be an arbitrary distribution in $Sat(\varphi_1)$. Define $L((s_1, s_2), e, \mu_1^\perp) = \top$.
$B_b(s_1, s_2)$			
$B_d(s_1, s_2)$			For all $\mu \in Dist(S)$, let $L((s_1, s_2), e, \mu) = \bot$.
$B_e(s_1, s_2)$	 s_1 $e, ?$ φ_1 s_2 $e, \top$ φ_2		
$B_c(s_1, s_2)$	 s_1 $e, \{?, \top\}$ φ_1 $\neq$ s_2 $e, ?$ φ_2	 (s_1, s_2) e $\widehat{\mu_1}$ $\notin_\mathcal{R} \varphi_2$	Let $\varphi_1 \in C(S_1)$ and $\varphi_2 \in C(S_2)$ such that $L_1(s_1, e, \varphi_1) \neq \bot$ and $L_2(s_2, e, \varphi_2) \neq \bot$. - If $e \in \mathsf{Break}(s_1, s_2)$, then let μ_1 be the distribution given in Lemma 8. - Else, let μ_1 be an arbitrary distribution in $Sat(\varphi_1)$ such that $\forall \mu_2 \in Sat(\varphi_2), \mu_1 \notin_\mathcal{R} \mu_2$. In both cases, let $L((s_1, s_2), e, \widehat{\mu_1}) = \top$.
$B_f(s_1, s_2)$	 s_1 $e, \top$ φ_1 $\neq$ s_2 $e, \top$ φ_2		

Table 1 Definition of the transition function L in P.

Remark that the conditions defined above are exactly the conditions for removing a pair of states (s_1, s_2) at step k of the algorithm for computing $\mathcal{R}$ defined in Section 4.1 above. Under the assumption that $V_1(s_1) \subseteq V_2(s_2)$ and $\mathsf{Ind}_\mathcal{R}(s_1, s_2) = k < \mathsf{Ind}(\mathcal{R})$, we can be sure that the set $\mathsf{Break}(s_1, s_2)$ is not empty. Moreover, we have the following lemma.

▶ **Lemma 8.** *For all pairs of states (s_1, s_2) in case 3 and for all actions $e \in (B_c(s_1, s_2) \cup B_f(s_1, s_2)) \cap \mathsf{Break}(s_1, s_2)$, there exist constraints φ_1 and φ_2 such that $L_1(s_1, e, \varphi_1) \neq \bot$ and $L_2(s_2, e, \varphi_2) \neq \bot$ and a distribution $\mu_1 \in Sat(\varphi_1)$ such that either μ_1 breaks φ_2 directly, or there exists $s_1' \in S_1, s_2' \in S_2$ such that $\mu_1(s_1') > 0, s_2' = \mathsf{succ}_{s_2,e}(s_1')$ and $\mathsf{Ind}_\mathcal{R}(s_1', s_2') < \mathsf{Ind}_\mathcal{R}(s_1, s_2)$.*

In other words, the above lemma ensures that a pair (s_1', s_2') such that $\mathsf{Ind}_\mathcal{R}(s_1', s_2') = 0$ can be reached within a bounded number of transitions for all pairs of states (s_1, s_2) in case 3. As explained above, this is a prerequisite for the correctness of the counterexample construction defined hereafter.

We now propose the main contribution of the section: a construction to build counterexamples. Let $N_1 = (S_1, A, L_1, AP, V_1, \{s_0^1\})$ and $N_2 = (S_2, A, L_2, AP, V_2, \{s_0^2\})$ be deterministic APAs in SVNF such that $N_1 \not\preceq N_2$. Let $\mathcal{R}$ be the maximal refinement relation between N_1 and N_2..

▶ **Definition 9.** The counterexample $P = (S, A, L, AP, V, s_0)$ is computed as follows:

- $S = S_1 \times (S_2 \cup \{\bot\})$,
- $s_0 = (s_0^1, s_0^2)$,
- $V(s_1, s_2) = v \in 2^{AP}$ such that $V_1(s_1) = \{v\}$ for all $(s_1, s_2) \in S$, and
- L is defined as follows. Let $(s_1, s_2) \in S$.
  - If (s_1, s_2) in case 1 or 2 or $s_2 = \bot$, then for all $a \in A$ and $\varphi_1 \in C(S_1)$ such that $L_1(s_1, a, \varphi_1) = \top$, let μ_1 be an arbitrary distribution in $Sat(\varphi_1)$ and let $L((s_1, s_2), a, \mu_1^\perp) =$

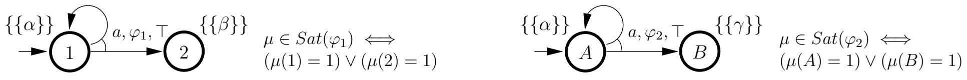
$\top$ with $\mu_1^{\perp} \in Dist(S)$ such that $\mu_1^{\perp}(s_1', s_2') = \mu_1(s_1')$ if $s_2' = \perp$ and 0 otherwise.

- Else, (s_1, s_2) is in case 3 and $B(s_1, s_2) \neq \emptyset$. For all $a \in A \setminus B(s_1, s_2)$ and $\varphi_1 \in C(S_1)$ such that $L_1(s_1, a, \varphi_1) = \top$, let μ_1 be an arbitrary distribution in $Sat(\varphi_1)$ and let $L((s_1, s_2), a, \mu_1^{\perp}) = \top$, with $\mu_1^{\perp}$ defined as above.
  In addition, for all $e \in B(s_1, s_2)$, let $L((s_1, s_2), e, .)$ be defined as in Table 1. In the table, given constraints $\varphi_1 \in C(S_1)$ and $\varphi_2 \in C(S_2)$ such that $L_1(s_1, e, \varphi_1) \neq \perp$ and $L_2(s_2, e, \varphi_2) \neq \perp$, and a distribution $\mu_1 \in Sat(\varphi_1)$, the distribution $\widehat{\mu_1} \in Dist(S)$ is defined as follows: $\widehat{\mu_1}(s_1', s_2') = \mu_1(s_1)$ if $s_2' = \mathsf{succ}_{s_2,e}(s_1')$ or $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$ and $s_2' = \perp$, and 0 otherwise.

▶ **Theorem 10.** *The counterexample PA P defined above is such that $P \models N_1$ and $P \not\models N_2$.*

5 Difference operators for APAs

We now try to compute an APA that represents the difference between the sets of implementations of two APAs. We first observe that such a set may not be representable by an APA, then we will propose over- and under-approximations. Consider the APAs N_1 and N_2 given in Figures 1a and 1b. Consider the difference of their sets of implementations. It is easy to see that this set contains all the PAs that can finitely loop on valuation α and then move into a state with valuation β. Since there is no bound on the time spent in the loop, there is no finite-state APA that can represent this set of implementations.


$\{\{\alpha\}\}$
$a, \varphi_1, \top$
$\{\{\beta\}\}$
1
2
$\mu \in Sat(\varphi_1) \iff (\mu(1) = 1) \vee (\mu(2) = 1)$
$\{\{\alpha\}\}$
$a, \varphi_2, \top$
$\{\{\gamma\}\}$
A
B
$\mu \in Sat(\varphi_2) \iff (\mu(A) = 1) \vee (\mu(B) = 1)$

(a) APA N_1

(b) APA N_2

Figure 1 APAs N_1 and N_2 such that $[\![N_1]\!] \setminus [\![N_2]\!]$ cannot be represented using a finite-state APA.

5.1 Over-approximation of the difference between APAs

In this section, we propose a construction $\setminus^*$ that over-approximates the difference between APAs in the following sense: given two deterministic APAs $N_1 = (S_1, A, L_1, AP, V_1, \{s_0^1\})$ and $N_2 = (S_2, A, L_2, AP, V_2, \{s_0^2\})$ in SVNF, such that $N_1 \not\preceq N_2$, we have $[\![N_1]\!] \setminus [\![N_2]\!] \subseteq [\![N_1 \setminus^* N_2]\!]$. We first observe that if $V_1(s_0^1) \neq V_2(s_0^2)$, i.e. (s_0^1, s_0^2) in case 2, then $[\![N_1]\!] \cap [\![N_2]\!] = \emptyset$. In such case, we define $N_1 \setminus^* N_2$ as N_1. Otherwise, we build on the reasons for which refinement fails between N_1 and N_2. We first give an informal intuition of how the construction works and then define it formally.

In our construction, states in $N_1 \setminus^* N_2$ will be elements of $S_1 \times (S_2 \cup \{\perp\}) \times (A \cup \{\varepsilon\})$. Our objective is to ensure that any implementation in our counterexample will satisfy N_1 and not N_2. In (s_1, s_2, e), states s_1 and s_2 keep track of executions of N_1 and N_2. Action e is the action of N_1 that will be used to break satisfaction with respect to N_2, i.e. the action that will be the cause for which any implementation of (s_1, s_2, e) cannot be accepted by N_2. Since satisfaction is defined recursively, the breaking is not necessarily immediate and can be postponed to successors. $\perp$ is used to represent states that can only be reached after breaking the satisfaction relation to N_2. In these states, we do not need to keep track of the corresponding execution in N_2, thus only focus on satisfying N_1. States of the form (s_1, s_2, ε) with $s_2 \neq \perp$ are states where the satisfaction is broken by a distribution that does not match constraints in N_2 (cases 3.c and 3.f). In order to invalidate these constraints, we still need to keep track of the corresponding execution in N_2, hence the use of ε instead of $\perp$.

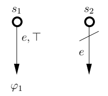
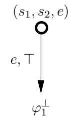
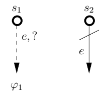
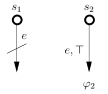
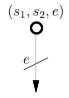
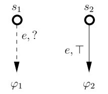
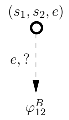
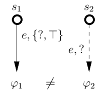
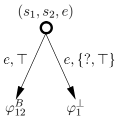
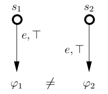
$e \in$	N_1, N_2	$N_1 \setminus^* N_2$	Formal Definition of L
$B_a(s_1, s_2)$			For all $a \neq e \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$, let $L((s_1, s_2, e), a, \varphi^\perp) = L_1(s_1, a, \varphi)$. In addition, let $L((s_1, s_2, e), e, \varphi_1^\perp) = \top$. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e), b, \varphi) = \perp$.
$B_b(s_1, s_2)$			
$B_d(s_1, s_2)$			For all $a \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$, let $L((s_1, s_2, e), a, \varphi^\perp) = L_1(s_1, a, \varphi)$. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e), b, \varphi) = \perp$.
$B_e(s_1, s_2)$	 s_1 s_2 $e, ?$ $e, \top$ φ_1 φ_2		For all $a \neq e \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$, let $L((s_1, s_2, e), a, \varphi^\perp) = L_1(s_1, a, \varphi)$. In addition, let $L((s_1, s_2, e), e, \varphi_{12}^B) =?$. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e), b, \varphi) = \perp$.
$B_c(s_1, s_2)$	 s_1 s_2 $e, \{?, \top\}$ $e, ?$ φ_1 $\neq$ φ_2	 (s_1, s_2, e) $e, \top$ $e, \{?, \top\}$ φ_{12}^B $\varphi_1^\perp$	For all $a \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$ (including e and φ_1), let $L((s_1, s_2, e), a, \varphi^\perp) = L_1(s_1, a, \varphi)$. In addition, let $L((s_1, s_2, e), e, \varphi_{12}^B) = \top$. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e), b, \varphi) = \perp$.
$B_f(s_1, s_2)$	 s_1 s_2 $e, \top$ $e, \top$ φ_1 $\neq$ φ_2		

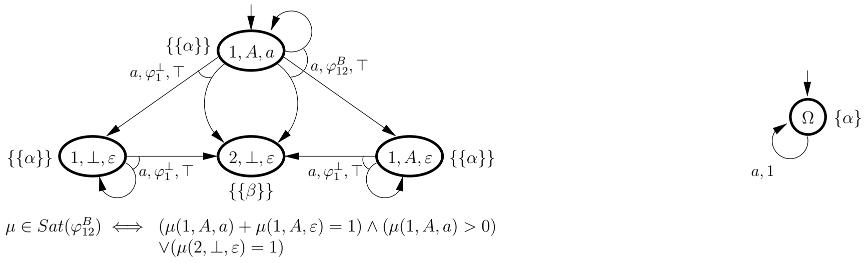
Table 2 Definition of the transition function L in $N_1 \setminus^* N_2$.

The transitions in our construction will match the different cases shown in the previous section, ensuring that in each state, either the relation is broken immediately or reported to at least one successor. Since there can be several ways of breaking the relation in state (s_0^1, s_0^2), each corresponding to an action $e \in B(s_0^1, s_0^2)$, the APA $N_1 \setminus^* N_2$ will have one initial state for each of them. Formally, if (s_0^1, s_0^2) is in case 3, we define the over-approximation of the difference of N_1 and N_2 as follows.

▶ **Definition 11.** Let $N_1 \setminus^* N_2 = (S, A, L, AP, V, S_0)$, where $S = S_1 \times (S_2 \cup \{\perp\}) \times (A \cup \{\varepsilon\})$, $V(s_1, s_2, a) = V(s_1)$ for all s_2 and a, $S_0 = \{(s_0^1, s_0^2, f) \mid f \in B(s_0^1, s_0^2)\}$, and L is defined by:

- If $s_2 = \perp$ or $e = \varepsilon$ or (s_1, s_2) in case 1 or 2, then for all $a \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$, let $L((s_1, s_2, e), a, \varphi^\perp) = L_1(s_1, a, \varphi)$, with $\varphi^\perp$ defined below. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e), b, \varphi) = \perp$.
- Else, we have (s_1, s_2) in case 3 and $B(s_1, s_2) \neq \emptyset$ by construction. The definition of L is given in Table 2, with the constraints $\varphi^\perp$ and φ_{12}^B defined hereafter.

Given $\varphi \in C(S_1)$, $\varphi^\perp \in C(S)$ is defined as follows: $\mu \in Sat(\varphi^\perp)$ iff $\forall s_1 \in S_1, \forall s_2 \neq \perp, \forall b \neq \varepsilon, \mu(s_1, s_2, b) = 0$ and the distribution $(\mu \downarrow_1 \colon s_1 \mapsto \mu(s_1, \perp, \varepsilon))$ is in $Sat(\varphi)$. Given a state $(s_1, s_2, e) \in S$ with $s_2 \neq \perp$ and $e \neq \varepsilon$ and two constraints $\varphi_1 \in C(S_1)$, $\varphi_2 \in C(S_2)$ such that $L_1(s_1, e, \varphi_1) \neq \perp$ and $L_2(s_2, e, \varphi_2) \neq \perp$, the constraint $\varphi_{12}^B \in C(S)$ is defined as follows: $\mu \in Sat(\varphi_{12}^B)$ iff (1) for all $(s_1', s_2', c) \in S$, we have $\mu(s_1', s_2', c) > 0 \Rightarrow s_2' = \perp$ if $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$ and $s_2' = \mathsf{succ}_{s_2,e}(s_1')$ otherwise, and $c \in B(s_1', s_2') \cup \{\varepsilon\}$, (2) the distribution $\mu_1 : s_1' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_2' \in S_2 \cup \{\perp\}} \mu(s_1', s_2', c)$ satisfies φ_1, and (3) either (a) there exists $(s_1', \perp, c)$ such that $\mu(s_1', \perp, c) > 0$ or (b) the distribution $\mu_2 : s_2' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_1' \in S_1} \mu(s_1', s_2', c)$ does not satisfy φ_2, or (c) there exists $s_1' \in S_1$, $s_2' \in S_2$ and $c \neq \varepsilon$ such that $\mu(s_1', s_2', c) > 0$. Informally, distributions in φ_{12}^B must (1) follow the corresponding execution is N_1 and N_2 if possible, (2) satisfy φ_1 and (3)



(a) $N_1 \setminus^* N_2$ **(b)** P

Figure 2 Over-approximating difference $N_1 \setminus^* N_2$ of APAs N_1 and N_2 from Figure 1 and PA P such that $P \models N_1 \setminus^* N_2$ and $P \models N_2$.

either (a) reach a state in N_1 that cannot be matched in N_2 or (b) break the constraint φ_2, or (c) report breaking the relation to at least one successor state. The following theorem shows that $N_1 \setminus^* N_2$ is an over-approximation of the difference of N_1 and N_2 in terms of sets of implementations.

▶ **Theorem 12.** *For all deterministic APAs N_1 and N_2 in SVNF such that $N_1 \not\preceq N_2$, we have $[\![N_1]\!] \setminus [\![N_2]\!] \subseteq [\![N_1 \setminus^* N_2]\!]$.*

The reverse inclusion in the above theorem does not hold. Intuitively, as explained in the construction of the constraint φ_{12}^B above, one can postpone the breaking of the satisfaction relation for N_2 to the next state (condition (3.c)). This assumption is necessary in order to produce an APA representing *all* counterexamples. However, when there are cycles in the execution of $N_1 \setminus^* N_2$, this assumption allows to postpone forever, thus allowing for implementations that will ultimately satisfy N_2. This is illustrated in the following example.

▶ **Example 13.** Consider the APAs N_1 and N_2 given in Fig. 1. Their over-approximating difference $N_1 \setminus^* N_2$ is given in Fig. 2a. On can see that the PA P in Fig. 2b satisfies both $N_1 \setminus^* N_2$ and N_2.

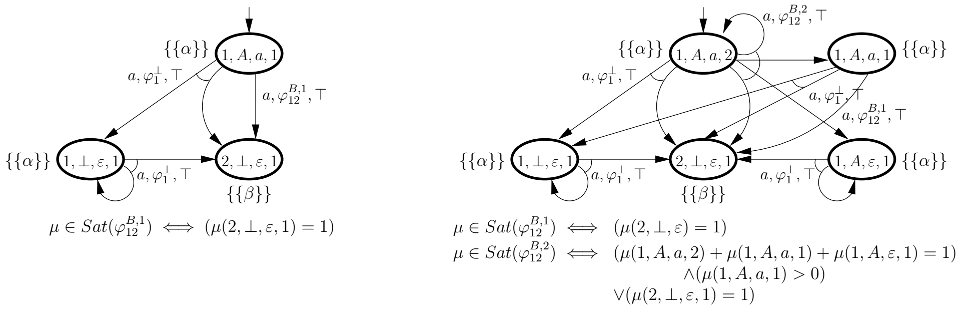
5.2 Under-approximation of the difference between APAs

We now propose a construction that under-estimates the difference between APAs. This construction resembles the over-approximation of the difference presented in the previous section, the main difference being that in the under-approximation, states are indexed with an integer that represents the maximal depth of the unfolding of counterexamples. The construction is as follows.

Let $N_1 = (S_1, A, L_1, AP, V_1, \{s_0^1\})$ and $N_2 = (S_2, A, L_2, AP, V_2, \{s_0^2\})$ be two deterministic APAs in SVNF such that $N_1 \not\preceq N_2$. Let $K \in \mathbb{N}$ be the parameter of our construction. As in Section 5.1, if $V_1(s_0^1) \neq V_2(s_0^2)$, i.e. (s_0^1, s_0^2) in case 2, then $[\![N_1]\!] \cap [\![N_2]\!] = \emptyset$. In this case, we define $N_1 \setminus^K N_2$ as N_1. Otherwise, the under-approximation is defined as follows.

▶ **Definition 14.** Let $N_1 \setminus^K N_2 = (S, A, L, AP, V, S_0^K)$, where $S = S_1 \times (S_2 \cup \{\perp\}) \times (A \cup \{\varepsilon\}) \times \{1, \ldots, K\}$, $V(s_1, s_2, a, k) = V(s_1)$ for all s_2, a, $k < K$, $S_0^K = \{(s_0^1, s_0^2, f, K) \mid f \in B(s_0^1, s_0^2)\}$, and L is defined as below, where we report to Table 2 for the structure. Let $(s_1, s_2, e, k) \in S$.

- If $s_2 = \perp$ or $e = \varepsilon$ or (s_1, s_2) in case 1 or 2, then for all $a \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$, let $L((s_1, s_2, e, k), a, \varphi^\perp) = L_1(s_1, a, \varphi)$, with $\varphi^\perp$ defined below. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e, k), b, \varphi) = \perp$.
- Else we have (s_1, s_2) in case 3 and $B(s_1, s_2) \neq \emptyset$ by construction. The definition of L is similar to the definition given in Table 2 (the full table being given in Appendix E for space reasons), where (s_1, s_2, e) and φ_{12}^B are replaced with (s_1, s_2, e, k) and $\varphi_{12}^{B,k}$ respectively. The constraints $\varphi^\perp$ and $\varphi_{12}^{B,k}$ are defined hereafter.



(a) $N_1 \setminus^1 N_2$ **(b)** $N_1 \setminus^2 N_2$

Figure 3 Under-approximations at level 1 and 2 of the difference of APAs N_1 and N_2 from Figure 1.

Given a constraint $\varphi \in C(S_1)$, the constraint $\varphi^\perp \in C(S)$ is defined as follows: $\mu \in Sat(\varphi^\perp)$ iff $\forall s_1 \in S_1, \forall s_2 \neq \perp, \forall b \neq \varepsilon, \forall k \neq 1, \mu(s_1, s_2, b, k) = 0$ and the distribution $(\mu \downarrow_1 \colon s_1 \mapsto \mu(s_1, \perp, \varepsilon, 1))$ is in $Sat(\varphi)$. Given a state $(s_1, s_2, e, k) \in S$ with $s_2 \neq \perp$ and $e \neq \varepsilon$ and two constraints $\varphi_1 \in C(S_1)$ and $\varphi_2 \in C(S_2)$ such that $L_1(s_1, e, \varphi_1) \neq \perp$ and $L_2(s_2, e, \varphi_2) \neq \perp$, the constraint $\varphi_{12}^{B,k} \in C(S)$ is defined as follows: $\mu \in Sat(\varphi_{12}^{B,k})$ iff (1) for all $(s_1', s_2', c, k') \in S$, if $\mu(s_1', s_2', c, k') > 0$, then $c \in B(s_1', s_2') \cup \{\varepsilon\}$ and either $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$, $s_2' = \perp$ and $k' = 1$, or $s_2' = \mathsf{succ}_{s_2,e}(s_1')$, (2) the distribution $\mu_1 : s_1' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_2' \in S_2 \cup \{\perp\}, k' \geq 1} \mu(s_1', s_2', c, k')$ satisfies φ_1, and (3) either (a) there exists $(s_1', \perp, c, 1)$ such that $\mu(s_1', \perp, c, 1) > 0$, or (b) the distribution $\mu_2 : s_2' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_1' \in S_1, k' \geq 1} \mu(s_1', s_2', c, k')$ does not satisfy φ_2, or (c) $k \neq 1$ and there exists $s_1' \in S_1$, $s_2' \in S_2$, $c \neq \varepsilon$ and $k' < k$ such that $\mu(s_1', s_2', c, k') > 0$. The construction is illustrated in Figure 3.

In the following theorem we show that $N_1 \setminus^K N_2$ is a correct under-approximation of the difference of N_1 and N_2 in terms of sets of implementations. The intuition of the theorem is that for K increasing, the under-approximation is getting better and better, so that in a set-theoretic sense, $\lim_{K \to \infty} [\![N_1 \setminus^K N_2]\!] = [\![N_1]\!] \setminus [\![N_2]\!]$.

▶ **Theorem 15.** *For all deterministic APAs N_1 and N_2 in SVNF such that $N_1 \not\preceq N_2$:*

1. *for all $K \in \mathbb{N}$, we have $N_1 \setminus^K N_2 \preceq N_1 \setminus^{K+1} N_2$,*
2. *for all $K \in \mathbb{N}$, $[\![N_1 \setminus^K N_2]\!] \subseteq [\![N_1]\!] \setminus [\![N_2]\!]$, and*
3. *for all PA $P \in [\![N_1]\!] \setminus [\![N_2]\!]$, there exists $K \in \mathbb{N}$ such that $P \in [\![N_1 \setminus^K N_2]\!]$.*

We now present the main result of this paper, using the above set-theoretic convergence theorem and the distance framework introduced in Section 3.

▶ **Theorem 16.** *Let N_1 and N_2 be two deterministic APAs in SVNF such that $N_1 \not\preceq N_2$. Then*

1. *the sequences $(N_1 \setminus^K N_2)_{K \in \mathbb{N}}$ and $([\![N_1 \setminus^K N_2]\!])_{K \in \mathbb{N}}$ both converge,*
2. *$\lim_{K \to \infty} d_t([\![N_1]\!] \setminus [\![N_2]\!], [\![N_1 \setminus^K N_2]\!] = 0$, and*
3. *$\lim_{K \to \infty} d(N_1 \setminus^* N_2, N_1 \setminus^K N_2) = 0$, so that*
4. *$d_t([\![N_1 \setminus^* N_2]\!], [\![N_1]\!] \setminus [\![N_2]\!]) = 0$.*

Hence the sequences $(N_1 \setminus^K N_2)_{K \in \mathbb{N}}$ and $([\![N_1 \setminus^K N_2]\!])_{K \in \mathbb{N}}$ converge with respect to the syntactic and thorough distances, respectively. A limit of $(N_1 \setminus^K N_2)_{K \in \mathbb{N}}$ is $N_1 \setminus^* N_2$ (recall that as d is not a metric, the sequence may have more than one limit), while $[\![N_1]\!] \setminus [\![N_2]\!]$ is also a metric (not only a set-theoretic) limit of $([\![N_1 \setminus^K N_2]\!])_{K \in \mathbb{N}}$. The last assertion shows that the thorough distance between the over-approximation of the difference presented in section 5.1 and the real difference is 0. This does not imply that they are equal, but from the distance perspective, they are indistinguishable, or infinitesimally close to each other.

References

1 H. Aljazzar and S. Leue. Directed explicit state-space search in the generation of counterexamples for stochastic model checking. *IEEE Trans. Software Eng.*, 36(1):37–60, 2010.

2 R. Alur, T. Feder, and T. A. Henzinger. The benefits of relaxing punctuality. *J. ACM*, 43(1):116–146, 1996.

3 M. E. Andrés, P. R. D'Argenio, and P. van Rossum. Significant diagnostic counterexamples in probabilistic model checking. In *Haifa Verification Conference*, LNCS, pp. 129–148. Springer, 2008.

4 S. S. Bauer, U. Fahrenberg, A. Legay, and C. Thrane. General quantitative specification theories with modalities. In *CSR*, vol. 7999 of *LNCS*, pp. 23–37. Springer, 2012.

5 S. S. Bauer, L. Juhl, K. G. Larsen, A. Legay, and J. Srba. Extending modal transition systems with structured labels. *MSCS*, 22:1–37, 2012.

6 B. Caillaud, B. Delahaye, K. G. Larsen, A. Legay, M. L. Pedersen, and A. Wasowski. Constraint Markov chains. *TCS*, 412(34):4373–4404, 2011.

7 R. Chadha and M. Viswanathan. A counterexample-guided abstraction-refinement framework for Markov decision processes. *ACM Trans. Comput. Log.*, 12(1):1, 2010.

8 L. de Alfaro and T. A. Henzinger. Interface automata. In *FSE*, pp. 109–120. ACM Press, 2001.

9 B. Delahaye, J.-P. Katoen, K. G. Larsen, A. Legay, M. L. Pedersen, F. Sher, and A. Wasowski. Abstract probabilistic automata. In *VMCAI*, vol. 6538 of *LNCS*, pp. 324–339. Springer, 2011.

10 B. Delahaye, J.-P. Katoen, K. G. Larsen, A. Legay, M. L. Pedersen, F. Sher, and A. Wasowski. New results on abstract probabilistic automata. In *ACSD*, pp. 118–127. IEEE, 2011.

11 B. Delahaye, K. G. Larsen, A. Legay, M. L. Pedersen, and A. Wasowski. APAC: A tool for reasoning about abstract probabilistic automata. In *QEST*, pp. 151–152. IEEE Computer Society, 2011.

12 B. Delahaye, K. G. Larsen, A. Legay, M. L. Pedersen, and A. Wąsowski. New results on constraint Markov chains. *Performance Evaluation*, 2011. To appear.

13 J. C. Godskesen, L. Song, and L. Zhang. Probabilistic bisimulation guided abstraction refinement. Personal communication, 2012.

14 T. Han, J.-P. Katoen, and B. Damman. Counterexample generation in probabilistic model checking. *IEEE Trans. Software Eng.*, 35(2):241–257, 2009.

15 H. Hansson and B. Jonsson. A logic for reasoning about time and reliability. *Formal Asp. Comput.*, 6(5):512–535, 1994.

16 H. Hermanns, U. Herzog, and J. Katoen. Process algebra for performance evaluation. *TCS*, 274(1-2):43–87, 2002.

17 H. Hermanns, B. Wachter, and L. Zhang. Probabilistic cegar. In *CAV*, vol. 5123 of *Lecture Notes in Computer Science*, pp. 162–175. Springer, 2008.

18 N. Jansen, E. Ábrahám, J. Katelaan, R. Wimmer, J.-P. Katoen, and B. Becker. Hierarchical counterexamples for discrete-time Markov chains. In *ATVA*, LNCS, pp. 443–452. Springer, 2011.

19 A. Komuravelli, C. S. Pasareanu, and E. M. Clarke. Assume-guarantee abstraction refinement for probabilistic systems. In P. Madhusudan and S. A. Seshia, editors, *CAV*, vol. 7358 of *LNCS*, pp. 310–326. Springer, 2012.

20 K. G. Larsen. Modal specifications. In *AVMS*, vol. 407 of *LNCS*, pp. 232–246, 1989.

21 K. G. Larsen, U. Fahrenberg, and C. Thrane. Metrics for weighted transition systems: Axiomatization and complexity. *TCS*, 412(28):3358–3369, 2011.

22 N. Lynch and M. R. Tuttle. An introduction to Input/Output automata. *CWI-quarterly*, 2(3), 1989.

23 Z. Manna and A. Pnueli. *The Temporal Logic of Reactive and Concurrent Systems*. Springer, 1992.

24 J.-B. Raclet. *Quotient de spécifications pour la réutilisation de composants*. PhD thesis, Université de Rennes I, december 2007. (In French).

25 M. Sassolas, M. Chechik, and S. Uchitel. Exploring inconsistencies between modal transition systems. *Software and System Modeling*, 10(1):117–142, 2011.

26 M. Schmalz, D. Varacca, and H. Völzer. Counterexamples in probabilistic LTL model checking for Markov chains. In *CONCUR*, vol. 5710 of *LNCS*, pp. 587–602, 2009.

27 R. Segala and N. Lynch. Probabilistic simulations for probabilistic processes. In *CONCUR*, vol. 836 of *LNCS*, pp. 481–496. Springer, 1994.

28 C. Thrane, U. Fahrenberg, and K. G. Larsen. Quantitative analysis of weighted transition systems. *JLAP*, 79(7):689–703, 2010.

29 R. Wimmer, B. Braitling, and B. Becker. Counterexample generation for discrete-time Markov chains using bounded model checking. In *VMCAI*, vol. 5403 of *LNCS*, pp. 366–380. Springer, 2009.

A Proof of Theorem 7

We now prove Theorem 7: for all APAs N_1 and N_2 in SVNF, it holds that $d_t(N_1, N_2) \leq d(N_1, N_2)$. For a distribution μ_1 and a constraint φ_2, we denote by

$$\mathsf{RD}(\mu_1, \varphi_2) := \{\delta : \mu_1 \Subset^{\delta} \mu_2 \mid \mu_2 \in Sat(\varphi_2)\}$$

the set of all simulations between μ_1 and distributions satisfying φ_2.

Proof. If $d(N_1, N_2) = 1$, we have nothing to prove. Otherwise, write $N_i = (S_i, A, L_i, AP, V_i, S_0^i)$ for $i = 1, 2$, and let $P_1 = (S_1', A, L_1', AP, V_1', \bar{S}_0^1) \in [\![N_1]\!]$ and $\eta > 0$; we need to expose $P_2 \in [\![N_2]\!]$ for which $d(P_1, P_2) \leq d(N_1, N_2) + \eta$. Note that by the triangle inequality, $d(P_1, N_2) \leq d(P_1, N_1) + d(N_1, N_2) \leq d(N_1, N_2)$. Define $P_2 = (S_2, A, L_2', AP, V_2, S_0^2)$, with L_2' given as follows:

For all $s_1' \in S_1'$, $a \in A$, $\mu_1 \in Dist(S_1')$ for which $L_1'(s_1', a, \mu_1) = \top$ and for all $s_2 \in S_2$, $\varepsilon < 1$ with $\varepsilon := d(s_1', s_2) < 1$: We must have $\varphi_2 \in Dist(S_2)$ such that $L_2(s_2, a, \varphi_2) \neq \bot$ and

$$\inf_{\delta \in \mathsf{RD}(\mu_1, \varphi_2)} \sum_{(t_1', t_2) \in S_1' \times S_2} \mu_1(t_1') \delta(t_1', t_2) d(t_1', t_2) \leq \lambda^{-1} \varepsilon\,,$$

so there must exist a redistribution $\delta \in \mathsf{RD}(\mu_1, \varphi_2)$ for which $\sum_{(t_1', t_2) \in S_1' \times S_2} \mu_1(t_1') \delta(t_1', t_2) d(t_1', t_2) \leq \lambda^{-1}\varepsilon + \lambda^{-1}\eta$. We let $\mu_2(s) = \sum_{s_1' \in S_1} \mu_1(s_1') \delta(s_1', s)$ and set $L_2'(s_2, a, \mu_2) = \top$ in P_2.

Similarly, for all $s_2 \in S_2$, $a \in A$, $\varphi_2 \in C(S_2)$ for which $L_2(s_2, a, \varphi_2) = \top$ and for all $s_1' \in S_1'$ with $\varepsilon := d(s_1', s_2) < 1$: We must have $\mu_1 \in Dist(S_1')$ for which $L_1'(s_1', a, \mu_1) = \top$ and

$$\inf_{\delta \in \mathsf{RD}(\mu_1, \varphi_2)} \sum_{(t_1', t_2) \in S_1' \times S_2} \mu_1(t_1') \delta(t_1', t_2) d(t_1', t_2) \leq \lambda^{-1} \varepsilon\,,$$

so there is $\delta \in \mathsf{RD}(\mu_1, \varphi_2)$ with $\sum_{(t_1', t_2) \in S_1' \times S_2} \mu_1(t_1') \delta(t_1', t_2) d(t_1', t_2) \leq \lambda^{-1}\varepsilon + \lambda^{-1}\eta$. Let again $\mu_2(s) = \sum_{s_1' \in S_1} \mu_1(s_1') \delta(s_1', s)$, and set $L_2'(s_2, a, \mu_2) = \top$ in P_2.

It is easy to see that $P_2 \in [\![N_2]\!]$: by construction of P_2, the identity relation $\{(s_2, s_2) \mid s_2 \in S_2\}$ provides a refinement $P_2 \preceq N_2$. To show that $d(P_1, P_2) \leq d(N_1, N_2) + \eta$, we define a function $d' : S_1' \times S_2 \to [0, 1]$ by $d'(s_1', s_2) = d(s_1', s_2) + \eta$ and show that d' is a pre-fixpoint to (1). Indeed, for s_1' and s_2 compatible, we have

$$\begin{aligned}
d'(s_1', s_2) &= d(s_1', s_2) + \eta \\
&= \max \begin{cases} \displaystyle\max_{a, \mu_1 : L_1'(s_1', a, \mu_1) = \top} \; \min_{\varphi_2 : L_2(s_2, a, \varphi_2) \neq \bot} \lambda D_{P_1, N_2}(\mu_1, \varphi_2, d) + \eta \\ \displaystyle\max_{a, \varphi_2 : L_2(s_2, a, \varphi_2) = \top} \; \min_{\mu_1 : L_1'(s_1', a, \mu_1) = \top} \lambda D_{P_1, N_2}(\mu_1, \varphi_2, d) + \eta \end{cases} \\
&= \max \begin{cases} \displaystyle\max_{a, \mu_1 : L_1'(s_1', a, \mu_1) = \top} \; \min_{\mu_2 : L_2'(s_2, a, \mu_2) = \top} \lambda D_{P_1, P_2}(\mu_1, \mu_2, d) + \eta \\ \displaystyle\max_{a, \mu_2 : L_2'(s_2, a, \mu_2) = \top} \; \min_{\mu_1 : L_1'(s_1', a, \mu_1) = \top} \lambda D_{P_1, P_2}(\mu_1, \mu_2, d) + \eta\,, \end{cases}
\end{aligned}$$

due to the construction of P_2 and the fact that the $\sup_{\mu_1 \in Sat(\mu_1)}$ is trivial in the formula for $D_{P_1, N_2}(\mu_1, \varphi_2, d)$,

$$\geq \max \begin{cases} \displaystyle\max_{a, \mu_1 : L_1'(s_1', a, \mu_1) = \top} \; \min_{\mu_2 : L_2'(s_2, a, \mu_2) = \top} \lambda D_{P_1, P_2}(\mu_1, \mu_2, d') \\ \displaystyle\max_{a, \mu_2 : L_2'(s_2, a, \mu_2) = \top} \; \min_{\mu_1 : L_1'(s_1', a, \mu_1) = \top} \lambda D_{P_1, P_2}(\mu_1, \mu_2, d')\,, \end{cases}$$

where the last inequality is a consequence of

$$\lambda D_{P_1, P_2}(\mu_1, \mu_2, d') = \lambda \sum_{t_1', t_2} \mu_1(t_1') \delta(t_1', t_2)(d(t_1', t_2) + \eta) = \lambda \sum_{t_1', t_2} \mu_1(t_1') \delta(t_1', t_2) d(t_1', t_2) + \lambda\eta.$$

◀

B Proof of Lemma 8

We prove the following lemma, which is equivalent to Lemma 8.

▶ **Lemma 17.** *For all pairs of states (s_1, s_2) in case 3 and for all actions $e \in (B_c(s_1, s_2) \cup B_f(s_1, s_2)) \cap \mathsf{Break}(s_1, s_2)$, there exist constraints φ_1 and φ_2 such that $L_1(s_1, e, \varphi_1) \neq \bot$ and $L_2(s_2, e, \varphi_2) \neq \bot$ and a distribution $\mu_1 \in Sat(\varphi_1)$ such that either*

1. *$\exists s_1' \in S_1$ such that $\mu_1(s_1') > 0$ and $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$, or*
2. *$\mu_1^2 : \left(s_2' \mapsto \sum_{\{s_1' \in S_1 \mid s_2' = \mathsf{succ}_{s_2,e}(s_1')\}} \mu_1(s_1')\right) \notin Sat(\varphi_2)$,*
3. *$\exists s_1' \in S_1, s_2' \in S_2$ such that $\mu_1(s_1') > 0, s_2' = \mathsf{succ}_{s_2,e}(s_1')$ and $\mathsf{Ind}_{\mathcal{R}}(s_1', s_2') < \mathsf{Ind}_{\mathcal{R}}(s_1, s_2)$.*

Proof. Let $\mathcal{R}$ be the maximal refinement relation between N_1 and N_2 and let $(s_1, s_2) \in S_1 \times S_2$ such that (s_1, s_2) is in case 3, i.e. $(s_1, s_2) \notin \mathcal{R}$ and $V_1(s_1) = V_2(s_2)$. Let $e \in A$ such that $e \in (B_c(s_1, s_2) \cup B_f(s_1, s_2)) \cap \mathsf{Break}(s_1, s_2)$.

Since $e \in B_c(s_1, s_2) \cup B_f(s_1, s_2)$, there exists $\varphi_1 \in C(S_1)$ and $\varphi_2 \in C(S_2)$ such that either $L_2(s_2, e, \varphi_2) = \top$ and $L_1(s_1, e, \varphi_1) = \top$ or $L_2(s_2, e, \varphi_2) =?$ and $L_1(s_1, e, \varphi_1) \neq \bot$. As a consequence, since $e \in \mathsf{Break}(s_1, s_2)$, we have that

$$\exists \mu_1 \in Sat(\varphi_1), \forall \mu_2 \in Sat(\varphi_2), \mu_1 \not\in_{\mathcal{R}_k} \mu_2. \tag{2}$$

Let K be the smallest index such that $\mathcal{R}_K = \mathcal{R}$. By construction, we know that $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) = k < K$, i.e. $(s_1, s_2) \in \mathcal{R}_k$ and $(s_1, s_2) \notin \mathcal{R}_{k+1}$. Consider the distribution μ_1 given by (2) above. We have that $\forall \mu_2 \in Sat(\varphi_2), \forall$ corresp. $\delta, \mu_1 \not\in^{\delta}_{\mathcal{R}_k} \mu_2$. Consider the function δ such that $\delta(s_1', s_2') = 1$ if $s_2' = \mathsf{succ}_{s_2,e}(s_1')$ and 0 otherwise. There are several cases.

- If there exists $s_1' \in S_1$ such that $\mu_1(s_1') > 0$ and $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$, then the lemma is proven.
- Else, δ is a correspondence function. Since $\forall \mu_2 \in Sat(\varphi_2), \mu_1 \not\in_{\mathcal{R}_k} \mu_2$, we know that either (1) $\mu_2 : s_2' \mapsto \sum_{s_1' \in S_1} \mu_1(s_1')\delta(s_1', s_2')$ does not satisfy φ_2, or (2) there exists s_1' and s_2' such that $\mu_1(s_1') > 0$, $\delta(s_1', s_2') > 0$ and $(s_1', s_2') \notin \mathcal{R}_k$.
  1. Assume that $\mu_2 : s_2' \mapsto \sum_{s_1' \in S_1} \mu_1(s_1')\delta(s_1', s_2')$ does not satisfy φ_2. Remark that the function μ_1^2 from Lemma 17 is equal to μ_2 defined above. As a consequence, $\mu_1^2 \notin \varphi_2$.
  2. Otherwise, assume that there exists s_1' and s_2' such that $\mu_1(s_1') > 0$, $\delta(s_1', s_2') > 0$ and $(s_1', s_2') \notin \mathcal{R}_k$. Since $(s_1', s_2') \notin \mathcal{R}_k$, we have that $\mathsf{Ind}_{\mathcal{R}}(s_1', s_2') < k$. As a consequence, there exists $s_1' \in S_1, s_2' \in S_2$ such that $\mu_1(s_1') > 0, s_2' = \mathsf{succ}_{s_2,e}(s_1')$ and $\mathsf{Ind}_{\mathcal{R}}(s_1', s_2') < \mathsf{Ind}_{\mathcal{R}}(s_1, s_2)$.

◀

C Proof of Theorem 10

Let $N_1 = (S_1, A, L_1, AP, V_1, \{s_0^1\})$ and $N_2 = (S_2, A, L_2, AP, V_2, \{s_0^2\})$ be deterministic APAs in SVNF such that $N_1 \not\preceq N_2$. Let $P = (S, A, L, AP, V, s_0)$ be the counterexample defined as in Section 4. We prove that $P \models N_1$ and $P \not\models N_2$.

Proof. $\mathbf{P \models N_1}$**.** Consider the relation $\mathcal{R}_s \subseteq S \times S_1$ such that $(s_1, s_2)\, \mathcal{R}_s\, s_1'$ iff $s_1 = s_1'$. We prove that $\mathcal{R}_s$ is a satisfaction relation. Let $t = (s_1, s_2) \in S$ and consider $(t, s_1) \in \mathcal{R}_s$.

- By construction, we have $V(s_1, s_2) \subseteq V_1(s_1)$.

- Let $a \in A$ and $\varphi_1 \in C(S_1$ such that $L_1(s_1, a, \varphi_1) = \top$. There are several cases.
  - If (s_1, s_2) in case 1 or 2 or $s_2 = \bot$, then by construction there exists $\mu_1^\bot \in Dist(S)$ such that $L((s_1, s_2), a, \mu_1^\bot) = \top$. By construction, we have that there exists $\mu_1 \in Sat(\varphi_1)$ such that $\mu_1^\bot \Subset_{\mathcal{R}_s} \mu_1$.
  - Else, (s_1, s_2) is in case 3 and $B(s_1, s_2) \neq \emptyset$. If $a \notin B(s_1, s_2)$, the result follows as above. Else, either $a \in B_a(s_1, s_2) \cup B_b(s_1, s_2)$ and the result follows again by construction, or $a \in B_c(s_1, s_2) \cup B_f(s_1, s_2)$. In this case, there exists a distribution $\widehat{\mu_1} \in Dist(S)$ such that $L((s_1, s_2), a, \widehat{\mu_1}) = \top$. By construction, $\widehat{\mu_1}$ is defined as follows:

$$\widehat{\mu_1}(s_1', s_2') = \begin{cases} \mu_1(s_1) & \text{if } s_2' = \mathsf{succ}_{s_2,e}(s_1') \\ & \quad \text{or } \mathsf{succ}_{s_2,e}(s_1') = \emptyset \text{ and } s_2' = \bot \\ 0 & \text{otherwise} \end{cases},$$

  where μ_1 is either the distribution given by Lemma 17 if $a \in \mathsf{Break}(s_1, s_2)$ or an arbitrary distribution in $Sat(\varphi_1)$. In both cases, $\mu_1 \in Sat(\varphi_1)$. Consider the function $\delta : S \times S_1 \to [0, 1]$ such that $\delta((s_1', s_2'), s_1'') = 1$ if $s_1' = s_1''$ and 0 otherwise. Using standard techniques, on can verify that δ is a correspondence function and that $\widehat{\mu_1} \Subset_{\mathcal{R}_s} \mu_1$.
- Let $a \in A$ and $\mu \in Dist(S)$ such that $L((s_1, s_2), a, \mu) = \top$. By construction of P, there must exists $\varphi_1 \in C(S_1)$ such that $L_1(s_1, a, \varphi_1) \neq \bot$ and μ is either of the form $\mu_1^\bot$ or $\widehat{\mu_1}$ for some $\mu_1 \in Sat(\varphi_1)$. As above, we can prove that in all cases, $\mu \Subset_{\mathcal{R}_s} \mu_1$.

  Finally $\mathcal{R}_s$ is a satisfaction relation. Moreover, we have $((s_0^1, s_0^2), s_0^1) \in \mathcal{R}_s$, thus $P \models N_1$.

$\mathbf{P \not\models N_2.}$ Let $\mathcal{R}_s \subseteq S \times S_2$ be the maximal satisfaction relation between P and N_2, and assume that $\mathcal{R}_s$ is not empty. Let $\mathcal{R} \subseteq S_1 \times S_2$ be the maximal refinement relation between N_1 and N_2 and let K be the smallest index such that $\mathcal{R}_K = \mathcal{R}$. We prove that for all $(s_1, s_2) \in S_1 \times S_2$, if $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) < K$, then $((s_1, s_2), s_2) \notin \mathcal{R}_s$. The proof is done by induction on $k = \mathsf{Ind}_{\mathcal{R}}(s_1, s_2)$. Let $(s_1, s_2) \in S_1 \times S_2$.

- **Base case.** If $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) = 0$, then there are several cases.
  - If (s_1, s_2) in case 2, i.e. $V_1(s_1) \neq V_2(s_2)$. In this case, we know that $V((s_1, s_2)) \in V_1(s_1)$. Thus, by SVNF of N_1 and N_2, we have that $V((s_1, s_2)) \notin V_2(s_2)$ and $((s_1, s_2), s_2) \notin \mathcal{R}_s$.
  - Else, if (s_1, s_2) in cases $3.a$ or $3.b$, then there exists $a \in A$ and $\mu_1^\bot \in Dist(S)$ such that $L((s_1, s_2), a, \mu_1^\bot) = \top$ and $\forall \varphi_2 \in C(S_2)$, we have $L_2(s_2, a, \varphi_2) = \bot$. As a consequence, $((s_1, s_2), s_2) \notin \mathcal{R}_s$.
  - Else, if (s_1, s_2) in cases $3.d$ or $3.d$, then there exists $a \in A$ and $\varphi_2 \in C(S_2)$ such that $L_2(s_2, a, \varphi_2) = \top$ and for all $\mu \in Dist(S)$, we have $L((s_1, s_2), a, \mu) = \bot$. As a consequence, $((s_1, s_2), s_2) \notin \mathcal{R}_s$.
  - Finally, if (s_1, s_2) in cases $3.c$ or $3.f$, there exists $e \in (B_c(s_1, s_2) \cup B_f(s_1, s_2)) \cap \mathsf{Break}(s_1, s_2)$. By Lemma 17, there exists constraints φ_1 and φ_2 such that $L_1(s_1, e, \varphi_1) \neq \bot$ and $L_2(s_2, e, \varphi_2) \neq \bot$ and a distribution $\mu_1 \in Sat(\varphi_1)$ such that either

    (I) $\exists s_1' \in S_1$ such that $\mu_1(s_1') > 0$ and $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$, or

    (II) $\mu_1^2 : \left(s_2' \mapsto \sum_{\{s_1' \in S_1 \mid s_2' = \mathsf{succ}_{s_2,e}(s_1')\}} \mu_1(s_1')\right) \notin Sat(\varphi_2)$,

    (III) $\exists s_1' \in S_1, s_2' \in S_2$ such that $\mu_1(s_1') > 0, s_2' = \mathsf{succ}_{s_2,e}(s_1')$ and $\mathsf{Ind}_{\mathcal{R}}(s_1', s_2') < \mathsf{Ind}_{\mathcal{R}}(s_1, s_2)$.

    By construction, we have that $L((s_1, s_2), e, \widehat{\mu_1}) = \top$ for μ_1 given above. Since $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) = 0$, case (III) above is not possible. From cases (I) and (II), we can deduce that for all $\mu_2 \in Sat(\varphi_2)$, we have $\widehat{\mu_1} \not\Subset_{\mathcal{R}_s} \mu_2$. Moreover, by determinism of N_2, φ_2 is the only constraint such that $L_2(s_2, e, \varphi_2) \neq \bot$. As a consequence, $((s_1, s_2), s_2) \notin \mathcal{R}_s$.

- **Inductive step.** Let $0 < k < K$ and assume that for all $k' < k$ and for all $(s_1', s_2) \in \S_1 \times S_2$, if $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) = k'$, then $((s_1, s_2), s_2) \notin \mathcal{R}_s$. Assume that $\mathsf{Ind}_{\mathcal{R}}(s_1, s_2) = k$. There are two cases.
  - If (s_1, s_2) in cases 2, $3.a$, $3.b$, $3.d$ or $3.d$, the same reasoning applies as for the base case. We thus deduce that $((s_1, s_2), s_2) \notin \mathcal{R}_s$.
  - Otherwise, if (s_1, s_2) in cases $3.c$ or $3.f$, then, as above, there exists $e \in (B_c(s_1, s_2) \cup B_f(s_1, s_2)) \cap \mathsf{Break}(s_1, s_2)$. By Lemma 17, there exists constraints φ_1 and φ_2 such that $L_1(s_1, e, \varphi_1) \neq \bot$ and $L_2(s_2, e, \varphi_2) \neq \bot$ and a distribution $\mu_1 \in Sat(\varphi_1)$ such that either

    (I) $\exists s_1' \in S_1$ such that $\mu_1(s_1') > 0$ and $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$, or

    (II) $\mu_1^2 : \left(s_2' \mapsto \sum_{\{s_1' \in S_1 \mid s_2' = \mathsf{succ}_{s_2,e}(s_1')\}} \mu_1(s_1')\right) \notin Sat(\varphi_2)$,

    (III) $\exists s_1' \in S_1, s_2' \in S_2$ such that $\mu_1(s_1') > 0, s_2' = \mathsf{succ}_{s_2,e}(s_1')$ and $\mathsf{Ind}_{\mathcal{R}}(s_1', s_2') < \mathsf{Ind}_{\mathcal{R}}(s_1, s_2)$.

    By construction, we have that $L((s_1, s_2), e, \widehat{\mu_1}) = \top$ for μ_1 given above. As above, if cases (I) or (II) apply, then we can deduce that $((s_1, s_2), s_2) \notin \mathcal{R}_s$. If case (III) applies, then there exists $(s_1', s_2') \in S$ such that $\widehat{\mu_1}(s_1', s_2') > 0$, $s_2' = \mathsf{succ}_{s_2,e}(s_1')$ and $\mathsf{Ind}_{\mathcal{R}}(s_1', s_2') < \mathsf{Ind}_{\mathcal{R}}(s_1, s_2)$. Since $s_2' = \mathsf{succ}_{s_2,e}(s_1')$, then, by determinism of N_2, all correspondence functions δ will be such that $\delta((s_1', s_2'), s_2') = 1$. However, we have that $\mathsf{Ind}_{\mathcal{R}}(s_1', s_2') < k$, thus by induction $((s_1', s_2'), s_2') \notin \mathcal{R}_s$. As a consequence, we have that for all $\mu_2 \in Sat(\varphi_2)$, we have $\widehat{\mu_1} \notin_{\mathcal{R}_s} \mu_2$. We can thus deduce that $((s_1, s_2), s_2) \notin \mathcal{R}_s$.

  Finally, we know that $\mathsf{Ind}_{\mathcal{R}}(s_0^1, s_0^2) < k$. As a consequence, we have $((s_0^1, s_0^2), s_0^2) \notin \mathcal{R}_s$ and thus $P \not\models N_2$.

◀

D Proof of Theorem 12

For all deterministic APAs N_1 and N_2 in SVNF such that $N_1 \not\preceq N_2$, we have $[\![N_1]\!] \setminus [\![N_2]\!] \subseteq [\![N_1 \setminus^* N_2]\!]$.

Proof. Let $N_1 = (S_1, A, L_1, AP, V_1, \{s_0^1\})$ and $N_2 = (S_2, A, L_2, AP, V_2, \{s_0^2\})$ be deterministic APAs in single valuation normal form such that $N_1 \not\preceq N_2$. Let $\mathcal{R}$ be the maximal weak refinement relation between N_1 and N_2. Let $P = (S_P, A, L_P, AP, V_P, s_0^P)$ be a PA such that $P \models N_1$ and $P \not\models N_2$. We prove that $P \models N_1 \setminus^* N_2$. Let $\mathcal{R}_1 \subseteq S_P \times S_1$ be the relation witnessing $P \models N_1$ and let $\mathcal{R}_2$ be the maximal satisfaction relation in $S_P \times S_2$. By construction, $(s_0^P, s_2) \notin \mathcal{R}_2$.
If $V_1(s_0^1) \neq V_2(s_0^2)$, then by construction $N_1 \setminus^* N_2 = N_1$ and thus $P \models N_1 \setminus^* N_2$.
Else, we have (s_0^1, s_0^2) in case 3, thus $N_1 \setminus^* N_2 = (S, A, L, AP, V, S_0)$ is defined as in Section 5.1. By construction, we also have (s_0^P, s_0^2) in case 3, thus there must exist $f \in B(s_0^P, s_0^2)$. Remark that by construction, we must have $B(s_0^P, s_0^2) \subseteq B(s_0^1, s_0^2)$. We will prove that $P \models N_1 \setminus^* N_2$.

Define the following relation $\mathcal{R}^{\setminus} \subseteq S_P \times S$:

$$p\,\mathcal{R}^{\setminus}(s_1, s_2, e) \iff \begin{cases} & (p\,\mathcal{R}_1\,s_1) \text{ and } (s_2 = \bot) \text{ and } (e = \varepsilon) \\ \text{or} & (p\,\mathcal{R}_1\,s_1) \text{ and } (p, s_2) \text{ in case 1 or 2 and and } (e = \varepsilon) \\ \text{or} & (p\,\mathcal{R}_1\,s_1) \text{ and } (p, s_2) \text{ in case 3 and } (e \in B(p, s_2)) \end{cases}$$

We now prove that $\mathcal{R}^{\setminus}$ is a satisfaction relation. Let $(p, (s_1, s_2, e)) \in \mathcal{R}^{\setminus}$.
If $s_2 = \bot$ or $e = \varepsilon$, then since $p\,\mathcal{R}_1\,s_1$, $\mathcal{R}^{\setminus}$ satisfies the axioms of a satisfaction relation by construction.
Else we have $s_2 \in S_2$ and $e \neq \varepsilon$, thus, by definition of $\mathcal{R}^{\setminus}$, we know that (p, s_2) is in case 3.

- By construction, we have $V_P(p) \in V_1(s_1) = V((s_1, s_2, e))$.
- Let $a \in A$ and $\mu_P \in Dist(S_P)$ such that $L_P(p, a, \mu_P) = \top$. There are several cases.
  - If $a \neq e$, then since $p\,\mathcal{R}_1\,s_1$, there exists $\varphi_1 \in C(S_1)$ such that $L_1(s_1, a, \varphi_1) \neq \bot$ and there exists $\mu_1 \in Sat(\varphi_1)$ such that $\mu_P \Subset_{\mathcal{R}^\setminus} \mu_1$. By construction, we have $L((s_1, s_2, e), a, \varphi_1^{\perp}) \neq \bot$ and there obviously exists $\mu \in Sat(\varphi_1^{\perp})$ such that $\mu_P \Subset_{\mathcal{R}^\setminus} \mu$.
  - If $a = e \in B_a(p, s_2)$, then, as above, there exists $\varphi \in C(S)$ such that $L((s_1, s_2, e), a, \varphi) \neq \bot$ and there exists $\mu \in Sat(\varphi)$ such that $\mu_P \Subset_{\mathcal{R}^\setminus} \mu$. Remark that $B_a(s_1, s_2) \subseteq B_a(p, s_2) \subseteq B_a(s_1, s_2) \cup B_b(s_1, s_2)$.
  - Else, we necessarily have $a = e \in B_c(p, s_2) \cup B_f(p, s_2)$. Remark that, by construction, $B_c(p, s_2) \subseteq B_c(s_1, s_2)$ and $B_f(p, s_2) \subseteq B_f(s_1, s_2)$. Since $p\,\mathcal{R}_1\,s_1$, there exists $\varphi_1 \in C(S_1)$ such that $L_1(s_1, e, \varphi_1) \neq \bot$ and there exists $\mu_1 \in Sat(\varphi_1)$ and a correspondence function $\delta_1 : S_P \to (S_1 \to [0, 1])$ such that $\mu_P \Subset_{\mathcal{R}_1}^{\delta_1} \mu_1$.

    Moreover, by construction of $N_1 \setminus^* N_2$, we know that the constraint φ_{12}^B such that $\mu \in Sat(\varphi_{12}^B)$ iff. (1) for all $(s_1', s_2', c) \in S$, we have $\mu(s_1', s_2', c) > 0 \Rightarrow s_2' = \bot$ if $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$ and $s_2' = \mathsf{succ}_{s_2,e}(s_1')$ otherwise, and $c \in B(s_1', s_2') \cup \{\varepsilon\}$, (2) the distribution $\mu_1 : s_1' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_2' \in S_2 \cup \{\bot\}} \mu(s_1', s_2', c)$ satisfies φ_1, and (3) either (b) the distribution $\mu_2 : s_2' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_1' \in S_1} \mu(s_1', s_2', c)$ does not satisfy φ_2, or (c) there exists $s_1' \in S_1$, $s_2' \in S_2$ and $c \neq \varepsilon$ such that $\mu(s_1', s_2', c) > 0$ is such that $L((s_1, s_2, e), e, \varphi_{12}^B) = \top$.

    We now prove that there exists $\mu \in Sat(\varphi_{12}^B)$ such that $\mu_P \Subset_{\mathcal{R}^\setminus} \mu$. Consider the function $\delta^\setminus : S_P \to (S \to [0, 1])$ defined as follows: Let $p' \in S_P$ such that $\mu_P(p') > 0$ and let $s_1' = \mathsf{succ}_{s_1,e}(p')$, which exists by $\mathcal{R}_1$.
    - If $\mathsf{succ}_{s_2,e}(p') = \emptyset$, then $\delta^\setminus(p')(s_1', \bot, \varepsilon) = 1$.
    - Else, let $s_2' = \mathsf{succ}_{s_2,e}(p')$. Then,
      - if $(p', s_2') \in \mathcal{R}_2$, then $\delta^\setminus(p')(s_1', s_2', \varepsilon) = 1$.
      - Else, (p', s_2') is in case 3 and $B(p', s_2') \neq \emptyset$. In this case, let $c \in B(p', s_2')$ and define $\delta^\setminus(p', (s_1', s_2', c)) = 1$. For all other $c' \in B(p', s_2')$, define $\delta^\setminus(p', (s_1', s_2', c)) = 0$.

    Remark that for all $p' \in S_P$ such that $\mu_P(p') > 0$, there exists a unique $s' \in S'$ such that $\delta^\setminus(p')(s') = 1$. Thus $\delta^\setminus$ is a correspondence function.

    We now prove that $\mu = \mu_P \delta^\setminus \in Sat(\varphi_{12}^B)$.

    1. Let $(s_1', s_2', c) \in S$ such that $\mu(s_1', s_2', c) > 0$. By construction, there exists $p' \in S_P$ such that $\mu_P(p') > 0$ and $\delta^\setminus(p')(s_1', s_2', c) > 0$. Moreover, $c \in B(s_1', s_2') \cup \{\varepsilon\}$, and $s_2' = \bot$ if $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$ and $s_2' = \mathsf{succ}_{s_2,e}(s_1')$ otherwise.
    2. Consider the distribution $\mu_1' : s_1' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_2' \in S_2 \cup \{\bot\}} \mu(s_1', s_2', c)$. By determinism (See Lemma 28 in [6]), we have that $\delta_1(p')(s_1') = 1 \iff s_1' = (succ)_{s_1,e}(p')$. As a consequence, we have that $\mu_1' = \mu_1 \in Sat(\varphi_1)$.
    3. Assume that for all $p' \in S_P$ such that $\mu_P(p') > 0$, we have $\mathsf{succ}_{s_2,e}(p') \neq \emptyset$ (the other case being trivial). Consider the distribution $\mu_2 : s_2' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_1' \in S_1} \mu(s_1', s_2', c)$ and let $\delta_2 : S_P \to (S_2 \to [0, 1])$ be such that $\delta_2(p')(s_2') = 1 \iff s_2' = \mathsf{succ}_{s_2,e}(p')$. By construction, δ_2 is a correspondence function and $\mu_2 = \mu_P \delta_2$. Since $e \in B_c(p, s_2) \cup B_f(p, s_2)$, we have that $\mu_P \not\Subset_{\mathcal{R}_2} \mu_2$. If $\mu_2 \notin Sat(\varphi_2)$, then we have $\mu \in Sat(\varphi_{12}^B)$. Else, there must exist $p' \in S_P$ and $s_2' \in S_2$ such that $\mu_P(p') > 0$, $\delta_2(p')(s_2') > 0$ and $(p', s_2') \notin \mathcal{R}_2$. As a consequence, (p', s_2') is in case 3 and there exists $c \neq \varepsilon$ such that $\delta^\setminus(p')(s_1', s_2', c) > 0$, thus $\mu(s_1', s_2', c) > 0$. As a consequence, $\mu \in Sat(\varphi_{12}^B)$.

    We thus conclude that there exists $\mu \in Sat(\varphi_{12}^B)$ such that $\mu_P \Subset_{\mathcal{R}^\setminus} \mu$.

  Finally, in all cases, there exists $\varphi \in C(S)$ such that $L((s_1, s_2, e), a, \varphi) \neq \bot$ and there exists $\mu \in Sat(\varphi)$ such that $\mu_P \Subset_{\mathcal{R}^\setminus} \mu$.
- Let $a \in A$ and $\varphi \in C(S)$ such that $L((s_1, s_2, e), a, \varphi) = \top$. As above, there are several cases.

- If $a \neq e$, then, by construction of $N_1 \setminus^* N_2$, there must exists $\varphi_1 \in C(S_1)$ such that $L_1(s_1, a, \varphi_1) = \top$. The rest of the proof is then as above.
- If $a = e \in B_a(p, s_2)$, then there exists $\mu_P \in \mathit{Dist}(S_P)$ such that $L_P(p, e, \mu_P) = \top$. The rest of the proof is then as above. Recall that $B_a(s_1, s_2) \subseteq B_a(p, s_2) \subseteq B_a(s_1, s_2) \cup B_b(s_1, s_2)$.
- Else, we necessarily have $a = e \in B_c(p, s_2) \cup B_f(p, s_2)$. Recall that, by construction, $B_c(p, s_2) \subseteq B_c(s_1, s_2)$ and $B_f(p, s_2) \subseteq B_f(s_1, s_2)$. Thus, there exists $\mu_P \in \mathit{Dist}(S_P)$ and $\varphi_2 \in C(S_2)$ such that $L_2(s_2, e, \varphi_2) \neq \bot$ and $\forall \mu_2 \in \mathit{Sat}(\varphi_2), \mu_P \not\Subset_{\mathcal{R}_2} \mu_2$. Since $e \in B_c(s_1, s_2) \cup B_f(s_1, s_2)$, there also exist $\varphi_1 \in C(S_1)$ such that $L_1(s_1, e, \varphi_1) \neq \bot$. By determinism, φ_1 and φ_2 are unique. The rest of the proof follows as above.

Thus, in all cases, there exists $\mu_P \in \mathit{Dist}(S_P)$ such that $L_P(p, a, \mu_P) = \top$ and there exists $\mu \in \mathit{Sat}(\varphi)$ such that $\mu_P \Subset_{\mathcal{R}^\setminus} \mu$.

Finally, $\mathcal{R}^\setminus$ is a satisfaction relation. Moreover, we have $s_0^P \, \mathcal{R}_1 \, s_0^1$, (s_0^P, s_0^2) in case 3 and $f \in B(s_0^P, s_0^2)$ by construction, thus $s_0^P \, \mathcal{R}^\setminus (s_0^1, s_0^2, f) \in S_0$.

We thus conclude that $P \models N_1 \setminus^* N_2$.

◀

E Appendix to Definition 14

Table 3 reports the full structure of the transition relation of the under-approximation construction given in Definition 14.

Recall the definition of constraints $\varphi^\bot$ and $\varphi_{12}^{B,k}$:

Given a constraint $\varphi \in C(S_1)$, the constraint $\varphi^\bot \in C(S)$ is defined as follows:

$$\mu \in \mathit{Sat}(\varphi^\bot) \iff \begin{cases} \forall s_1 \in S_1, \forall s_2 \neq \bot, \forall b \neq \varepsilon, \forall k \neq 1, \mu(s_1, s_2, b, k) = 0 \\ (\mu \downarrow_1 \colon s_1 \mapsto \mu(s_1, \bot, \varepsilon, 1)) \in \mathit{Sat}(\varphi) \end{cases}$$

Given a state $(s_1, s_2, e, k) \in S$ with $s_2 \neq \bot$ and $e \neq \varepsilon$ and two constraints $\varphi_1 \in C(S_1)$ and $\varphi_2 \in C(S_2)$ such that $L_1(s_1, e, \varphi_1) \neq \bot$ and $L_2(s_2, e, \varphi_2) \neq \bot$, the constraint $\varphi_{12}^{B,k} \in C(S)$ is defined as follows: $\mu \in \mathit{Sat}(\varphi_{12}^{B,k})$ iff

1. for all $(s_1', s_2', c, k') \in S$, if $\mu(s_1', s_2', c, k') > 0$, then $c \in B(s_1', s_2') \cup \{\varepsilon\}$ and either $\mathsf{succ}_{s_2,e}(s_1') = \emptyset$, $s_2' = \bot$ and $k' = 1$, or $s_2' = \mathsf{succ}_{s_2,e}(s_1')$,
2. the distribution $\mu_1 : s_1' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_2' \in S_2 \cup \{\bot\}, k' \geq 1} \mu(s_1', s_2', c, k')$ satisfies φ_1, and
3. either
   a. there exists $(s_1', \bot, c, 1)$ such that $\mu(s_1', \bot, c, 1) > 0$, or
   b. the distribution $\mu_2 : s_2' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_1' \in S_1, k' \geq 1} \mu(s_1', s_2', c, k')$ does not satisfy φ_2, or
   c. $k \neq 1$ and there exists $s_1' \in S_1$, $s_2' \in S_2$, $c \neq \varepsilon$ and $k' < k$ such that $\mu(s_1', s_2', c, k') > 0$.

F Proof of Theorem 15

For all deterministic APAs N_1 and N_2 in SVNF such that $N_1 \not\preceq N_2$, we have that

1. for all $K \in \mathbb{N}$, $[\![N_1 \setminus^K N_2]\!] \subseteq [\![N_1]\!] \setminus [\![N_2]\!]$, and
2. for all PA $P \in [\![N_1]\!] \setminus [\![N_2]\!]$, there exists $K \in \mathbb{N}$ such that $P \in [\![N_1 \setminus^K N_2]\!]$.

Proof. For the first claim, consider the relation $\mathcal{R} \subseteq (S_1 \times (S_2 \cup \{\bot\}) \times (A \cup \{\varepsilon\}) \times \{1, \ldots, K\}) \times (S_1 \times (S_2 \cup \{\bot\}) \times (A \cup \{\varepsilon\}) \times \{1, \ldots, K{+}1\})$ such that $\mathcal{R} = \{((s_0^1, s_0^2, e, K), (s_0^1, s_0^2, e, K{+}1)) \mid e \in B(s_0^1, s_0^2)\} \cup \mathcal{R}_{\mathsf{id}}$, where $\mathcal{R}_{\mathsf{id}}$ denotes the identity relation. One can verify that, by construction, $\mathcal{R}$ is a refinement relation witnessing $N_1 \setminus^K N_2 \preceq N_1 \setminus^{K+1} N_2$.

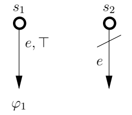
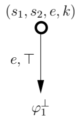
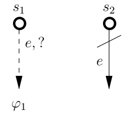
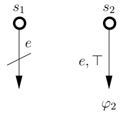
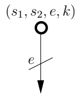
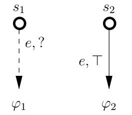
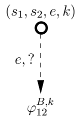
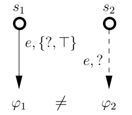
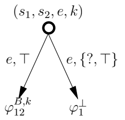
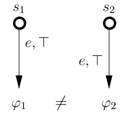
$e \in$	N_1, N_2	$N_1 \setminus^K N_2$	Formal Definition of L
$B_a(s_1, s_2)$			For all $a \neq e \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$, let $L((s_1, s_2, e, k), a, \varphi^\perp) = L_1(s_1, a, \varphi)$. In addition, let $L((s_1, s_2, e, k), e, \varphi_1^\perp) = \top$. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e, k), b, \varphi) = \perp$.
$B_b(s_1, s_2)$			
$B_d(s_1, s_2)$			For all $a \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$, let $L((s_1, s_2, e, k), a, \varphi^\perp) = L_1(s_1, a, \varphi)$. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e, k), b, \varphi) = \perp$.
$B_e(s_1, s_2)$	 s_1 s_2 $e, ?$ $e, \top$ φ_1 φ_2		For all $a \neq e \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$, let $L((s_1, s_2, e, k), a, \varphi^\perp) = L_1(s_1, a, \varphi)$. In addition, let $L((s_1, s_2, e, k), e, \varphi_{12}^{B,k}) = ?$. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e, k), b, \varphi) = \perp$.
$B_c(s_1, s_2)$	 s_1 s_2 $e, \{?, \top\}$ $e, ?$ φ_1 $\neq$ φ_2	 (s_1, s_2, e, k) $e, \top$ $e, \{?, \top\}$ $\varphi_{12}^{B,k}$ $\varphi_1^\perp$	For all $a \in A$ and $\varphi \in C(S_1)$ such that $L_1(s_1, a, \varphi) \neq \perp$ (including e and φ_1), let $L((s_1, s_2, e, k), a, \varphi^\perp) = L_1(s_1, a, \varphi)$. In addition, let $L((s_1, s_2, e, k), e, \varphi_{12}^{B,k}) = \top$. For all other $b \in A$ and $\varphi \in C(S)$, let $L((s_1, s_2, e, k), b, \varphi) = \perp$.
$B_f(s_1, s_2)$	 s_1 s_2 $e, \top$ $e, \top$ φ_1 $\neq$ φ_2		

Table 3 Definition of the transition function L in $N_1 \setminus^K N_2$.

Let $N_1 = (S_1, A, L_1, AP, V_1, \{s_0^1\})$ and $N_2 = (S_2, A, L_2, AP, V_2, \{s_0^2\})$ be deterministic APAs in single valuation normal form such that $N_1 \not\preceq N_2$. Let $\mathcal{R}$ be the maximal weak refinement relation between N_1 and N_2.

1. We first prove that for all $K \in \mathbb{N}$, $[\![N_1 \setminus^K N_2]\!] \subseteq [\![N_1]\!] \setminus [\![N_2]\!]$.
If $V_1(s_0^1) \neq V_2(s_0^2)$, then for all $K \in \mathbb{N}$, we have $N_1 \setminus^K N_2 = N_1$ and the result holds.
Otherwise, assume that (s_0^1, s_0^2) is in case 3 and let $K \in \mathbb{N}$. We have $N_1 \setminus^K N_2 = (S, A, L, AP, V, S_0^K)$ defined as in Section 5.2. Let $P = (S_P, A, L_P, AP, V_P, s_0^P)$ be a PA such that $P \models N_1 \setminus^K N_2$. Let $\mathcal{R}^\setminus \subseteq S_P \times S$ be the associated satisfaction relation and let $f \in B(s_0^1, s_0^2)$ be such that $s_0^P \, \mathcal{R}^\setminus (s_0^1, s_0^2, f, K)$. We show that $P \models N_1$ and $P \not\models N_2$.

We start by proving that $P \models N_1$. Consider the relation $\mathcal{R}_1 \subseteq S_P \times S_1$ such that $p \, \mathcal{R}_1 \, s_1 \iff \exists s_2 \in (S_2 \cup \{\perp\}), \exists e \in (A \cup \{\varepsilon\}), \exists n \leq K$ s.t. $p \, \mathcal{R}^\setminus (s_1, s_2, e, n)$. We prove that $\mathcal{R}_1$ is a satisfaction relation. Let p, s_1, s_2, e, n such that $p \, \mathcal{R}_1 \, s_1$ and $p \, \mathcal{R}^\setminus (s_1, s_2, e, n)$.

- By construction, we have $V_P(p) \in V((s_1, s_2, e, n)) = V_1(s_1)$.
- Let $a \in A$ and $\mu_P \in Dist(S_P)$ be such that $L_P(p, a, \mu_P) = \top$. By $\mathcal{R}^\setminus$, there exists $\varphi \in C(S)$ such that $L((s_1, s_2, e, n), a, \varphi) \neq \perp$ and there exists $\mu \in Sat(\varphi)$ such that $\mu_P \Subset_{\mathcal{R}^\setminus} \mu$.

  If $s_2 = \bot$ or $e = \varepsilon$ or $a \neq e$, then by construction of $N_1 \setminus^K N_2$, there exists $\varphi_1 \in C(S_1)$ such that $\varphi = \varphi_1^{\bot}$ and $L_1(s_1, a, \varphi_1) \neq \bot$. As a consequence, the distribution $\mu\downarrow_1$: $s_1' \mapsto \mu(s_1', \bot, \varepsilon, 1)$ is in $Sat(\varphi_1)$ and it follows that $\mu_P \Subset_{\mathcal{R}_1} \mu\downarrow_1$.
  Otherwise, assume that $s_2 \in S_2$, $e \in A$ and $a = e$. There are several cases.

  - If $e \in B_a(s_1, s_2) \cup B_b(s_1, s_2)$, then by construction of $N_1 \setminus^K N_2$, there exists $\varphi_1 \in C(S_1)$ such that $L_1(s_1, e, \varphi_1) \neq \bot$ and $\varphi = \varphi_1^{\bot}$. As above, we thus have $\mu_P \Subset_{\mathcal{R}_1} \mu\downarrow_1$.
  - Else, if $e \in B_e(s_1, s_2)$, then there exists $\varphi_1 \in C(S_1)$ and $\varphi_2 \in C(S_2)$ such that $L_1(s_1, e, \varphi_1) = ?$ and $L_2(s_2, e, \varphi_2) = \top$. Moreover, φ is of the form φ_{12}^{B}, and $\mu' \in Sat(\varphi_{12}^{B})$implies that the distribution $\mu_1' : s_1' \mapsto \sum_{c\in A\cup\{\varepsilon\}, s_2'\in S_2\cup\{\bot\}, k'\geq 1} \mu(s_1', s_2', c, k')$ satisfies φ_1. Thus, the distribution $\mu_1 : s_1' \mapsto \sum_{c\in A\cup\{\varepsilon\}, s_2'\in S_2\cup\{\bot\}, k'\geq 1} \mu(s_1', s_2', c, k')$ satisfies φ_1. Let $\delta_1 : S_P \to (S_1 \to [0,1])$ be such that $\delta_1(p')(s_1') = 1$ if $\mu_P(p') > 0$ and $s_1' = \mathsf{succ}_{s_1,e}(p')$ and 0 otherwise. By construction, δ_1 is a correspondence function and we have $\mu_P\delta_1 = \mu_1$.
    Thus there exists $\mu_1 \in Sat(\varphi_1)$ such that $\mu_P \Subset_{\mathcal{R}_1} \mu_1$.
  - Finally, if $e \in B_c(s_1, s_2) \cup B_f(s_1, s_2)$, then there exists $\varphi_1 \in C(S_1)$ such that $L(s_1, e, \varphi_1) \neq \bot$, and either $\varphi = \varphi_1^{\bot}$ or $\varphi = \varphi_{12}^{B}$ as in the case above. In both cases, as proven before, there exists $\mu_1 \in Sat(\varphi_1)$ such that $\mu_P \Subset_{\mathcal{R}_1} \mu_1$.

- Let $a \in A$ and $\varphi_1 \in C(S_1)$ such that $L_1(s_1, a, \varphi_1) = \top$.
  If $s_2 = \bot$ or $e = \varepsilon$ or $a \neq e$, then by construction of $N_1 \setminus^K N_2$, the constraint $\varphi_1^{\bot}$ is such that $L((s_1, s_2, e, n), a, \varphi_1^{\bot}) = \top$. As a consequence, there exists a distribution $\mu_P \in Dist(S_P)$ such that $L_P(p, a, \mu_P) = \top$ and there exists $\mu \in Sat(\varphi_1^{\bot})$ such that $\mu_P \Subset_{\mathcal{R}^{\setminus}} \mu$. Moreover, by construction of $\varphi_1^{\bot}$, the distribution $\mu\downarrow_1$: $s_1' \mapsto \mu(s_1', \bot, \varepsilon, 1)$ is in $Sat(\varphi_1)$ and it follows that $\mu_P \Subset_{\mathcal{R}_1} \mu\downarrow_1$.
  Otherwise, assume that $s_2 \in S_2$, $e \in A$ and $a = e$. Since $L_1(s_1, a, \varphi_1) = \top$, (s_1, s_2) can only be in cases $3.a, 3.c$ or $3.f$. As a consequence, $e \in B_a(s_1, s_2) \cup B_c(s_1, s_2) \cup B_f(s_1, s_2)$. By construction, in all of these cases, we have $L((s_1, s_2, e, n), a, \varphi_1^{\bot}) = \top$. Thus, there exists a distribution $\mu_P \in Dist(S_P)$ such that $L_P(p, a, \mu_P) = \top$ and there exists $\mu \in Sat(\varphi_1^{\bot})$ such that $\mu_P \Subset_{\mathcal{R}^{\setminus}} \mu$. As above, it follows that $\mu_P \Subset_{\mathcal{R}_1} \mu\downarrow_1$.

Finally, $\mathcal{R}_1$ is a satisfaction relation. Moreover, by hypothesis, we have $s_0^P \,\mathcal{R}^{\setminus} (s_0^1, s_0^2, f, K)$, thus $s_0^P \,\mathcal{R}_1\, s_0^1$ and $P \models N_1$.

We now prove that $P \not\models N_2$. Assume the contrary and let $\mathcal{R}_2 \subseteq S_P \times S_2$ be the smallest satisfaction relation witnessing $P \models N_2$ (i.e. containing only reachable states). We prove the following by induction on the value of n, for $1 \leq n \leq K$: $\forall p \in S_P, s_2 \in S_2$, if there exists $s_1 \in S_1$ and $e \in A$ such that $p\,\mathcal{R}^{\setminus}(s_1, s_2, e, n)$, then $(p, s_2) \notin \mathcal{R}_2$.

- **Base Case ($n = 1$).** Let p, s_1, s_2, e such that $p\,\mathcal{R}^{\setminus}(s_1, s_2, e, 1)$. If $e \in B_a(s_1, s_2) \cup B_b(s_1, s_2) \cup B_d(s_1, s_2)$, then by construction there is an e transition in either P or N_2 that cannot be matched by the other. Thus $(p, s_2) \notin \mathcal{R}_2$. The same is verified if $e \in B_e(s_1, s_2)$ and there is no distribution $\mu_P \in Dist(S_P)$ such that $L_P(p, e, \mu_P) = \top$. Else, $e \in B_e(s_1, s_2) \cup B_c(s_1, s_2) \cup B_f(s_1, s_2)$ and there exists $\mu_P \in Dist(S_P)$ such that $L_P(p, e, \mu_P) = \top$. Let $\varphi_1 \in C(S_1)$ and $\varphi_2 \in C(S_2)$ be the corresponding constraints in N_1 and N_2. Consider the corresponding constraint $\varphi_{12}^{B,1} \in C(S)$. By $\mathcal{R}^{\setminus}$, there exists $\mu \in Sat(\varphi_{12}^{B,1})$ such that $\mu_P \Subset_{\mathcal{R}^{\setminus}} \mu$. By construction of $\varphi_{12}^{B,1}$, we know that either (3.a) there exists $(s_1', \bot, \varepsilon, 1)$ such that $\mu(s_1', \bot, \varepsilon, 1) > 0$ or (3.b) the distribution $\mu_2 : s_2' \mapsto \sum_{c\in A\cup\{\varepsilon\}, s_1'\in S_1, k'\geq 1} \mu(s_1', s_2', c, k')$ does not satisfy φ_2. If there exists $(s_1', \bot, \varepsilon, 1)$ such that $\mu(s_1', \bot, \varepsilon, 1) > 0$, then there exists $p' \in S_P$ such that $\mu_P(p') > 0$ and $\mathsf{succ}_{s_2,e}(p') = \emptyset$. Thus there cannot exists $\mu_2' \in Sat(\varphi_2)$ such that $\mu_P \Subset_{\mathcal{R}_2} \mu_2'$. Otherwise, by determinism of N_2, we know that the only possible correspondence function for μ_P and $\mathcal{R}_2$ is $\delta_2 : S_P \to (S_2 \to [0,1])$ such that $\delta_2(p')(s_2') = 1$ if $s_2' = \mathsf{succ}_{s_2,e}(p')$ and 0 otherwise. By construction, we have

$\mu_P\delta_2 = \mu_2$ and thus there is no distribution $\mu_2' \in Sat(\varphi_2)$ such that $\mu_P \Subset_{\mathcal{R}_2} \mu_2'$. Consequently, $(p, s_2) \notin \mathcal{R}_2$.

- **Induction.** Let $1 < n \leq K$ and assume that for all $k < n$, for all $p' \in S_P, s_2' \in S_2$, whenever there exists $s_1' \in S_1$ and $e \in A$ such that $p' \mathcal{R}^{\setminus}(s_1', s_2', e, k)$, we have $(p', s_2') \notin \mathcal{R}_2$. Let p, s_1, s_2, e such that $p \mathcal{R}^{\setminus}(s_1, s_2, e, n)$. If $e \in B_a(s_1, s_2) \cup B_b(s_1, s_2) \cup B_d(s_1, s_2)$, then by construction there is an e transition in either P or N_2 that cannot be matched by the other. Thus $(p, s_2) \notin \mathcal{R}_2$. The same is verified if $e \in B_e(s_1, s_2)$ and there is no distribution $\mu_P \in Dist(S_P)$ such that $L_P(p, e, \mu_P) = \top$. Else, $e \in B_e(s_1, s_2) \cup B_c(s_1, s_2) \cup B_f(s_1, s_2)$ and there exists $\mu_P \in Dist(S_P)$ such that $L_P(p, e, \mu_P) = \top$. Let $\varphi_1 \in C(S_1)$ and $\varphi_2 \in C(S_2)$ be the corresponding constraints in N_1 and N_2.

  Consider the corresponding constraint $\varphi_{12}^{B,n} \in C(S)$. By $\mathcal{R}^{\setminus}$, there exists $\mu \in Sat(\varphi_{12}^{B,n})$ such that $\mu_P \Subset_{\mathcal{R}^{\setminus}} \mu$. By construction of $\varphi_{12}^{B,n}$, we know that either (3.a) there exists $(s_1', \perp, c, 1)$ such that $\mu(s_1', \perp, c, 1) > 0$ or (3.b) the distribution $\mu_2 : s_2' \mapsto \sum_{c \in A \cup \{\varepsilon\}, s_1' \in S_1, k' \geq 1} \mu(s_1', s_2', c, k')$ does not satisfy φ_2, or (3.c) there exists $s_1' \in S_1$, $s_2' \in S_2$, $c \neq \varepsilon$ and $k < n$ such that $\mu(s_1', s_2', c, k) > 0$. If case (3.a) or (3.b) holds, then as in the base case, there is no distribution $\mu_2' \in Sat(\varphi_2)$ such that $\mu_P \Subset_{\mathcal{R}_2} \mu_2'$. Otherwise, if (3.c) holds, then there exists $p' \in S_P$ such that $\mu_P(p') > 0$ and $p' \mathcal{R}^{\setminus}(s_1', s_2', c, k)$. By induction, we thus know that $(p', s_2') \notin \mathcal{R}_2$ and by construction and determinism of N_2, we have that $\mathsf{succ}_{s_2,e}(p') = \{s_2'\}$. Thus there is no distribution $\mu_2' \in Sat(\varphi_2)$ such that $\mu_P \Subset_{\mathcal{R}_2} \mu_2'$. Consequently, $(p, s_2) \notin \mathcal{R}_2$.

By hypothesis, we have $s_0^P \mathcal{R}^{\setminus}(s_0^1, s_0^2, f, K)$. As a consequence, we have that $(s_0^P, s_0^2) \notin \mathcal{R}_2$, implying that $P \not\models N_2$.

2. We now prove that for all PA $P \in [\![N_1]\!] \setminus [\![N_2]\!]$, there exists $K \in \mathbb{N}$ such that $P \in [\![N_1 \setminus^K N_2]\!]$. If $V_1(s_0^1) \neq V_2(s_0^2)$, then for all $K \in \mathbb{N}$, we have $N_1 \setminus^K N_2 = N_1$ and the result holds.

Otherwise, assume that (s_0^1, s_0^2) is in case 3. Let $P = (S_P, A, L_P, AP, V_P, s_0^P)$ be a PA such that $P \models N_1$ and $P \not\models N_2$. Let $\mathcal{R}_1$ be the satisfaction relation witnessing $P \models N_1$ and $\mathcal{R}_2$ be the maximal satisfaction relation between P and N_2. Assume that $\mathcal{R}_2$ is computed as described in Section 4.1. Let $\mathsf{Ind}_{\mathcal{R}_2}$ be the associated index function and let K be the minimal index such that $\mathcal{R}_{2K} = \mathcal{R}_2$. We show that $P \models N_1 \setminus^K N_2$. Let $N_1 \setminus^K N_2 = (S, A, L, AP, V, S_0)$ be defined as in Section 5.2.

Let $\mathcal{R}^{\setminus} \subseteq S_P \times S_2$ be the relation such that

$$p\,\mathcal{R}^{\setminus}(s_1, s_2, e, k) \iff \left\{ \begin{array}{ll} & (p\,\mathcal{R}_1\, s_1) \text{ and } (s_2 = \perp) \text{ and } (e = \varepsilon) \text{ and } (k = 1) \\ \text{or} & (p\,\mathcal{R}_1\, s_1) \text{ and } (p, s_2) \text{ in case 1 or 2 and } \text{ and } (e = \varepsilon) \text{ and } (k = 1) \\ \text{or} & \left\{ \begin{array}{r} (p\,\mathcal{R}_1\, s_1) \text{ and } (p, s_2) \text{ in case 3 and } (e \in \mathsf{Break}(p, s_2)) \\ \text{and } (k = \mathsf{Ind}_{\mathcal{R}_2}(p, s_2) + 1) \end{array} \right. \end{array} \right.$$

Remark that whenever (p, s_2) is in case 3, we know that $\mathsf{Ind}_{\mathcal{R}_2}(p, s_2) < K$, thus $\mathsf{Ind}_{\mathcal{R}_2}(p, s_2) + 1 \leq K$.

We prove that $\mathcal{R}^{\setminus}$ is a satisfaction relation. Let $p\,\mathcal{R}^{\setminus}(s_1, s_2, e, k)$.
If $s_2 = \perp$ or $e = \varepsilon$, then since $p\,\mathcal{R}_1\, s_1$, $\mathcal{R}^{\setminus}$ satisfies the axioms of a satisfaction relation by construction.
Else we have $s_2 \in S_2$ and $e \neq \varepsilon$, thus, by definition of $\mathcal{R}^{\setminus}$, we know that (p, s_2) is in case 3. The rest of the proof is almost identical to the proof of Theorem 12. In the following, we report to this proof and only highlight the differences.

- By construction, we have $V_P(p) \in V_1(s_1) = V((s_1, s_2, e, k))$.

- Let $a \in A$ and $\mu_P \in Dist(S_P)$ such that $L_P(p, a, \mu_P) = \top$. There are several cases.
  - If $a \neq e$, or $a = e \in B_a(p, s_2)$, the proof is identical to the proof of Theorem 12.
  - Else, we necessarily have $a = e \in B_c(p, s_2) \cup B_f(p, s_2)$. Remark that, by construction, $B_c(p, s_2) \subseteq B_c(s_1, s_2)$ and $B_f(p, s_2) \subseteq B_f(s_1, s_2)$. Since $p \, \mathcal{R}_1 \, s_1$, there exists $\varphi_1 \in C(S_1)$ such that $L_1(s_1, e, \varphi_1) \neq \bot$ and there exists $\mu_1 \in Sat(\varphi_1)$ and a correspondence function $\delta_1 : S_P \to (S_1 \to [0, 1])$ such that $\mu_P \Subset^{\delta_1}_{\mathcal{R}_1} \mu_1$.

    Moreover, by construction of $N_1 \setminus^K N_2$, we know that the constraint $\varphi_{12}^{B,k}$ is such that $L((s_1, s_2, e, k), e, \varphi_{12}^{B,k}) = \top$.

    We now prove that there exists $\mu \in Sat(\varphi_{12}^{B,k})$ such that $\mu_P \Subset_{\mathcal{R}^\setminus} \mu$. Consider the function $\delta : S_P \to (S \to [0, 1])$ defined as follows: Let $p' \in S_P$ such that $\mu_P(p') > 0$ and let $s'_1 = \mathsf{succ}_{s_1,e}(p')$, which exists by $\mathcal{R}_1$.

    - If $\mathsf{succ}_{s_2,e}(p') = \emptyset$, then $\delta(p')(s'_1, \bot, \varepsilon, 1) = 1$.
    - Else, let $s'_2 = \mathsf{succ}_{s_2,e}(p')$. Then,
      - if $(p', s'_2) \in \mathcal{R}_2$, then $\delta(p')(s'_1, s'_2, \varepsilon, 1) = 1$.
      - Else, (p', s'_2) is in case 3 and $\mathsf{Break}(p', s'_2) \neq \emptyset$. In this case, let $c \in \mathsf{Break}(p', s'_2)$ and define $\delta(p', (s'_1, s'_2, c, \mathsf{Ind}_{\mathcal{R}_2}(p', s'_2) + 1)) = 1$. For all other $c' \in A$ and $1 \leq k' \leq K$, define $\delta(p', (s'_1, s'_2, c', k')) = 0$.

    Remark that for all $p' \in S_P$ such that $\mu_P(p') > 0$, there exists a unique $s' \in S'$ such that $\delta(p')(s') = 1$. Thus δ is a correspondence function.

    We now prove that $\mu = \mu_P \delta \in Sat(\varphi_{12}^{B,k})$.

    1. Let $(s'_1, s'_2, c, k') \in S$ such that $\mu(s'_1, s'_2, c, k') > 0$. By construction, there exists $p' \in S_P$ such that $\mu_P(p') > 0$ and $\delta(p')(s'_1, s'_2, c, k') > 0$. Moreover, $c \in B(s'_1, s'_2) \cup \{\varepsilon\}$, $s'_2 = \bot$ if $\mathsf{succ}_{s_2,e}(s'_1) = \emptyset$ and $s'_2 = \mathsf{succ}_{s_2,e}(s'_1)$ otherwise.
    2. Consider the distribution $\mu'_1 : s'_1 \mapsto \sum_{c \in A \cup \{\varepsilon\}, s'_2 \in S_2 \cup \{\bot\}, k' \geq 1} \mu(s'_1, s'_2, c, k')$. By determinism (See Lemma 28 in [6]), we have that $\delta_1(p')(s'_1) = 1 \iff s'_1 = (succ)_{s_1,e}(p')$. As a consequence, we have that $\mu'_1 = \mu \delta_1 = \mu_1 \in Sat(\varphi_1)$.
    3. Depending on k, there are 2 cases.
       - If $k > 1$, assume that for all $p' \in S_P$ such that $\mu_P(p') > 0$, we have $\mathsf{succ}_{s_2,e}(p') \neq \emptyset$ (the other case being trivial). Since $c \in (B_c(p, s_2) \cup B_f(p, s_2)) \cap \mathsf{Break}(p, s_2)$ by $\mathcal{R}^\setminus$, we can apply Lemma 17. As a consequence, either (2) the distribution $\mu_1^2 : \left(s'_2 \mapsto \sum_{p' \in P \mid s'_2 = \mathsf{succ}_{s_2,e}(p')} \mu_P(p')\right) \notin \varphi_2$, or (3) there exists $p' \in S_P$ and $s'_2 \in S_2$ such that $\mu_P(p') > 0$, $s'_2 = \mathsf{succ}_{s_2,e}(p')$ and $\mathsf{Ind}_{\mathcal{R}_2}(p', s'_2) < \mathsf{Ind}_{\mathcal{R}_2}(p, s_2)$.

         In the first case (2), consider the distribution μ_2 defined as follows:

         $$\mu_2 : s'_2 \mapsto \sum_{c \in A \cup \{\varepsilon\}, s'_1 \in S_1, k' \geq 1} \mu(s'_1, s'_2, c, k').$$

         We have the following: for all $s'_2 \in S_2$,

$$
\begin{aligned}
\mu_2(s_2') &= \sum_{c \in A \cup \{\varepsilon\}, s_1' \in S_1, k' \geq 1} \mu(s_1', s_2', c, k') \\
&= \sum_{c \in A \cup \{\varepsilon\}, s_1' \in S_1, k' \geq 1} \sum_{p' \in S_P} \mu_P(p')\delta(p')((s_1', s_2', c, k')) \\
&= \sum_{p' \in S_P} \mu_P(p') \sum_{c \in A \cup \{\varepsilon\}, s_1' \in S_1, k' \geq 1} \delta(p')((s_1', s_2', c, k')) \\
&= \sum_{p' \in S_P \ \mid \ s_2' = \mathsf{succ}_{s_2,e}(p')} \mu_P(p')\delta(p')((\mathsf{succ}_{s_1,e}(p'), s_2', c, \mathsf{Ind}_{\mathcal{R}_2}(p', s_2'))) \\
&\qquad \text{for } c \in \mathsf{Break}(p', s_2') \text{ fixed as above} \\
&= \sum_{p' \in S_P \ \mid \ s_2' = \mathsf{succ}_{s_2,e}(p')} \mu_P(p') \\
&= \mu_1^2(s_2')
\end{aligned}
$$

As a consequence, $\mu_2 \notin Sat(\varphi_2)$ and $\mu \in Sat(\varphi_{12}^{B,k})$.

In the second case (3), we have $\delta(p')((s_1', s_2', c, k')) > 0$ for $s_1' = \mathsf{succ}_{s_1,e}(p')$, $c \in \mathsf{Break}(p', s_2')$ fixed above, and $k' = \mathsf{Ind}_{\mathcal{R}_2}(p', s_2') + 1 < \mathsf{Ind}_{\mathcal{R}_2}(p, s_2) + 1 = k$. As a consequence, we thus have $\mu(s_1', s_2', c, k') > 0$ for $k' < k$ and $c \neq \varepsilon$, thus $\mu \in Sat(\varphi_{12}^{B,k})$.

- On the other hand, if $k = 1$, then $\mathsf{Ind}_{\mathcal{R}_2}(p, s_2) = 0$ and either (1) there exists $p' \in S_P$ such that $\mu_P(p') > 0$ and $\mathsf{succ}_{s_2,e}(p') = \emptyset$, or (2) the distribution the distribution $\mu_1^2 : \left(s_2' \mapsto \sum_{p' \in P \ \mid \ s_2' = \mathsf{succ}_{s_2,e}(p')} \mu_P(p')\right) \notin \varphi_2$. In both cases, as above, we can prove that $\mu \in Sat(\varphi 12^{B,k}$.

In both cases, we have $\mu \in Sat(\varphi_{12}^{B,k})$.

We thus conclude that there exists $\mu \in Sat(\varphi_{12}^{B,k})$ such that $\mu_P \Subset_{\mathcal{R}^{\setminus}} \mu$.

- Let $a \in A$ and $\varphi \in C(S)$ such that $L((s_1, s_2, e), a, \varphi) = \top$. As in the proof of Theorem 12, there are several cases that all boil down to the same arguments as above.

Finally, $\mathcal{R}^{\setminus}$ is a satisfaction relation.

Let $c \in \mathsf{Break}_{\mathcal{R}_2}(s_0^P, s_0^2)$ and consider the relation $\mathcal{R}^{\setminus\prime} = \mathcal{R}^{\setminus} \cup \{(s_0^P, (s_0^1, s_0^2, c, K))\}$. Due to the fact that $K \geq \mathsf{Ind}_{\mathcal{R}_2}(s_0^P, s_0^2)$, one can verify that the pair $(s_0^P, (s_0^1, s_0^2, c, K))$ also satisfies the axioms of a satisfaction relation. The proof is identical to the one presented above. As a consequence, $\mathcal{R}^{\setminus\prime}$ is also a satisfaction relation. Moreover, we now have that $(s_0^P, (s_0^1, s_0^2, c, K)) \in \mathcal{R}^{\setminus\prime}$, with $(s_0^1, s_0^2, c, K) \in S_0$, thus $P \models N_1 \setminus^K N_2$.

◀

G Proof of Theorem 16

Let N_1 and N_2 be two deterministic APAs in SVNF such that $N_1 \not\preceq N_2$. The following holds:

1. the sequences $(N_1 \setminus^K N_2)_K$ and $([\![N_1 \setminus^K N_2]\!])_K$ both converge,
2. $\lim_{K \to \infty} d_t([\![N_1]\!] \setminus [\![N_2]\!], [\![N_1 \setminus^K N_2]\!] = 0$, and
3. $\lim_{K \to \infty} d(N_1 \setminus^* N_2, N_1 \setminus^K N_2) = 0$, so that
4. $d_t([\![N_1 \setminus^* N_2]\!], [\![N_1]\!] \setminus [\![N_2]\!]) = 0$.

Proof. Let $N_1 = (S_1, A, L_1, AP, V_1, \{s_0^1\})$ and $N_2 = (S_2, A, L_2, AP, V_2, \{s_0^2\})$ be two deterministic APAs in SVNF such that $N_1 \not\preceq N_2$.

1. The proof of the convergence of both sequences $(N_1 \setminus^K N_2)_K$ and $(\llbracket N_1 \setminus^K N_2 \rrbracket)_K$ is done as follows. We show in Lemma 18 that the sequence $(N_1 \setminus^K N_2)_K$ is *bi-Cauchy* (i.e. both forward-Cauchy and backwards-Cauchy)

▶ **Lemma 18.** *Let $N_1 = (S_1, A, L_1, AP, V_1, \{s_0^1\})$ and $N_2 = (S_2, A, L_2, AP, V_2, \{s_0^2\})$ be two deterministic APAs in SVNF such that $N_1 \not\preceq N_2$. Let $1 \leq K_1 \leq K_2$ be integers. The distance between $N_1 \setminus^{K_2} N_2$ and $N_1 \setminus^{K_1} N_2$ is bounded as follows:*

$$d(N_1 \setminus^{K_2} N_2, N_1 \setminus^{K_1} N_2) \leq \lambda^{K_1}.$$

Proof. Let $N_1 \setminus^{K_i} N_2 = N^i = (S^i, A, L^i, AP, V^i, T_0^i)$. The proof is in several steps.

- We first remark that for all $(s_1, s_2, e) \in S_1 \times (S_2 \cup \bot) \times (A \cup \varepsilon)$ and for all $k \leq K_1$, the distance between State $(s_1, s_2, e, k)^1 \in S^1$ and $(s_1, s_2, e, k)^2 \in S^2$ is 0. Indeed, if k is the same in both states, then they are identical by construction.
- We now prove by induction on $1 \leq k_1 \leq K_1$ and $k_1 \leq k_2 \leq K_2$ that $d((s_1, s_2, e, k_2)^2, (s_1, s_2, e, k_1)^1) \leq \lambda^{k_1}$.
  - **Base case:** $k_1 = 1$. By construction, $t_1 = (s_1, s_2, e, k_1)^1$ and $t_2 = (s_1, s_2, e, k_2)^2$ have the same outgoing transitions. The only distinction is in the constraints $\varphi_{12}^{B,1}$ and φ_{12}^{B,k_2} when $e \in B_c(s_1, s_2) \cup B_e(s_1, s_2) \cup B_f(s_1, s_2)$. As a consequence, the states t_1 and t_2 are compatible, thus

    $$d(t_2, t_1) = \max \begin{cases} \max\limits_{a,\varphi' \mid L^2(t_2,a,\varphi')\neq\bot} \left(\min\limits_{\varphi \mid L^1(t_1,a,\varphi)\neq\bot} \lambda D_{N^2,N^1}(\varphi', \varphi, d) \right) \\ \max\limits_{a,\varphi \mid L^1(t_1,a,\varphi)=\top} \left(\min\limits_{\varphi' \mid L^2(t_2,a,\varphi')=\top} \lambda D_{N^2,N^1}(\varphi', \varphi, d) \right) \end{cases}$$

    Moreover, we know by construction that $D_{N^2,N^1}(\varphi', \varphi, d) \leq 1$ for all φ' and φ. As a consequence, $d(t_2, t_1) \leq \lambda = \lambda^{k_1}$.
  - **Induction.** Let $t_1 = (s_1, s_2, e, k_1)^1$ and $t_2 = (s_1, s_2, e, k_2)^2$, with $1 < k_1 \leq k_2$. Again, if $e \notin B_c(s_1, s_2) \cup B_e(s_1, s_2) \cup B_f(s_1, s_2)$, then t_1 and t_2 are identical by construction and the result holds. Otherwise, the pair of constraints for which the distance is maximal will be constraints $\varphi_{12}^{B,k_1} \in C(S^1)$ and $\varphi_{12}^{B,k_2} \in C(S^2)$. Assume that $d((s_1, s_2, e, k_2')^2, (s_1, s_2, e, k_1')^1) \leq \lambda^{k_1'}$ for all $k_1' < k_1$ and $k_1' \leq k_2' \leq K_2$. By definition, we have

    $$D_{N^2,N^1}(\varphi_{12}^{B,k_2}, \varphi_{12}^{B,k_1}, d) = \sup_{\mu_2 \in Sat(\varphi_{12}^{B,k_2})} \left[\inf_{\delta \in \mathsf{RD}(\mu_2, \varphi_{12}^{B,k_1})} \left(\sum_{t_2', t_1' \in S^2 \times S^1} \mu_2(t_2') \delta(t_2', t_1') d(t_2', t_1') \right) \right]$$

    Consider the function $\delta : S^2 \times S^1 \to [0,1]$ such that

    $$\delta((s_1', s_2', f, k_2'), (s_1'', s_2'', f', k_1')) = \begin{cases} 1 & \text{if } s_1' = s_1'' \wedge s_2' = s_2'' \wedge f' = f \\ & \quad \wedge k_1' = k_2' \wedge k_2' < k_1 \\ 1 & \text{if } s_1' = s_1'' \wedge s_2' = s_2'' \wedge f' = f \\ & \quad \wedge k_1' = k_1 - 1 \wedge k_1 \leq k_2' \\ 0 & \text{otherwise} \end{cases}$$

    Let $\mu_2 \in Sat(\varphi_{12}^{B,k_2})$. One can verify that $\delta \in \mathsf{RD}(\mu_2, \varphi_{12}^{B,k_1})$ as follows:

1. Let $t_2' = (s_1', s_2', f, k_2')$ be such that $\mu_2(t_2') > 0$. By definition, we always have $\sum_{t_1' \in S^1} \delta(t_2', t_1') = 1$.
2. δ preserves all the conditions for satisfying φ_{12}^{B,k_2}. In particular, all states $t_2' = (s_1', s_2', f, k_2')^2$ such that $k_2' < k_2$ are redistributed to states $(s_1', s_2', f, k_1')^1$ with $k_1' < k_1$. As a consequence, the distribution $\mu_1 : t_1' \mapsto \sum_{t_2' \in S^2} \mu_2(t_2')\delta(t_2', t_1')$ satisfies φ_{12}^{B,k_1}.

As a consequence, for all $\mu_2 \in Sat(\varphi_{12}^{B,k_2})$, we have

$$\begin{aligned}
&\inf_{\delta \in \mathsf{RD}(\mu_2, \varphi_{12}^{B,k_1})} \left(\sum_{t_2', t_1' \in S^2 \times S^1} \mu_2(t_2')\delta(t_2', t_1')d(t_2', t_1') \right) \\
&\leq \sum_{\substack{(s_1', s_2', f, k_2') \in S^2 \\ k_2' < k_1}} \mu_2(s_1', s_2', f, k_2') d((s_1', s_2', f, k_2')^2, (s_1', s_2', f, k_2')^1) \\
&\qquad + \sum_{\substack{(s_1', s_2', f, k_2') \in S^2 \\ k_1 \leq k_2'}} \mu_2(s_1', s_2', f, k_2') d((s_1', s_2', f, k_2')^2, (s_1', s_2', f, k_1 - 1)^1) \\
&\leq \sum_{\substack{(s_1', s_2', f, k_2') \in S^2 \\ k_1 \leq k_2'}} \mu_2(s_1', s_2', f, k_2') d((s_1', s_2', f, k_2')^2, (s_1', s_2', f, k_1 - 1)^1) \\
&\leq \sum_{\substack{(s_1', s_2', f, k_2') \in S^2 \\ k_1 \leq k_2'}} \mu_2(s_1', s_2', f, k_2') \lambda^{k_1 - 1} \qquad\qquad \textbf{by induction} \\
&\leq \lambda^{k_1 - 1}
\end{aligned}$$

Since this is true for all $\mu_2 \in Sat(\varphi_{12}^{B,k_2})$, we thus have

$$D_{N^2,N^1}(\varphi_{12}^{B,k_2}, \varphi_{12}^{B,k_1}, d) \leq \lambda^{k_1 - 1}.$$

Finally, we have $d(t_2, t_1) \leq \lambda\lambda^{k_1 - 1} = \lambda^k$, which proves the induction.

- For all state $t_0^2 = (s_0^1, s_0^2, e, K_2) \in T_0^2$, there exists a state $t_0^1 = (s_0^1, s_0^2, e, K_1) \in T_0^1$ such that $d(t_0^2, t_0^1) \leq \lambda^{K_1}$. As a consequence, we have $d(N_1 \setminus^{K_2} N_2, N_1 \setminus^{K_1} N_2) \leq \lambda^{K_1}$.

◀

Let $\varepsilon > 0$. Since $\lambda < 1$, there exists $K \in \mathbb{N}$ such that $\lambda^K < \varepsilon$. As a consequence, by the above lemma, we have that for all $K \leq K_1 \leq K_2$,

$$d(N_1 \setminus^{K_2} N_2, N_1 \setminus^{K_1} N_2) \leq \lambda^{K_1} \leq \lambda^K < \varepsilon.$$

The sequence $(N_1 \setminus^K N_2)_K$ is thus bi-Cauchy. Hence, because of Theorem 7, the sequence (of sets of PA) $([\![N_1 \setminus^K N_2]\!])_K$ is also bi-Cauchy. The other two items show that they converge.

2. Theorem 15 shows that the sequence $([\![N_1 \setminus^K N_2]\!])_K$ converges in a set-theoretic sense (as a direct limit), and that $\lim_{K\to\infty}[\![N_1 \setminus^K N_2]\!] = [\![N_1]\!] \setminus [\![N_2]\!]$. Hence $d_t([\![N_1]\!] \setminus [\![N_2]\!], \lim_{K\to\infty}[\![N_1 \setminus^K N_2]\!] = 0$, and by continuity of d_t, $\lim_{K\to\infty} d_t([\![N_1]\!] \setminus [\![N_2]\!], [\![N_1 \setminus^K N_2]\!] = 0$.

3. Finally, we prove that $\lim_{K\to\infty} d(N_1 \setminus^* N_2, N_1 \setminus^K N_2) = 0$. This proof is very similar to the proof of Lemma 18 above: we can show that the distance between $N_1 \setminus^* N_2$ and $N_1 \setminus^K N_2$ is bounded as follows:

$$d(N_1 \setminus^* N_2, N_1 \setminus^K N_2) \leq \lambda^K.$$

Let $N_1 \setminus^K N_2 = N^K = (S^K, A, L^K, AP, V^K, T_0^K)$ and $N_1 \setminus^* N_2 = N^* = (S^*, A, L^*, AP, V^*, T_0^*)$. We start by proving by induction on $1 \leq k \leq K$ that for all $(s_1, s_2, e) \in S_1 \times (S_2 \cup \bot) \times (A \cup \varepsilon)$, we have $d((s_1, s_2, e)^*, (s_1, s_2, e, k)) \leq \lambda^k$. The only difference with the proof of Lemma 18 is in the choice of the function $\delta : S^* \times S^K \to [0, 1]$ in the induction part. Here, we choose δ as follows:

$$\delta((s'_1, s'_2, f), (s''_1, s''_2, f', k')) = \begin{cases} 1 & \text{if } s'_1 = s''_1 \wedge s'_2 = s''_2 \wedge f' = f \wedge k' = k - 1 \\ 0 & \text{otherwise} \end{cases}$$

The rest of the proof is identical, and we obtain that for all $1 \leq k \leq K$ and for all $(s_1, s_2, e) \in S_1 \times (S_2 \cup \bot) \times (A \cup \varepsilon)$, we have $d((s_1, s_2, e)^*, (s_1, s_2, e, k)) \leq \lambda^k$. In particular, this is also true for initial states. As a consequence, for all state $t_0^* = (s_0^2, s_0^1, e) \in T_0^*$, there exists a state $t_0^K = (s_0^1, s_0^2, e, K) \in T_0^K$ such that $d(t_0^*, t_0^K) \leq \lambda^K$. As a consequence, we have $d(N_1 \setminus^* N_2, N_1 \setminus^K N_2) \leq \lambda^K$.

As a consequence, we obtain:

$$\lim_{K \to \infty} d(N_1 \setminus^* N_2, N_1 \setminus^K N_2) = 0.$$

4. The last part of the theorem is now a direct consequence of the other items and Lemma 16. We know that $d_t(\llbracket N_1 \setminus^K N_2 \rrbracket, \llbracket N_1 \rrbracket \setminus \llbracket N_2 \rrbracket) = 0$ for every K, and since $\lim_{K \to \infty} d(N_1 \setminus^* N_2, N_1 \setminus^K N_2) = 0$, it follows by the triangle inequality that $\lim_{K \to \infty} d_t(\llbracket N_1 \setminus^* N_2 \rrbracket, \llbracket N_1 \rrbracket \setminus \llbracket N_2 \rrbracket) \leq \lim_{K \to \infty} d_t(\llbracket N_1 \setminus^* N_2 \rrbracket, \llbracket N_1 \setminus^K N_2 \rrbracket) + \lim_{K \to \infty} d_t(\llbracket N_1 \setminus^K N_2 \rrbracket, \llbracket N_1 \rrbracket \setminus \llbracket N_2 \rrbracket) = 0$. ◀