Statistic Analysis for Probabilistic Processes[1]

Michel de Rougemont, Mathieu Tracol[2]
University Paris-South, LRI-CNRS
mdr@lri.fr, tracol@lri.fr

[1]An Extended Abstract of this paper appeared in LICS 2009 [8]. This research was funded by the French program on Computer Security **ANR-07-SESU-013**
[2]Corresponding Author. Mail: tracol@lri.fr. Tel.: +33 (0)1 69 15 76 43. Fax: 33 (0)1 69 15 65 86

Abstract. We associate a statistical vector to a trace and a geometrical embedding to a Markov Decision Process, based on a distance on words, and study basic Membership and Equivalence problems. The Membership problem for a trace w and a Markov Decision Process $\mathcal{S}$ decides if there exists a strategy on $\mathcal{S}$ which generates with high probability traces close to w. We prove that Membership of a trace is *testable* and Equivalence of MDPs is polynomial time approximable. For Probabilistic Automata, Membership is not testable, and approximate Equivalence is undecidable. We give a class of properties, based on results concerning the structure of the tail σ-field of a finite Markov chain, which characterizes equivalent Markov Decision Processes in this context.

Keywords: Markov Decision Processes, Probabilistic Automata, State Action frequency, Tail σ-Field, Property Testing, Approximation, Approximate Verification.

Contents

1. Introduction

We consider probabilistic systems with both non deterministic and probabilistic transitions, as Markov Decision Processes, and basic questions concerning their traces such as *Statistical membership* for a given system and *Equivalence* of two systems which are known to be hard in this context [3, 17]. We study their approximation, based on Property testing, with a natural distance dist on words, and show that *Statistical membership* becomes *testable* and *Equivalence* polynomial time computable in the size of the system, but remain hard for *Probabilistic Automata*.

Property testing [24, 13] is a classical method to approximate decision problems, given a distance between two inputs. An ε-*tester* for a property P on words, is a randomized algorithm $\mathcal{A}$ which takes a word w_n of size n as input, and distinguishes with high probability between w_n satisfies P and w_n ε-far from P. A property P is *testable* if there exists a randomized algorithm such that for every $\varepsilon > 0$, $\mathcal{A}(\varepsilon)$ is an ε-tester for P whose time complexity only depends on ε, *i.e.* is independent of the size n.

Let $\mathcal{S}$ be an MDP (Markov Decision Process) of size m, $\lambda \leq 1$ a threshold value and $0 \leq \varepsilon \leq 1$. Given an input w_n of size n, the *Statistical membership* decides if there exists a strategy σ, which assigns decisions on each non deterministic state, such that $Prob_\sigma[\mathsf{dist}(r_n, w_n) \leq \varepsilon] \geq \lambda$, i.e. the probability to observe a trace r_n which is ε-close to w_n is greater than λ. If $\mathcal{S}$ is part of the input, we prove that an exact version of a problem close to statistical membership problem is PSPACE-hard. Given a fixed $\mathcal{S}$, the problem is testable, i.e. can be approximated in time independent of the length of the word n. We present a method which generalizes the approach introduced in [12] for non deterministic systems and considers the statistical behavior of a system. It associates a statistical vector x to

w_n and a convex set $\mathcal{H}(\mathcal{S})$ of vectors to $\mathcal{S}$ in such a way that the geometrical distance between x and $\mathcal{H}(\mathcal{S})$ is close to the distance between w_n and the set of traces of $\mathcal{S}$. We define a distance between two MDPs $\mathcal{S}_1$ and $\mathcal{S}_2$ as the geometrical distance between $H(\mathcal{S}_1)$ and $H(\mathcal{S}_2)$ and show how to approximate it in polynomial time. A motivation for this statistical analysis is to decide if there are runs with some statistic constraints, such as the proportion of action a greater than ten percent in the run.

In [28], Tzeng studied the equivalence between two probabilistic automata and proved that the exact equivalence is in PTIME. This result was extended in [11], where the authors study labeled Markov chains in a context close to ours. Tzeng defined the *approximate equivalence* in a natural way: two probabilistic automata are ε-close if for all words w, the probabilities to be accepted are ε-close. The undecidability of this problem is proved in [17]. In this context, given a word w_n, the *membership* simply decides if $Prob_A[w_n \text{ is accepted}] \geq \lambda$. We show that this property is not testable.

There are many other approaches which associate distances to probabilistic systems. In [29, 10], distances generalize the classical probabilistic bisimulation between two states and in [19, 26] the $\bar{d}$ generalizes the Trace Equivalence between two Markov chains. The distance we introduce distinguishes systems which have a different long term behavior, and is most relevant for systems which are not supposed to stop. Our main results are: a generalization of Derman's theorem to higher order statistics **(Theorem 1)**. The *Statistical membership* on MDPs is testable **(Theorem 2)**, whereas it is not testable **(Theorem 3)** for Probabilistic Automata. Approximate Equivalence is polynomial time computable for MDPs **(Proposition 4)** whereas it is undecidable for Probabilistic Automata. *Ultimate properties* characterize the MDPs and the Markov chains at distance 0 **(Theorem 5,6,7).**

In section 2, we review the main definitions for Testers, MDPs and State-Action Frequencies, Probabilistic Automata. In section 3, we generalize known results on MDPs and statistics to higher order statistics. In section 4 we define the *Statistical membership* and *Equivalence* problems, and prove positive results for MDPs. In section 5 we prove negative results for Probabilistic Automata. In section 6 we present a class of properties (ultimate properties), which characterizes exactly the equivalence and simulation relations between MDPs induced by the distance.

2. Preliminaries

2.1. Testers and Statistics on words

An *elementary operation* on a word w_n of size n on an alphabet Σ is an insertion, a deletion, a substitution of a single letter, or the *move* of a whole subword of w to another position. The *edit distance with moves* $\mathsf{dist}(w, w')$ between w and w' is the minimal number of elementary operations performed on w to obtain w', divided by $max\{|w|, |w'|\}$ as we only consider relative distances. The distance between w and a language L, noted $\mathsf{dist}(w, L)$ is the minimum distance $\mathsf{dist}(w, w')$ for $w' \in L$.

Definition 1 (Property Tester [24, 13])**.** *Let $\varepsilon > 0$. An ε-*tester *for a language L is a randomized algorithm $\mathcal{A}$ such that, for all words w_n as input:*
(1) If $w_n \in L$, then $\mathcal{A}$ accepts with probability at least $2/3$,
(2) If w_n is ε-far from L, then $\mathcal{A}$ rejects with probability at least $2/3$.

Applying a polynomial number of times the algorithm $\mathcal{A}$ on the same input w, and using a majority vote, the $2/3$ bound could be replaced by any real number $\rho > 1/2$. A *query* asks for the value of $w[i]$ for some i. The *query complexity* is the number of queries made to the word, and the *time complexity* is the usual definition, where we assume that arithmetic operations, a uniformly random choice of an integer from any finite range not larger than the input size, and a query to the input, take constant time. A language L is *testable*, if there exists a randomized algorithm $\mathcal{A}$ such that, for every real $\varepsilon > 0$ as input, $\mathcal{A}(\varepsilon)$ is an ε-tester of L, and the query and time complexities of $\mathcal{A}$ depend only on ε.

The $\mathsf{ustat}_k(w_n)$ vector of w_n, of dimension $|\Sigma|^k$, also called the k-gram of w_n is a vector whose u component, $\mathsf{ustat}_k(w_n)[u]$ for u a word of size k, is the number of different occurrences of u in w_n divided by $n - k + 1$, the number of blocks of size k in w_n. It is also the probability to find u in a uniform random block. For instance, for $k = 2$ and $\Sigma = \{0, 1\}$ there are 4 possible words u of length k, which we take in lexicographic order. For $w_6 = 101101 \in \Sigma^*$ we get $ustat_2(w_6) = (0, 2/5, 2/5, 1/5)$. We will use the result of [12], which relates dist to the L_1 distance between ustat vectors:

Proposition 1. *For large enough words $w, w' \in \Sigma^*$ and $k \in \mathbb{N}$, if $\epsilon = \ulcorner 1/k \urcorner$ and $\delta > 0$, then:*

- *if* $\mathsf{dist}(w, w') \leq \delta$*, then* $||ustat_k(w) - ustat_k(w')||_1 \leq 7 \cdot \sqrt{\delta}$
- *if* $||ustat_k(w) - ustat_k(w')||_1 \leq \delta$*, then* $\mathsf{dist}(w, w') \leq \ 7 \cdot \delta$

2.2. Markov Decision Processes and Probabilistic Automata

2.2.1. Definitions

All the MDPs and automata are on a finite alphabet Σ. If S is finite set, we write $\Delta(S)$ for the set of distributions on S.

Definition 2. *A Markov Decision Process (MDP) is a triple $\mathcal{S} = (S, \Sigma, P)$ where S is a finite set of states, Σ is a set of* actions, *and $P : S \times \Sigma \times S \to [0; 1]$ is the transition relation. The value $P(s, a, t)$, also written $P(t|s, a)$, is the probability to arrive in t in one step when the current state is s and action $a \in \Sigma$ is chosen for the transition.*

Our MDPs do not have rewards, as we only study properties of their traces. If action a is not allowed from state s, $P(t|s, a) = 0$ for all $t \in S$. The initial state of the system is chosen randomly according to an initial probability distribution α on its state space. A *history*, or *run*, on $\mathcal{S}$ is a finite or infinite alternating sequence of states and actions, which begins with a state and ends with a state when finite. We write Ω^* for the set of finite runs, Ω for the set of infinite runs on $\mathcal{S}$. If $n \in \mathbb{N}$ and $r \in \Omega$ we write $r_{|n}$ for the sequence of the first $n-1$ state action couples in r and the n-th state in r. The *trace* $Tr(r)$ of a run r is the sequence of actions. If $n \in \mathbb{N}$, X_n and Y_n are the random variables on Ω which associate to a run r its n-th state and its n-th action. A *policy* on $\mathcal{S}$, see [30, 25], is a function $\sigma : \Omega^* \to \Delta(\Sigma)$. A policy resolves the non determinism of the system by choosing a distribution on the set of available actions from the last state of the given history. We write HR for the set of History dependent and Randomized policies. A policy is *deterministic* when for all history $h = (s_1, a_1, ..., , a_{i-1}, s_i)$ on $\mathcal{S}$, $\sigma(h) \in \Sigma$. We write HD for the set of History dependent Deterministic policies.

If $k \in \mathbb{N}$, a policy σ is said to have *memory* k if for any history $h = (s_1, a_1, ..., , a_{i-1}, s_i)$ of length at least k we have $\sigma((s_1, a_1, ..., , a_{i-1}, s_i)) = \sigma((s_{i-k}, a_{i-k}, ..., , a_{i-1}, s_i))$. We write $MR(k)$ for the set of Randomized policies with Memory less than k. A policy is *stationary*, or *memoryless*, if it has memory 0, i.e. for any history $h = (s_1, a_1, ..., , a_{i-1}, s_i)$ we have $\sigma(h) = \sigma(s_i)$. We write SR for the set of Stationary Randomized policies, and we write SD for the set of Stationary Deterministic policies.

A policy σ and an initial distribution α induce a probability distribution $\mathbb{P}^{\sigma,\alpha}$ on the σ-field $\mathcal{F}$ of Ω generated by the cones $C_\rho = \{r \in \Omega \mid r_{|\rho|} = \rho\}$, for $\rho \in \Omega^*$ (see [5, 30]). If the initial distribution α is concentrated on a state, that is if there exists $s \in S$ such that $\alpha(s) = 1$, we may write $\mathbb{P}^{\sigma,s}$ instead of $\mathbb{P}^{\sigma,\alpha}$.

Let $\mathcal{S}$ be an MDP and s be a state of $\mathcal{S}$. The set $Leave(s) \subseteq \Omega$ is the set of runs which do not cross s after a finite number of steps. That is,

$$Leave(s) = \{r \in \Omega \mid \exists k \in \mathbb{N} \; s.t. \; \forall l \geq k \;\; X_l(r) \neq s\}$$

Given a policy σ on $\mathcal{S}$, the state s is said to be *transient* under σ if $\mathbb{P}^{\sigma,s}[Leave(s)] = 1$. That is, s is transient for σ if with probability one the system visits s only a finite number of times, when the system is initiated in s and when policy σ is used.

As for Markov chains, the communication properties between the states of an MDP are important. An MDP is *weakly communicating*, see [22], if the set of states can be partitioned into a set of states that are accessible from each other (i.e., for any two states s and s' in that set, there exists a policy under which there is a positive probability to reach s' from s), and a set of states which are transient under *all* policies.

For a general MDP, there is a decomposition (see [6] and subsection 3.2.2) into maximal disjoint end components (MECs) $S_1, ..., S_l$ of the state space such that $S = S_0 \cup S_1 \cup ... \cup S_l$, where S_0 is the set of states which are transient for any policy on $\mathcal{S}$ (see section 3.2.2).

If $T \subseteq S$, $Reach(T)$ is the event: $\{r \in \Omega | \exists k \in \mathbb{N} \; s.t. \; X_k(r) \in T\}$, which is a measurable event [5, 30]. The maximal reachability problem [5, 7] asks, given a set $T \subset S$ of *destination states*, and an initial distribution α on S, for:

$$MaxReach_{\mathcal{S}}(\alpha, T) = Sup_{\sigma \in HR(\mathcal{S})} \mathbb{P}^{\alpha,\sigma}(Reach(T))$$

The following proposition is proved in [7].

Proposition 2 ([7])**.** *The value $MaxReach_{\mathcal{S}}(\alpha, T)$ can be computed in polynomial time. Moreover, the sup limit is reached using an optimal policy σ which is deterministic, and we can compute σ in polynomial time as well.*

A *probabilistic automaton (PA)*, [23], is an MDP with an extra set of final states $F \subseteq S$, usually given with a probability threshold $\lambda \in [0; 1]$ for the acceptance condition, which reads an input w_n. We will study if approximate algorithms for Membership and Equivalence problems generalize.

2.2.2. State-Action Frequencies

Statistics of runs on MDPs have been introduced in [18, 9, 22, 26] as random variables for the empirical state-action frequency vectors. We consider a run on an MDP with state space S as a sequence of couples in S and Σ. The statistics of a run will be taken on this alphabet.

Definition 3 (Expected state action frequency vector)**.** *Let σ be a policy on $\mathcal{S}$, $k \in \mathbb{N}$ and $T \geq 0$. The vector $\hat{x}_k^T$ is the random variable on the set of histories Ω, which associates to all $r \in \Omega$ the k-gram of its prefix of length T. That is, $\hat{x}_k^T = ustat_k(r_{|T})$. Given an initial distribution α, the* Expected state action frequency vector *$x_{\sigma,\alpha,k}^T$ is $\mathbb{E}_{\sigma,\alpha}[\hat{x}_k^T]$, i.e. the expectation of $\hat{x}_k^T$.*

We may forget the k in the notations when $k = 1$. The vector $x_{\sigma,\alpha,1}^T$ is in $[0; 1]^{S \times \Sigma}$, with non-negative components which sum to one. The value $x_{\sigma,\alpha,1}^T(s, a)$ is the expected frequency, up to time T, of taking state-action (s, a), given the initial distribution is α and the non determinism is resolved by σ. If σ is a policy on $\mathcal{S}$, then $x_{\sigma,\alpha,k}^\infty$ is the empty set if $x_{\sigma,\alpha,k}^T$ does not converge as $T \to +\infty$, and the limit point if $x_{\sigma,\alpha,k}^T$ converges. If K is a class of policies ($K = SD, HR$...), let:

$$H_k^K(\alpha) = \bigcup_{\sigma \in K} x_{\sigma,\alpha,k}^\infty$$

2.2.3. The polytope H for communicating MDPs

For $k = 1$, let $\mathcal{H}(\mathcal{S})$ be the set of vectors $x \in \Delta(S \times \Sigma)$ which satisfy for all $s' \in S$ the linear equations:

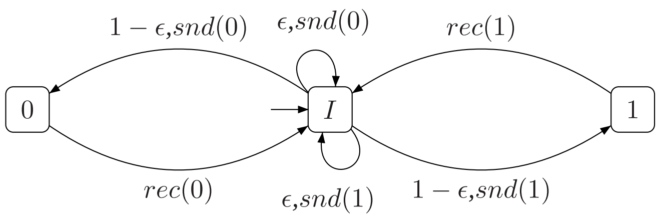
$$\sum_{s \in S} \sum_{a \in \Sigma} P(s'|s, a) \cdot x(s, a) = \sum_{a' \in \Sigma} x(s', a') \tag{1}$$

If $\overline{H}$ is the convex closure of H, the following proposition is an improvement by [18, 22, 14] of a first result of [9].

Proposition 3. *Let $\mathcal{S}$ be a weakly communicating MDP. Then for all distribution α on $\mathcal{S}$,*

$$H = H^{HR}(\alpha) = \overline{(H^{SD}(\alpha))},$$

Example 1 (A lossy communication channel)**.**



We will be interested in the actions appearing in the runs, on the alphabet $\Sigma = \{snd(1), snd(0), rec(1), rec(0)\}$. In order to get vectors of reasonable size, we take the statistics of the traces and do not consider the states. Non determinism is only present on state I, from which the system chooses betwen $snd(0)$ and $snd(1)$. It induces two possible stationary and deterministic policies. One chooses $snd(0)$ and leads to the limit statistic vector of order one on Σ: $x_1 = (snd(0) : \frac{1}{2-\epsilon}, rec(0) : \frac{1-\epsilon}{2-\epsilon}, snd(1) : 0, rec(1) : 0)$. The other chooses $snd(1)$ in I, and gives the symmetric point $x_2 = (snd(0) : 0, rec(0) : 0, snd(1) : \frac{1}{2-\epsilon}, rec(1) : \frac{1-\epsilon}{2-\epsilon})$. As the system is weakly communicating, the projection of $\mathcal{H}(\mathcal{S})$ in $\mathbb{R}^{\Sigma^k}$ is the segment between x_1 and x_2.

3. State-action frequencies and MDPs of higher order

3.1. k-statistic polytopes for MDPs

We first generalize the results of [18, 9, 22] to higher order statistics. In this section k is a natural number greater than 0. We fix an MDP $\mathcal{S}$ and α an initial distribution on $\mathcal{S}$. The analogous of Derman's theorem ([9], chapter 7), is not true any more when we consider statistics of higher orders: if $k \geq 2$, in general $\mathcal{H}_k^{HR}(\alpha)$ is not the convex hull of $\mathcal{H}_k^{SR}(\alpha)$. Still, we will see in the future that in that case $\mathcal{H}_k^{HR}(\alpha) = \overline{(\mathcal{H}_k^{MR(k)}(\alpha))}$.

Example 2. *Consider the following MDP:*

$$\mathcal{S} : \begin{tikzcd} s_1 \arrow[r, bend left, "{a,1}"] & s_2 \arrow[l, bend left, "{c,1}"] \arrow[loop right, "{b,1}"] \end{tikzcd} \quad s_1 \textit{ initial state.}$$

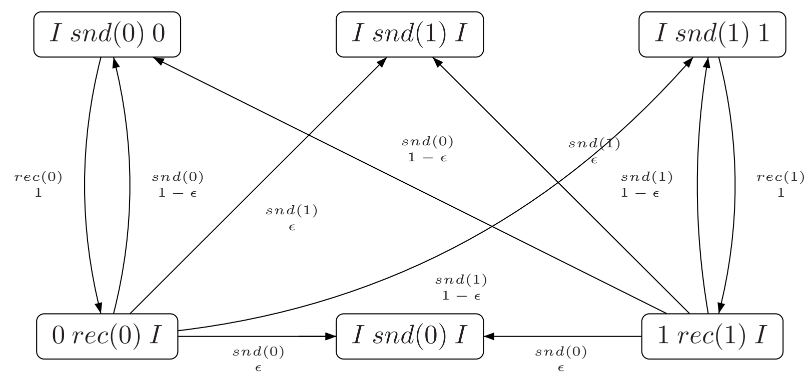
On $\mathcal{S}$, consider the policy $\sigma \in HR$ such that the choice on state s_2 depends on the history: if the system was in s_2 before, then σ chooses action c with probability one. If the system was in s_1 and just arrives in s_2, σ chooses action b with probability one. Then, $x^T_{\sigma,\alpha,2}$ converges to a point $x \in \Delta((S \times \Sigma)^2)$, such that $x[s_1 a s_2 b] = x[s_2 b s_2 c] = x[s_2 c s_1 a] = 1/3$, and all the other coordinates are zero. If $x \in H_2^{SR}(\alpha)$, $x[s_1 a s_2 b] > 0$ and $x[s_2 c s_1 a] > 0$ implies $x[s_1 a s_2 c] > 0$. This proves that $x \notin H_2^{SR}(\alpha)$, and in fact we have $d_{L_1}(x, H_2^{SR}(\alpha)) \geq 1/6$, which proves that we do not have $H_2^{HR}(\alpha) = \overline{H_2^{SR}(\alpha)}$. We will prove that in that case, $H_2^{HR}(\alpha) = \overline{H_2^{MR(2)}(\alpha)}$.

We describe now the construction of the k-th self product of an MDP. Our goal is the following: the state-action frequency vectors on the k-th self product $\mathcal{S}^k$ of $\mathcal{S}$ should correspond to the order k state-action frequency vectors on $\mathcal{S}$.

Definition 4. *The k self product of an MDP $\mathcal{S} = (S, \Sigma, P)$ is the MDP $\mathcal{S}^k = (S', \Sigma, P')$ where:*
$S' = (\prod_{i=1}^{k-1} S \times \Sigma) \times S$,
If $t' = (t_1, b_1 ..., b_{k-1}, t_k)$ and $s' = (s_1, a_1 ..., a_{k-1}, s_k)$ are in S' and $a \in \Sigma$,

$$P'(t'|s', a) = \begin{cases} P(t_k|s_k, a) \ \textit{if } (t_1, b_1 ..., b_{k-2}, t_{k-1}) = \\ \qquad (s_2, a_2 ..., a_{k-1}, s_k) \textit{ and } a = b_{k-1} \\ 0 \textit{ if not.} \end{cases}$$

Example 3. *As an example for the last definition we present the second power of the lossy channel of example 1:*



If $\mathcal{S}$ is a communicating MDP, then $\mathcal{S}^k$ is communicating as well. A run on $\mathcal{S}^k$ is a sequence of couples in S' and Σ, which can be seen as a sequence of couples in S and Σ, i.e. a run on $\mathcal{S}$. Also, given a policy σ on $\mathcal{S}$, there is a naturally associated policy σ' on $\mathcal{S}^k$, which takes the same actions given the same histories. However, several policies on $\mathcal{S}$ may be associated to the same policy σ' on $\mathcal{S}^k$. Indeed a policy on $\mathcal{S}^k$ does not consider runs on $\mathcal{S}$ of length smaller than k. We can use the notion of state-action statistic vector for the MDP $\mathcal{S}^k$. If σ' is a policy on $\mathcal{S}^k$, $T \geq 1$, and α' is an initial distribution on $\mathcal{S}^k$, then $x^T_{\sigma',\alpha',1}(\mathcal{S}^k) \in \mathbb{R}^{(S' \times \Sigma)} = \mathbb{R}^{(S\times\Sigma)^k}$.

An initial distribution α and a policy σ on $\mathcal{S}$ induce a unique initial distribution $\alpha'(\sigma, \alpha)$ on $\mathcal{S}^k$, such that for all state $(s_1, a_1..., s_k)$ in $\mathcal{S}^k$,

$$\alpha'(\sigma, \alpha)((s_1, a_1..., s_k)) = \mathbb{P}^{\sigma,\alpha}(C_{s_1,a_1...,s_k}).$$

Lemma 1. *Let $\mathcal{S}$ be an MDP, α an initial distribution, $k \geq 1$, and σ a policy on $\mathcal{S}$. We write σ' for the policy on $\mathcal{S}^k$ associated to σ, and $\alpha' = \alpha'(\sigma, \alpha)$ for the initial distribution on $\mathcal{S}^k$ associated to σ and α. Then for all $T \geq 1$, we have:*

$$x^{T+k-1}_{\sigma,\alpha,k}(\mathcal{S}) = x^T_{\sigma',\alpha',1}(\mathcal{S}^k).$$

Proof. Both vectors are in $\mathbb{R}^{(S\times\Sigma)^k}$. Write $\mathcal{S}^k = (S', \Sigma, P')$. Recall $S' = (S \times \Sigma)^{k-1} \times S$. If $w = (s_0, a_0..., a_{k-2}, s_{k-1}) \in S'$, we write $Prec(w)$ for the set of states in $\mathcal{S}^k$ of the form $(s_{-1}, a_{-1}, s_0, a_0, ..., a_{k-3}, s_{k-2})$. If $b \in \Sigma$, we consider (w, b) as a couple in $S' \times \Sigma$, or as a sequence of k couples $(s_0, a_0), ..., (s_{k-1}, b)$ in $(S \times \Sigma)^k$, depending on the context. If $n \in \mathbb{N}$, $X'_n : \Omega \to S'$ associates to a run r the subword with k states and $k-1$ letters which starts after n steps in r. (For instance, if $r = (s_1, a_1, s_2, a_2, s_3, a_3, s_4)$, $k = 2$ and $n = 2$, $X'_2(r) = (s_2, a_2, s_3)$). If fact, X'_n associates to a run on $\mathcal{S}^k$ its n-th state in S'. If $n \in \mathbb{N}$, $Y'_n : \Omega \to \Sigma$ associates to a run $r \in \Omega$ its n-th action when we consider r as a run on $\mathcal{S}^k$, hence $Y'_n(r)$ is the $n + k - 1$ action on r (In the previous example, $Y_2(r) = a_3$).

Let $w = (s_0, a_0..., a_{k-2}, s_{k-1}) \in S'$, and $b \in \Sigma$. By definition,

$$x^{T+k-1}_{\sigma,\alpha,k}(\mathcal{S})[w, b] = \tfrac{1}{T+1} \textstyle\sum_{t=0}^T \mathbb{P}^{\sigma,\alpha}(X'_t = w, Y'_t = b)$$
$$x^T_{\sigma',\alpha',1}(\mathcal{S}^k)[w, b] = \tfrac{1}{T+1} \textstyle\sum_{t=0}^T \mathbb{P}^{\sigma',\alpha'}(X'_t = w, Y'_t = b)$$

To prove the lemma, we prove by induction on $t \in \mathbb{N}$ that for all $w = (s_0, a_0..., a_{k-2}, s_{k-1}) \in S'$ and $b \in \Sigma$,

$$\forall t \geq 0, \mathbb{P}^{\sigma,\alpha}(X'_t = w, Y'_t = b) = \mathbb{P}^{\sigma',\alpha'}(X'_t = w, Y'_t = b).$$

- Suppose $t = 0$, then $\mathbb{P}^{\sigma',\alpha'}(X'_0 = w, Y'_0 = b) = \alpha'(w) \cdot \sigma'(w)(b)$. By definition, $\alpha'(w) = \mathbb{P}^{\sigma,\alpha}(X'_0 = w)$ and $\sigma'(w) = \sigma(w)$, hence $\mathbb{P}^{\sigma',\alpha'}(X'_0 = w, Y'_0 = b) = \mathbb{P}^{\sigma,\alpha}(X'_0 = w) \cdot \sigma(w)(b) = \mathbb{P}^{\sigma,\alpha}(X'_0 = w, Y'_0 = b)$.

- Suppose the lemma true for $t \geq 0$. By definition,

  $$\mathbb{P}^{\sigma',\alpha'}(X'_{t+1} = w, Y'_{t+1} = b) = \sum_{w' \in Prec(w)} \mathbb{P}^{\sigma',\alpha'}(X'_t = w', Y'_t = a_{k-2}) \times P'(w|w', a_{k-2}) \cdot \sigma'(w)(b)$$

  We write $w' = (s_{-1}, a_{-1}, a_0, s_0, ..., a_{k-3}, s_{k-2})$ when $w' \in Prec(w)$. Then, by the induction hypothesis and the definition of P' and σ',

  $$\begin{aligned}\mathbb{P}^{\sigma',\alpha'}(X'_{t+1} = w, Y'_{t+1} = b) = {} & \textstyle\sum_{(s_{-1},\, a_{-1})\, \in\, S \times \Sigma} \mathbb{P}^{\sigma,\alpha}(X'_t = (s_{-1}, a_{-1}, .., s_{k-2}), Y'_t = a_{k-2}) \\ & \times P(s_{k-1}|s_{k-2}, a_{k-2}) \cdot \sigma(w)(b)\end{aligned}$$

  That is,

  $$\mathbb{P}^{\sigma',\alpha'}(X'_{t+1} = w, Y'_{t+1} = b) = \sum_{s_{-1} \in S, a_{-1} \in \Sigma} \mathbb{P}^{\sigma,\alpha}(X'_t = (s_{-1}, ..., s_{k-2}), Y'_t = a_{k-2}, X_{t+1} = s_{k-1}, Y_{t+1} = b)$$

  I.e. $\mathbb{P}^{\sigma',\alpha'}(X'_{t+1} = w, Y'_{t+1} = b) = \mathbb{P}^{\sigma,\alpha}(X'_{t+1} = w, Y'_{t+1} = b)$.

□

The following theorem and corollary generalize proposition 3 to the context of higher order statistics.

Theorem 1. *Let $\mathcal{S}$ be a weakly communicating MDP and $k \geq 1$. Then $H^{HR}_1(\mathcal{S}^k) = H^{HR}_k(\mathcal{S})$, and for all initial distribution α on $\mathcal{S}$,*

$$\bigcup_{\sigma\in MR(k)(\mathcal{S})} H^{SR}(\mathcal{S}^k)(\alpha'(\sigma,\alpha)) = H_k^{MR(k)}(\mathcal{S})(\alpha)$$

Proof.

- For the first statement, as $\mathcal{S}$ is communicating, $\mathcal{S}^k$ is communicating as well, and by proposition 3, $H^{HR}(\mathcal{S}^k)$ and $H_k^{HR}(\mathcal{S})$ do not depend on the initial distributions on $\mathcal{S}$ or $\mathcal{S}^k$. We have $H_k^{HR}(\mathcal{S}) = \bigcup_{\sigma\in HR} x^\infty_{\sigma,\alpha,k}(\mathcal{S})$. By lemma 1, $x^\infty_{\sigma,\alpha,k}(\mathcal{S}) = x^\infty_{\sigma',\alpha',1}(\mathcal{S}^k)$, where σ' is the policy on $\mathcal{S}^k$ associated to σ, and α' is the initial distribution on $\mathcal{S}^k$ associated to α and σ. This proves $H_k^{HR}(\mathcal{S}) \subseteq H^{HR}(\mathcal{S}^k)$.

  For the other inclusion, let $x \in H^{HR}(\mathcal{S}^k)$. Let $\sigma_0 \in MR(k)(\mathcal{S})$, α an initial distribution on $\mathcal{S}$, and $\alpha' = \alpha'(\sigma_0,\alpha)$ the initial distribution on $\mathcal{S}^k$ associated to α and σ_0. We have that $H^{HR}(\mathcal{S}^k)$ does not depend on the initial distribution, and in particular $H^{HR}(\mathcal{S}^k) = H^{HR}(\mathcal{S}^k)(\alpha')$. Hence there exists $\sigma' \in HR(\mathcal{S}^k)$ such that $x = x^\infty_{\sigma',\alpha',1}$. Let σ be a policy on $\mathcal{S}$ such that on histories of length smaller than k, σ coincides with σ_0, and σ coincides with σ' on histories with greater length. By construction, α' is also the initial distribution on $\mathcal{S}^k$ which is associated to α and σ . By lemma 1, $x^\infty_{\sigma,\alpha,k}(\mathcal{S}) = x^\infty_{\sigma',\alpha',1} = x$. This proves $H^{HR}(\mathcal{S}^k) \subseteq H_k^{HR}(\mathcal{S})$.

- For the second statement, the proof that $H_k^{MR(k)}(\mathcal{S})(\alpha)$ is a subset of $\bigcup_{\sigma_0\in MR(k)(\mathcal{S})} H^{SR}(\mathcal{S}^k)(\alpha'((\sigma_0,\alpha)))$ is analogous to the previous argument. We just have to point out that if $\sigma \in MR(k)(\mathcal{S})$, then the associated $\sigma' \in HR(\mathcal{S}^k)$ is in fact in $SR(\mathcal{S}^k)$.

  For the other inclusion, let $\sigma_0 \in MR(k)(\mathcal{S})$, let α' be the initial distribution on $\mathcal{S}^k$ associated to α and σ_0, and let $x \in H^{SR}(\mathcal{S}^k)(\alpha')$. Let σ be the policy on $\mathcal{S}$ associated to σ_0, σ' as before. Since $\sigma' \in SR(\mathcal{S}^k)$, we have $\sigma \in MR_k(\mathcal{S})$. Moreover, by lemma 1, $x^\infty_{\sigma,\alpha,k}(\mathcal{S}) = x^\infty_{\sigma',\alpha',1} = x$. This proves $x \in H_k^{MR(k)}(\mathcal{S})(\alpha)$

□

Corollary 1. *If $\mathcal{S}$ is weakly communicating,*

$$H_1(\mathcal{S}^k) = H_1^{HR}(\mathcal{S}^k) = \mathcal{H}_k^{HR}(\mathcal{S}) = \overline{[H_1^{SR}(\mathcal{S}^k)]} = \overline{[H_k^{MR(k)}(\mathcal{S})]}.$$

Proof.

- $H_1(\mathcal{S}^k) = H_1^{HR}(\mathcal{S}^k)$ comes from proposition 3.
- $H_1^{HR}(\mathcal{S}^k) = \mathcal{H}_k^{HR}(\mathcal{S})$ comes from the last theorem. ($H_1^{HR}(\mathcal{S}^k(\sigma_0))$ is independent of σ, since $H_1^{HR}(\mathcal{S}^k)(\alpha'$) is independent of the initial distribution α'.)
- $H_1^{HR}(\mathcal{S}^k) = \overline{[H_1^{SR}(\mathcal{S}^k)]}$ comes from proposition 3 applied to the communicating MDP $\mathcal{S}^k$.
- $\overline{[H_1^{SR}(\mathcal{S}^k)]} = \overline{[H_k^{MR(k)}(\mathcal{S})]}$ comes from the last theorem.

□

So far the polytope associated to an MDP lies in a vector space whose dimension depends on the state space of the system. We eliminate this dependence, in order to be able to compare systems with very different state spaces. We introduce the linear projection $\pi : \mathbb{R}^{(S\times\Sigma)^k} \to \mathbb{R}^{\Sigma^k}$ such that if $x \in \mathbb{R}^{(S\times\Sigma)^k}$, on a component $v \in \Sigma^k$ we have

$$\pi(x)[v] = \sum_{u\in(S\times\Sigma)^k \ s.t.\ Tr(u)=v} x[u] \tag{2}$$

In the future, if $i,k \in \mathbb{N}$, we write $H_k^i(\mathcal{S})$ for $H_k^{MR(i)}(\mathcal{S})$, and $\Pi_k^i(\mathcal{S})$ for $\pi(H_k^i(\mathcal{S}))$.

3.2. Distance between MDPs

3.2.1. The weakly communicating case

In this paragraph we compute the distance between a statistic vector and a polytope $\Pi_k^k(\mathcal{S})$, and define a distance between MDPs naturally associated with their polytopes.

Definition 5 (Distance $d_k(x, \mathcal{S})$)**.** *If $\mathcal{S}$ is a weakly communicating MDP, $k \in \mathbb{N}$, and $x \in \mathbb{R}^{\Sigma^k}$, let*

$$d_k(x, \mathcal{S}) = Inf_{y \in \Pi_k^k(\mathcal{S})} ||x - y||_1$$

The following proposition shows that $d_k(x, \mathcal{S})$ can be computed in time polynomial in $(|S| \cdot |\Sigma|)^k$. The crucial point is that we get non exponential bounds because the polytope $H_k^k(\mathcal{S})$ is characterized by a number of linear equations polynomial in $|\mathcal{S}|$ (see equation 1), even if it may have an exponential number of extremal points (corresponding to the exponential number of possible stationary and deterministic policies).

Proposition 4. *The distance $d_k(x, \mathcal{S})$ can be computed in time polynomial in $(|S| \cdot |\Sigma|)^k$.*

Proof. Let C be the matrix of size $|\Sigma|^k \times |S \times \Sigma|^k$ corresponding to the projection $\pi : \mathbb{R}^{(S \times \Sigma)^k} \to \mathbb{R}^{\Sigma^k}$. (see equation 2). Using the linear characterization of the polytope $H_k^k(\mathcal{S})$ (see equation 1), we know that we can find A, a square matrix of size $(|S| \cdot |\Sigma|)^k$, and B a column vector of size $(|S| \cdot |\Sigma|)^k$ such that:

$$\Pi_k^k(\mathcal{S}) = \{C \cdot X | \begin{pmatrix} A \cdot X = B \\ X \geq 0,\ X \in \mathbb{R}^{(|S| \cdot |\Sigma|)^k} \end{pmatrix}\} \tag{1}$$

Now,

$$d_k(x, \mathcal{S}) = \begin{pmatrix} Inf ||Y - x||_1 \\ Y = C \cdot X \\ A \cdot X = B,\ X \geq 0 \end{pmatrix}$$

If E is the column vector in $\mathbb{R}^{\Sigma^k}$ with only ones on its components, we get:

$$d_k(x, \mathcal{S}) = \begin{pmatrix} Inf_{Z \in \mathbb{R}^{\Sigma^k}} . {}^t Z \cdot E \\ Z \geq C \cdot X - x \\ Z \geq -C \cdot X + x \\ A \cdot X = B \\ X \geq 0 \end{pmatrix}$$

This is a regular form of linear programming, with variable $X \in \mathbb{R}^{(S \times \Sigma)^k}$ and $Z \in \mathbb{R}^{\Sigma^k}$, and with constraints of size $(|S| \cdot |\Sigma|)^k$. Hence $d_k(x, \mathcal{S})$ can be computed in $PTIME((|S| \cdot |\Sigma|)^k)$.

□

We define now the k-th order distance between two weakly communicating MDPs as the Hausdorff distance between the associated order k polytopes.

Definition 6 (The distance $d_k(\mathcal{S}_1, \mathcal{S}_2)$)**.** *If $\mathcal{S}_1$ and $\mathcal{S}_2$ are two weakly communicating MDPs let $d_k(\mathcal{S}_1, \mathcal{S}_2)$ be the Hausdorff distance (with respect to the norm L_1) between their polytopes for k statistics:*

$$d_k(\mathcal{S}_1, \mathcal{S}_2) = \frac{Sup_{x \in \Pi_k^k(\mathcal{S}_1)} Inf_{y \in \Pi_k^k(\mathcal{S}_2)} ||x - y||_1}{2} + \frac{Sup_{x \in \Pi_k^k(\mathcal{S}_2)} Inf_{y \in \Pi_k^k(\mathcal{S}_1)} ||x - y||_1}{2}$$

For instance, with two lossy channels $\mathcal{S}_1, \mathcal{S}_2$ with respective parameters ϵ_1 and ϵ_2, it is not difficult to see that the distance between the order one polytopes is $|\epsilon_1 - \epsilon_2|/2$. (It is $|\epsilon_1 - \epsilon_2|$ for the metric of [10], for instance, and we will see that in general our distance is far from the metric of [10]). This distance is difficult to compute and hard to approximate to any fixed ratio $\beta > 0$, but we give a $|\Sigma^k|$-approximation ([15] proves that even the L_1-diameter of a polytope is not computable in PTIME, and that it is not well approximable). We use the fact that the Hausdorff distance with the norm L_∞ is computable using a linear program of polynomial size, to approximate the L_1 Hausdorff distance within a factor $|\Sigma|^k$.

Proposition 5. *Suppose $\mathcal{S}_1$ and $\mathcal{S}_2$ weakly communicating. Then we can compute the distance $d_k(\mathcal{S}_1, \mathcal{S}_2)$ within a factor $|\Sigma|^k$ in $PTIME(((|S_1| + |S_2|) \cdot |\Sigma|)^k \cdot |\Sigma|)$.*

Proof. We compute

$$d_k^\infty(\mathcal{S}_1,\mathcal{S}_2) = \frac{Sup_{x\in\Pi_k^k(\mathcal{S}_1)} Inf_{y\in\Pi_k^k(\mathcal{S}_2)} ||x-y||_\infty}{2} + \frac{Sup_{x\in\Pi_k^k(\mathcal{S}_2)} Inf_{y\in\Pi_k^k(\mathcal{S}_1)} ||x-y||_\infty}{2}$$

in time polynomial in $|\Sigma|^k \cdot |S\times\Sigma|^k$, and use the fact that $d_k^\infty(\mathcal{S}_1,\mathcal{S}_2) \le d_k(\mathcal{S}_1,\mathcal{S}_2) \le |\Sigma|^k \cdot d_k^\infty(\mathcal{S}_1,\mathcal{S}_2)$ to approximate the L_1 Hausdorff distance. In order to compute $Sup_{x\in\Pi_k^k(\mathcal{S}_1)} Inf_{y\in\Pi_k^k(\mathcal{S}_2)} ||x-y||_\infty$, we just have to compute the components of the extremal points of the polytopes $\Pi_k^k(\mathcal{S}_1)$ and $\Pi_k^k(\mathcal{S}_2)$ on all the coordinates $u\in\Sigma^k$. If $u\in\Sigma^k$, the values $Max_{x\in\Pi_k^k(\mathcal{S}_1)} x[u]$ and $Min_{x\in\Pi_k^k(\mathcal{S}_1)} x[u]$ can be computed using linear programs of size $(|S_1|\cdot|\Sigma|)^k$:

$$Max_{x\in\Pi_k^k(\mathcal{S}_1)} x[u] = Max_{x\in\Pi_k^k(\mathcal{S}_1)} x\cdot E_u,$$
$$Min_{x\in\Pi_k^k(\mathcal{S}_1)} x[u] = -Max_{x\in\Pi_k^k(\mathcal{S}_1)} (-x\cdot E_u),$$

where E_u is the vector in $\mathbb{R}^{\Sigma^k}$ with value 1 on coordinate u, and 0 elsewhere. The linear constraints $x\in\Pi_k^k(\mathcal{S}_1)$ are given by the equations in (1). The two equations are valid since all the coordinates of any $x\in\Pi_k^k(\mathcal{S}_1)$ are non-negative. For $u\in\Sigma^k$ and $i\in\{1,2\}$, let

$$\begin{cases} t_i^+(u) = Max_{x\in\Pi_k^k(\mathcal{S}_i)} x[u] \\ t_i^-(u) = Min_{x\in\Pi_k^k(\mathcal{S}_i)} x[u] \end{cases}$$

Then, by a simple comparison between the extremal coordinates in $\Pi_k^k(\mathcal{S}_1)$ and $\Pi_k^k(\mathcal{S}_2)$,

$$Sup_{x\in\Pi_k^k(\mathcal{S}_1)} Inf_{y\in\Pi_k^k(\mathcal{S}_2)} |x[u]-y[u]| = Max[t_1^+ - t_2^+, t_2^- - t_1^-, 0].$$

Finally, since

$$Sup_{x\in\Pi_k^k(\mathcal{S}_1)} Inf_{y\in\Pi_k^k(\mathcal{S}_2)} \|x-y\|_\infty = Max_{u\in\Sigma^k} Sup_{x\in\Pi_k^k(\mathcal{S}_1)} Inf_{y\in\Pi_k^k(\mathcal{S}_2)} |x[u]-y[u]|,$$

we can apply $|\Sigma|^k$ times the previous algorithm to each coordinate $u\in\Sigma^k$, which proves the result. □

Instead of approximating $d_k(\mathcal{S}_1,\mathcal{S}_2)$ by approximating the L_1 norm by the L_∞ norm, we could have discretized the problem. For instance we could have discretized the set of vectors in $\mathbb{R}^{\Sigma^k}$ with positive components which sum to one, with a grid of radius $\epsilon>0$. For each point y of the grid we can compute the L_1 distance d between the affine subspaces $\pi_1^{-1}(y)$ and $H_k^k(\mathcal{S}_1)$ in $PTIME(|(S_1\cdot\Sigma)|^k)$ with a linear program. Since d is also the distance between y and $\Pi_k^k(\mathcal{S}_1)$, we can compute the last distance in $PTIME(|(S_1\cdot\Sigma)|^k)$. The same holds for the distance between $\pi_2^{-1}(y)$ and $H_k^k(\mathcal{S}_2)\|$. The π_1 and π_2 are the projection from $\mathbb{R}^{(S_1\times\Sigma)^k}$ and $\mathbb{R}^{(S_2\times\Sigma)^k}$ to $\mathbb{R}^{\Sigma^k}$. Since there exists $(1/\epsilon)^{|\Sigma|^k}$ vertices in the grid, we would get an ϵ-approximation of $d_k(\mathcal{S}_1,\mathcal{S}_2)$ in time $(1/\epsilon)^{|\Sigma|^k}\cdot P((|(S_1+S_2)\cdot\Sigma|^k))$, where P is a polynomial.

3.2.2. The non weakly communicating case

In general, for a non weakly communicating MDP, the set of the limit statistics is not a polytope, and it is not characterized by a set of linear equations. Instead it is the union of the polytopes associated to the maximal end components of the MDP, that we present now. We need some previous notions, presented in [6].

Let $\mathcal{S}=(S,\Sigma,P)$ be an MDP. Given $s\in S$ and $a\in\Sigma$, let $Succ(s,a)=\{t\in S|P(t|s,a)>0$. A *sub-MDP* of $\mathcal{S}$ is a pair (C,D) where $C\subset S$ and D is a function that associates to each $s\in C$ a set $D(s)\subseteq\Sigma$ of actions. Given an MDP (C,D) of $\mathcal{S}$, the relation $\rho_{(C,D)}$ is defined as:

$$\rho_{(C,D)} = \{(s,t)\in C\times S|\exists a\in\Sigma\ s.t.\ P(t|s,a)>0\}$$

Definition 7 (End component of an MDP ([6])). *A sub-MDP (C,D) of $\mathcal{S}$ is an* end component *if:*

- *$Succ(s,a)\subseteq C$ for all $s\in C$ and $a\in D(s)$;*
- *The graph $(C,\rho_{(C,D)})$ is strongly connected.*

We say that an end component (C, D) of $\mathcal{S}$ is *maximal* if there is no other end component (C', D') of $\mathcal{S}$ different from (C, D) and such that $C \subseteq C'$ and $\rho_{(C,D)} \subseteq \rho_{(C',D')}$.

A state $s \in S$ is *transient for a policy* σ on $\mathcal{S}$ if with probability one on the set of runs, when started from s, after a finite number of steps a run will leave s and never reach it again in the future. The *fundamental theorem of end components* of [6] implies that the set S_0 of states in S which belong to no maximal end component is exactly the set of states which are transient for all policies. Thus, we can decompose the state space S of a given MDP $\mathcal{S}$ into maximal disjoint end components (MECs) $S_1, ..., S_l$ of the state space such that $S = S_0 \cup S_1 \cup ... \cup S_l$, where S_0 is the set of states which are transient for any policy on $\mathcal{S}$.

Given two general MDPs $\mathcal{S}_1, \alpha_1$ and $\mathcal{S}_2, \alpha_2$, we partition their state spaces into MECs. Write $S_j = S_j^0 \cup S_j^1 \cup ... \cup S_j^{l_j}$, for $j = 1, 2$. A sub-MDP is in particular an MDP, hence for all $j \in \{1, 2\}$ and $i \in [1; l_j]$ we can consider $\mathcal{S}_j^i$, the MDP associated to the MEC with state space S_j^i. If $x \in \mathbb{R}^{\Sigma^k}$, $\delta \geq 0$ and $j \in \{1, 2\}$, $V_j(x, \delta)$ is the union of the state spaces of the $\mathcal{S}_j^i$'s whose polytope is δ-close to x. That is:

$$V_j(x, \delta) = \bigcup \{S_i^j | d_k(x, \mathcal{S}_j^i) \leq \delta\}.$$

Now we have two parameters to differentiate two systems: one to measure the difference of the polytopes, and a second one to measure the different probabilities to reach the polytopes. This gives a notion of simulation.

Definition 8 ((ϵ, δ)-simulation). *If $\epsilon, \delta \in [0; 1]$, $\mathcal{S}_1, \alpha_1$ will be said to be (ϵ, δ)-simulated by $\mathcal{S}_2, \alpha_2$ for order k statistics, written $\mathcal{S}_1, \alpha_1 \prec_{(\epsilon,\delta)}^k \mathcal{S}_2, \alpha_2$, if for all $x \in \mathbb{R}^{\Sigma^k}$,*

$$MaxReach_1(\alpha_1, V_1(x, 0)) \leq MaxReach_2(\alpha_2, V_2(x, \epsilon)) + \delta$$

That is, $\mathcal{S}_2, \alpha_2$ (ϵ, δ)-simulates $\mathcal{S}_1, \alpha_1$ if we can generate traces on the second system which are ϵ close to the traces generated on the first, with probabilities δ close. This notion induces naturally two pseudo metrics, which extend d_k on the set of general MDPs.

Definition 9 (Pseudometrics d_k and $d_k^{\prec}$). *Let $d_k^{\prec}(\mathcal{S}_1, \mathcal{S}_2) = \mathrm{Inf}\{\epsilon > 0 | \mathcal{S}_1, \alpha_1 \prec_{(\epsilon,\epsilon)} \mathcal{S}_2, \alpha_2\}$, and $d_k(\mathcal{S}_1, \mathcal{S}_2) = Min(d_k^{\prec}(\mathcal{S}_1, \mathcal{S}_2), d_k^{\prec}(\mathcal{S}_2, \mathcal{S}_1))$.*

Notice that d_k is symmetric, whereas $d_k^{\prec}$ is not in general. In the case of weakly communicating MDPs, the two definitions for $d_k(\mathcal{S}_1, \mathcal{S}_2)$ coincide. Both $d_k^{\prec}(\mathcal{S}_1, \mathcal{S}_2)$ and $d_k(\mathcal{S}_1, \mathcal{S}_2)$ can be approximated in $PTIME((|S_1| \cdot |\Sigma| + |S_1| \cdot |\Sigma|)^k)$, within a factor $|\Sigma|^k$.

Our distances between MDPs cannot be compared to the metrics defined in [29, 10]. These metrics are in fact metrics between states of a given system, which may induce metrics between systems, by taking their initial states. In our approach, we do not rely on states, and the distance between two communicating systems does not depend on the initial distributions.

4. Problems on MDPs

We consider two classes of problems concerning generation of words by MDPs. First "membership" type problems, where we want to know if a word or a statistic vector can be generated by an MDP, and second "comparison" type problems, where the input consists in two MDPs that we want to compare. In the following, if $\epsilon \geq 0$, $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $x \in \mathbb{R}^{\Sigma^k}$, let

$$B_n^k(x, \epsilon) = \{r \in \Omega | \|ustat_k(r_{|n}) - x\| \leq \epsilon\}$$

As well, if $w \in \Sigma^n$, let

$$B_n^k(w, \epsilon) = \{r \in \Omega | \|(ustat_k(r_{|n}) - ustat_k(w)\| \leq \epsilon\}$$

We may forget the ϵ in the notation when $\epsilon = 0$.

4.1. The statistical λ-membership problem.

Many standard problems in the context of MDP optimization, (see [4, 21]), can be presented as follows. We are given an MDP $\mathcal{S}$, a length $n \in \mathbb{N}$, a probability threshold $\lambda \in [0;1]$, and an objective x. The question is to find a policy σ on $\mathcal{S}$ such that the probability of the set of runs of length n which satisfy the objective is greater than λ. Suppose that our objective is to generate words with a given statistic. We prove that the associated decision problem is PSPACE-hard, even when we consider order one statistics.

Problem 1 (Statistical λ-membership problem)**.**
A threshold $\lambda \in [0;1]$ is fixed.
Input: *An MDP $\mathcal{S}$ on Σ of size $n \in \mathbb{N}$, $w \in \Sigma^*$ of length $O(n)$.*
Question: *$\exists \sigma \in HR(\mathcal{S})$ such that $\mathbb{P}^\sigma(B^1_n(w)) \geq \lambda$?*

We prove that the problem 1 is PSPACE hard. We reduce a variation of SSAT to this problem. We recall first the definition of the SSAT problem.

Problem 2 (Stochastic satisfiability problem (SSAT))**.**
Input: *$p_1, q_1, p_2, q_2, ..., p_m, q_m$ boolean variables, $\phi = C_1 \wedge ... \wedge C_l$ a 3-CNF formula, where the C_i are 3-disjunctions.*
Question: *Can we choose the value of p_1 in $\{0,1\}$ such that, if q_1 is taken with uniform probability in $\{0,1\}$, we can choose $p_2 \in \{0,1\}$ such that if q_2 is chosen randomly, etc..., such that the probability to satisfy ϕ is at least $1/2$. In our notation: $\exists p_1 R q_1 \exists p_2 R q_2 ... \exists p_m R q_m \; \mathbb{P}(\phi = true) \geq 1/2$?*

Recall from [20] that problem SSAT is PSPACE-complete.

Proposition 6. *The statistical λ-membership problem is PSPACE hard.*

Proof. The proof is in two steps: first we prove the PSPACE hardness of the same problem when we ask the policy to be deterministic. In a second step, we present a generalization of the SSAT problem, where all the variables are chosen randomly. We prove that this generalization is computationally PTIME-equivalent to SSAT, which can be of independent interest, and next reduce it to our problem Statistical λ-membership, which proves it PSPACE hardness.

First reduction.
We reduce SSAT to the following variation of problem 1: Given an MDP $\mathcal{S}$ on Σ of size $n \in \mathbb{N}$, and $w \in \Sigma^*$ of length $O(n)$, does there exist a *deterministic* policy $\sigma \in HD(\mathcal{S})$ such that $\mathbb{P}^\sigma(B^1_n(w)) \geq \lambda$?

Let $\phi = C_1 \wedge ... \wedge C_l$, where the C_i are 3-disjunctions on variables $p_1, q_1, p_2, q_2, ..., p_m, q_m$, be an instance of SSAT. We build an MDP $\mathcal{S}$ and a word w, of sizes polynomial in m, l, such that the statistical λ-membership problem is satisfied on input $(\mathcal{S}, w)$ with deterministic policies iff SSAT is satisfied.

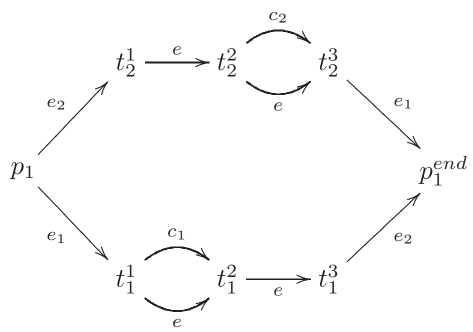
Let Σ be the union of the $l+3$ different letters $e, e_1, e_2, c_1, c_2, ..., c_l$. The word w will contain exactly once each letters $c_1, ..., c_l$, plus a set of extra buffer letters. Intuitively, a policy on $\mathcal{S}$ will give a valuation for the p_i, and the letter c_i will belong to the trace of a run if the valuation associated to the underlying policy satisfies the clause C_i. Our MDP $\mathcal{S}$ is built as a sequence of basic units, which are directed acyclic graphs (DAGs), one associated with each variables p_i and q_i. Each unit reflects whether the associated variable is existential or random, and whether it occurs positively or negatively in $C_1, ..., C_l$. Thus, the depth of each DAG is proportional to l, the number of clauses. For all variable $v \in \{p_i, q_i\}$, consider the following DAG, that we call U_v:

- The root is labelled by v. The root has two children, labelled t^1_1, t^1_2.
  - If $v \in \{p_1, ..., p_m\}$, then from v there exists a transition with label e_1 to state t^1_1, and a transition with label e_2 to t^1_2.
  - If $v \in \{q_1, ..., q_m\}$, then from v there exists a probabilistic transition with label e to states t^1_1 and t^1_2, with same probability $1/2$ for these two states.
- If v appears positively in C_1 then
  - t^1_1 has a unique child t^2_1, that he can reach with a transition with label c_1, or with a transition with label e.
  - t^1_2 has a unique child t^2_2, that he can reach with a transition with label e.

- If v appears negatively in C_1 then
  - t_1^1 has a unique child t_1^2, that he can reach with a transition with label e.
  - t_2^1 has a unique child t_2^2, that he can reach with a transition with label c_1, and with a transition with label e.
- If v does not appear in C_1 then
  - t_1^1 has a unique child t_1^2, that he can reach with a transition with label e.
  - t_2^1 has a unique child t_2^2, that he can reach with a transition with label e.
- The same holds from states t_1^2 and t_2^2, changing the clause C_1 by C_2, and continues for all clauses.
- Finally, from states t_1^{l+1} and t_2^{l+1} there exists transitions with label e_1 and e_2 respectively, to a terminal state v^{end}. Hence, any run on the top part of the unit contains exactly the same number of letters e_1 and e_2 as a run on the low part.

The height of this DAG is $l+3$.

For instance, suppose $\phi = (p_1 \vee \neg q_1 \vee p_2) \wedge (\neg p_1 \vee q_2 \vee p_3)$. The DAG associated to variable p_1 is:



Our MDP is a sequence of such DAGs. The root of our MDP, which will be its initial state, is the variable p_1. The terminal state p_1^{end} of the tree U_{p_1} is identified to the first state q_1 of the tree U_{q_1}. Next, the last state q_1^{end} of U_{q_1} is identified with the initial state p_2 of U_{p_2}, and so on, until state q_m^{end}. The initial distribution on $\mathcal{S}$ is $\alpha = \delta_{p_1}$, the Dirac distribution on state p_1. Notice that on $\mathcal{S}$, only probabilistic transitions are on the states $q_1, ..., q_m$. A path from the initial state of $\mathcal{S}$ to the state q_m^{end} has length $2 \cdot m \cdot (l+3)$.

Intuitively, an "accepting run" on the MDP will be a run which contains in its trace exactly one time each letter $c_i, i \in [1;l]$. The bridges from a state t_1^j where both transition labels e and c_i are authorized will allow not to take label c_i when we have already seen it in the run.

Given $w \in \Sigma^*$, we consider its 1-gram as a multiset of letters in Σ, such that the union of the letters in this multiset are exactly the letters of w. Let $w \in \Sigma^{2 \cdot m \cdot (l+3)}$ be such that it contains exactly one time each letters $c_i, i \in [1;l]$, m times the letter e_1, m times the letter e_2, $2 \cdot m$ times the letter e, and $2 \cdot m \cdot (l+3) - l - 4 \cdot m$ times the letter e. We prove the following claim:

Claim:

$$\left\{ \begin{array}{l} \exists \sigma \in HD(\mathcal{S}) | \mathbb{P}^\sigma(B^1_{2 \cdot m \cdot (l+3)}(w)) \geq 1/2 \\ \text{iff } \exists p_1 R q_1 ... \exists p_m R q_m \ \mathbb{P}(\phi = true) \geq 1/2. \end{array} \right.$$

Suppose first that $\exists p_1 R q_1 ... \exists p_m R q_m \ \mathbb{P}(\phi = true) \geq 1/2$. We define our policy $\sigma \in HD(\mathcal{S})$ as follows. If a finite history h ends on a state of the type p_i, there exists the possibility between a transition with label e_1, and a transition with label e_2. If $b \in \{0,1\}$ is the choice of p_i once $p_1, q_1, ..., p_{i-1}, q_{i-1}$ have been chosen as in the history h on $\exists p_1 R q_1 ... \exists p_m R q_m \ \mathbb{P}(\phi = true) \geq 1/2$, then σ chooses the transition with label e_b. If a finite history h ends on a state of the type t_j^i, then the choice of σ depends on the occurrence of c_j in h: if c_j is already in h, then σ chooses label e, whereas if c_j is not in h, σ chooses label c_j. The goal is that on each run of length $2 \cdot m \cdot (l+3)$ on the MDP, the $c_i, i \in [1; l]$ appear exactly once.

Then, if we resolve the nondeterminism on $\mathcal{S}$ by σ, the probability to get run of length $2 \cdot m \cdot (l+3)$ whose trace contains exactly the letters of w, is exactly the probability to satisfy ϕ.

Conversely, suppose $\exists \sigma \in HD(\mathcal{S})$ such that $\mathbb{P}^\sigma(B^1_{2\cdot m\cdot(l+3)}(w)) \geq 1/2$. The choices of the p_i for $\exists p_1 R q_1 ... \exists p_m R q_m \; \mathbb{P}(\phi = true) \geq 1/2$ are the same as the choices that σ makes on states of type p_i. As before, the two probabilities to satisfy ϕ and to get a trace δ-close to $ustat_1(w)$ coincide, hence the claim.

Second reduction: The problem SSSAT.

This second part of the proof introduces a generalization of the SSAT problem, that we call *Super SSAT* (SSSAT). We prove that SSSAT is computationally PTIME equivalent to SSAT, which can be of independent interest. The reduction between SSSAT and SSAT does not imply any MDPs. We next prove that we can reduce SSSAT to problem 1, i.e. when we consider *randomized* policies on our MDP. In SSSAT, we allow random choices on the p_i. Instead of choosing the p_i in $\{0, 1\}$, we give $\lambda_i \in [0; 1]$ and generate p_i randomly with probabilities given by λ_i: p_i is one with probability λ_i, and zero with probability $1 - \lambda_i$. The SSSAT problem can be presented as: given variables p_i, q_j and formula ϕ as for SSAT, does there exists $\lambda_1 \in [0; 1]$ such that if p_1 is chosen randomly according to threshold λ_1 and if q_1 is chosen randomly, there exists λ_2 such that if p_2 is chosen randomly according to threshold λ_2,..., such that the probability to satisfy ϕ is at least $1/2$. We use the notation:

$$\exists \lambda_1 R q_1 \exists \lambda_2 R q_2 ... \exists \lambda_m R q_m \; \mathbb{P}(\phi = true) \geq 1/2.$$

We prove that SSSAT and SSAT are equivalent, with the identity for reduction: On an input $p_1, ..., p_m, q_1, ..., q_m, \phi$, we have:

$$\begin{gathered} \exists p_1 R q_1 ... \exists p_m R q_m \; \mathbb{P}(\phi = true) \geq 1/2 \\ \text{iff } \exists \lambda_1 R q_1 \exists \lambda_2 R q_2 ... \exists \lambda_m R q_m \; \mathbb{P}(\phi = true) \geq 1/2. \end{gathered}$$

The way $\Rightarrow$ is trivial, by taking the λ_i in $\{0, 1\}$.

For $\Leftarrow$, suppose that

$$\exists \lambda_1 R q_1 \exists \lambda_2 R q_2 ... \exists \lambda_m R q_m \; \mathbb{P}(\phi = true) \geq 1/2.$$

We define a function f which takes values in $\{0, 1\}$ and which associates to all sequences $p_1, q_1, ..., p_i, q_i, \; i \in [1; m-1]$, a choice for the variable p_{i+1}, in order to satisfy $\exists p_1 R q_1 ... \exists p_m R q_m \; \mathbb{P}(\phi = true) \geq 1/2$ with probability at least $1/2$. Given a valuation for $p_1, q_1, ..., p_i, q_i$, for $i \in [1; m]$, we write $\mathbb{P}(\phi = 1 | (p_1, q_1, ..., p_i, q_i))$ for the probability that ϕ is satisfied, given this valuation for the first variables.

We prove by induction on $i \in [1; m]$ that we can define f such that for all $i \in [1; m]$ we have:

$$\mathbb{P}(\phi = true | p_1 = f(.), p_2 = f(p_1, q_1), ..., p_i = f(p_1, q_1, ..., p_{i-1}, q_{i-1})) \geq 1/2.$$

- Suppose first $i = 1$. Then $\mathbb{P}(\phi = true) = \mathbb{P}(\phi = true | p_1 = 1) \cdot \mathbb{P}(p_1 = 1) + \mathbb{P}(\phi = true | p_1 = 0) \cdot \mathbb{P}(p_1 = 0)$. That is, $\mathbb{P}(\phi = true) = \mathbb{P}(\phi = true | p_1 = 1) \cdot \lambda_1 + \mathbb{P}(\phi = true | p_1 = 0) \cdot (1 - \lambda_1)$. Since $\mathbb{P}(\phi = true) \geq 1/2$, this proves that either $\mathbb{P}(\phi = true | p_1 = 1) \geq 1/2$, in which case we define $f(.) = 1$, or $\mathbb{P}(\phi = true | p_1 = 0) \geq 1/2$, in which case we define $f(.) = 0$.

- Let $i \in [1; m]$. Suppose the claim true for i. We use the simple result from probability theory that, given three events A, B, C, we have:

  $$\mathbb{P}(A|B) = \mathbb{P}(A | (B \wedge C)) \cdot \mathbb{P}(C|B) + \mathbb{P}(A | (B \wedge \overline{C})) \cdot \mathbb{P}(\overline{C}|B))$$

  Where $\overline{C}$ is the negation of C.

  We apply the previous result with the following associations:

  – A represents the formula $\phi = 1$,
  – B is $(p_1 = f(.), p_2 = f(p_1, q_1), ..., p_i = f(p_1, q_1, ..., p_{i-1}, q_{i-1}))$,
  – C is $p_{i+1} = 1$,

– and $\overline{C}$ is $p_{i+1} = 0$

Write λ_{i+1} for $\lambda_{i+1}(p_1, q_1, ..., p_i, q_i)$. Notice that $\mathbb{P}(C|B) = \lambda_{i+1}$, and $\mathbb{P}(\overline{C}|B) = 1 - \lambda_{i+1}$. Then,
$\mathbb{P}(\phi = true|(p_1 = f(.), p_2 = f(p_1, q_1), ..., p_i = f(p_1, q_1, ..., p_{i-1}, q_{i-1})) =$
$\mathbb{P}(A|B) =$
$\mathbb{P}(A|(B \wedge C)) \cdot \mathbb{P}(C|B) + \mathbb{P}(A|(B \wedge \overline{C})) \cdot \mathbb{P}(\overline{C}|B)) =$
$\mathbb{P}(A|(B \wedge C)) \cdot \lambda_{i+1} + \mathbb{P}(A|(B \wedge \overline{C})) \cdot (1 - \lambda_{i+1})$.
Now, since by induction hypothesis $\mathbb{P}(A|B) \geq 1/2$, either $\mathbb{P}(A|(B \wedge C)) \geq 1/2$, in which case we define $f(p_1, q_1, ..., p_i, q_i) = 1$, or $\mathbb{P}(A|(B \wedge \overline{C})) \geq 1/2$, in which case we define $f(p_1, q_1, ..., p_i, q_i) = 0$.

Finally, for $i = m$, we get a choice procedure for the p_i, given by f, such that ϕ is satisfied with probability at least $1/2$. That is, $\exists p_1 R q_1 ... \exists p_m R q_m \; \mathbb{P}(\phi = true) \geq 1/2$.

Once we have proved this result, it is easy to see that our previous reduction works for randomized policies as before, and that with the same construction for the MDP we have:

$$\exists \sigma \in HR(\mathcal{S}) | \mathbb{P}^\sigma(B^1_{2 \cdot m \cdot (l+3)}(w)) \geq 1/2$$
$$\text{iff } \exists \lambda_1 R q_1 ... \exists \lambda_m R q_m \; \mathbb{P}(\phi = true) \geq 1/2.$$

This proves that the Statistical membership problem is PSPACE hard. □

4.2. The approximate membership problem

Given $w \in \Sigma^n$ and $\lambda, \delta \geq 0$, let

$$B_n(w, \delta) = \{r \in \Omega | dist(w, Tr(r_{|n})) \leq \delta\}$$

The (λ, δ)-*membership problem* is: given an MDP $\mathcal{S}$ and a word w, is there a policy $\sigma \in HR(\mathcal{S})$ such that $\mathbb{P}^\sigma(B_n(w, \delta)) \geq \lambda$?

We define the associated language as:

$$\mathcal{L}^\lambda_\delta = \{w \in \Sigma^*, \; |w| = n \in \mathbb{N} | \exists \sigma \in HR(\mathcal{S}) \; s.t. \; \mathbb{P}^\sigma(B_n(w, \delta)) \geq \lambda\}$$

Using proposition 1 which links the distance $dist$ between words and the statistics, deciding if a word belongs to $\mathcal{L}^\lambda_\delta$ can be done approximately by deciding if a word belongs to the language $\mathcal{L}^{\lambda,k}_\delta$, where $k = \ulcorner \frac{1}{\epsilon} \urcorner$ and

$$\mathcal{L}^{\lambda,k}_\delta = \{w \in \Sigma^*, \; |w| = n \in \mathbb{N} | \exists \sigma \in HR(\mathcal{S}) \; s.t. \; \mathbb{P}^\sigma(B^k_n(w, \delta)) \geq \lambda\}$$

That is, $\mathcal{L}^{\lambda,k}_\delta$ is the set of finite words w for which there exists a policy on $\mathcal{S}$ such that the k-order statistics of the generated words of length $|w|$ are δ-close to w with probability at least λ.

We know that given $\mathcal{S}$, λ, δ, it is intractable to decide if a word w belongs to $\mathcal{L}^{\lambda,1}_\delta$. For the following we fix $k = \ulcorner \frac{1}{\epsilon} \urcorner$.

Definition 10. *If $\epsilon \in [0; 1]$, we say that a word $w \in \Sigma^n$ is ϵ-close to $\mathcal{L}^{\lambda,k}_\delta$ if there exists $w' \in \Sigma^n$ such that*

$$d(w, w') \leq \epsilon, \; \textit{and} \; w' \in \mathcal{L}^{\lambda \cdot (1-\epsilon), k}_{\delta \cdot (1+\epsilon)}$$

We prove that $\mathcal{L}^{\lambda,k}_\delta$ is constant time testable (see [24, 13] for a definition). That is, there exists an algorithm $\mathcal{T}_\epsilon$, which depends on a parameter $\epsilon \in [0; 1]$, such that for all $\epsilon \in]0; 1]$, for long enough input $w \in \Sigma^*$ we have:

- If $w \in \mathcal{L}^{\lambda,k}_\delta$, then $\mathcal{T}_\epsilon$ answers YES with probability at least $2/3$.
- If w is ϵ-far from $\mathcal{L}^{\lambda,k}_\delta$, $\mathcal{T}_\epsilon$ answers NO with probability at least $2/3$.
- The time complexity of $\mathcal{T}_\epsilon$ does not depend on the length of w.

Fix $\epsilon > 0$. The construction of $\mathcal{T}_\epsilon$ is simple: Let $S = S_0 \cup S_1 \cup ... \cup S_l$ be the decomposition of $\mathcal{S}$ into MECs. For all $i \in [1; l]$, compute the set of linear equations which give $\Pi_k^k(\mathcal{S}_i)$. As before, if $x \in \mathbb{R}^{\Sigma^k}$, $V(x, \delta) = \bigcup\{S_i | d_k(x, \mathcal{S}_i) \leq \delta\}$.

Algorithm 1 (The tester $\mathcal{T}_\epsilon(w)$)**.** *Fix an MDP $\mathcal{S}$ with initial distribution α.*
Input: $w \in \Sigma^*$.

- *Sample w, to obtain $\hat{x}$ an approximation of $ustat_k(w)$, in constant time.*
- *Compute $p = MaxReach(\alpha, V(\hat{x}, \delta \cdot (1 + \epsilon)))$.*
- *If $p \geq \lambda \cdot (1 - \epsilon/2)$, then $\mathcal{T}_\epsilon$ answers YES, if not, $\mathcal{T}_\epsilon$ answers NO.*

This algorithm decides approximately the (λ, δ)-membership problem in time independent of $|w|$. The following lemma, proved in [12], and based on a Chernoff bound, proves that we can sample $\hat{x}$ efficiently.

Lemma 2. *There exists a probabilistic algorithm which works in constant time on inputs $w \in \Sigma^*$ and which produces a vector $\hat{x} \in \mathbb{R}^{\Sigma^k}$ such that $\mathbb{P}(\|ustat_k(w) - \hat{x}\| < \delta \cdot \epsilon/14) \geq 2/3$.*

To prove the correctness of the tester, we use the results in [27] and [18]. Theorem 2 of [27], and theorem 5.1. of [18], generalized to general MDPs, give the following proposition:

Proposition 7. *Let $S = S_0 \cup S_1 \cup ... \cup S_l$ be the decomposition of a general MDP $\mathcal{S}$ into MECs. Then for all $i \in [1; l]$ we can find two constants C_1^i, C_2^i, such that for all $x \in \mathbb{R}^{\Sigma^k}$, $n \in \mathbb{N}, \alpha_i \in \Delta(S_i)$, and $\epsilon > 0$ we have:*

- *If $d_k(x, \mathcal{S}_i) \leq \delta \cdot (1 + \epsilon/2)$, then $\exists \sigma_i \in HR(\mathcal{S}_i)$ s.t. $\mathbb{P}^{\sigma_i, \alpha_i}(B_n(x, \delta \cdot (1 + \epsilon))) \geq 1 - C_1^i \cdot e^{-C_2^i \cdot n \cdot \epsilon^2}$*
- *If $d_k(x, \mathcal{S}_i) > \delta \cdot (1 + \epsilon)$, then $\forall \sigma_i \in HR(\mathcal{S}_i)$, $\mathbb{P}^{\sigma_i, \alpha_i}(B_n(x, \delta \cdot (1 + \epsilon/2))) \leq C_1^i \cdot e^{-C_2^i \cdot n \cdot \epsilon^2}$*

As before, given $n \in \mathbb{N}$, X_n is the random variable which associates to a run its n-th state. By definition of S_0, we know that we can find two constants C_1^0, C_2^0 such that for any policy σ on $\mathcal{S}$, all $\alpha \in \Delta(S)$ and all $n \in \mathbb{N}$, $\mathbb{P}^{\sigma, \alpha}(X_n \in S_0) \leq C_1^0 \cdot e^{-C_1^0 \cdot n}$. In the following, $q, N \in \mathbb{N}$ are such that

$$C_1^0 \cdot e^{-C_1^0 \cdot q} < \frac{\lambda \cdot \epsilon}{4}; \ \sum_{i=1}^{l} C_1^i \cdot e^{-C_2^i \cdot (N-q) \epsilon^2 1} < \frac{\lambda \cdot \epsilon}{4}$$

Lemma 3. *Let $x \in \mathbb{R}^{\Sigma^k}, n \geq 2 \cdot N$.*

1. *If $MaxReach(\alpha, V(x, \delta \cdot (1 + \epsilon))) \leq \lambda \cdot (1 - \epsilon/2)$, then $\forall \sigma \in HR(\mathcal{S})$, $\mathbb{P}^{\sigma, \alpha}(B_n(x, \delta \cdot (1 + \epsilon/2))) \leq \lambda$.*
2. *If $MaxReach(\alpha, V(x, \delta \cdot (1 + \epsilon))) \geq \lambda \cdot (1 - \epsilon/2)$, then $\exists \sigma \in HR(\mathcal{S})$ s.t. $\mathbb{P}^{\sigma, \alpha}(B_n(x, \delta \cdot (1 + \epsilon))) \geq \lambda \cdot (1 - \epsilon)$.*

Proof. Suppose first $MaxReach(\alpha, V(x, \delta \cdot (1 + \epsilon))) \leq \lambda \cdot (1 - \epsilon/2)$, and let $\sigma \in HR(\mathcal{S})$. We use the following decomposition:

$$B_n(x, \delta \cdot (1 + \epsilon/2)) \subseteq \{X_n \in S_0\} \cup \{X_n \in V(x, \delta \cdot (1 + \epsilon))\} \cup \{X_n \in (S \backslash V(x, \delta \cdot (1 + \epsilon)))\} \cap B_n(x, \delta \cdot (1 + \epsilon/2))$$

We prove that the terms in the right hand side have small probabilities. By the choice of N and n, and by proposition 7,

$$\mathbb{P}^{\sigma, \alpha}(\{X_n \in S_0\}) \leq \lambda \cdot \epsilon/4$$

and $\mathbb{P}^{\sigma, \alpha}(\{X_n \in (S \backslash V(x, \delta \cdot (1 + \epsilon)))\} \cap B_n(x, \delta \cdot (1 + \epsilon/2))) \leq \lambda \cdot \epsilon/4$. Moreover, by hypothesis, $\mathbb{P}^{\sigma, \alpha}(\{X_n \in V(x, \delta \cdot (1 + \epsilon))\}) \leq \lambda \cdot (1 - \epsilon/2)$. Finally,

$$\mathbb{P}^{\sigma, \alpha}(B_n(x, \delta \cdot (1 + \epsilon/2))) \leq 2 \cdot \lambda \cdot \epsilon/4 + \lambda \cdot (1 - \epsilon/2) \leq \lambda$$

Which proves the first point.

For the second point, using proposition 2, there exists σ_0 a deterministic policy on $\mathcal{S}$ such that

$$\mathbb{P}^{\sigma_0,\alpha}(Reach(V(x,\lambda\cdot(1+\epsilon)))\geq\lambda\cdot(1-\epsilon/2)$$

We define σ as follows: until a state in $\bigcup_{i\in[1;l]}S_i$ has been reached, σ follows σ_0. Suppose one of the S_i has been entered, for $i\in[1;l]$. If $d_k(x,\mathcal{S}_i)>\delta\cdot(1+\epsilon/2)$, define σ to be anything. Suppose $d_k(x,\mathcal{S}_i)\leq\delta\cdot(1+\epsilon/2)$. Let $x_i\in\Pi_k^k(\mathcal{S}_i)$ such that $\|x-x_i\|\leq\delta\cdot(1+\epsilon/2)$. Then, by proposition 7, there exists a policy σ_i on $\mathcal{S}_i$ such that for all $m\geq 0$ and all initial distribution α_i on S_i we have $\mathbb{P}^{\sigma_i,\alpha_i}(B_m(x,\delta\cdot(1+\epsilon)))\geq 1-C_1^i\cdot e^{-C_2^i\cdot m\cdot\epsilon^2}$. Now, for all $m\in\mathbb{N}$, we have:

$$\mathbb{P}^{\sigma,\alpha}(B_n(x,\delta\cdot(1+\epsilon))\geq\mathbb{P}^{\sigma,\alpha}(B_n(x,\delta\cdot(1+\epsilon))|\{X_m\in S_i\})\cdot\mathbb{P}^{\sigma,\alpha}(\{X_m\in V(x,\delta\cdot(1+\epsilon/2))\})$$

If $m\geq N$, since σ behaves as σ_0 until a state in $\bigcup\{S_i|i\in[1;l]\}$ is reached, and since

$$\mathbb{P}^{\sigma_0,\alpha}(Reach(V(x,\delta\cdot(1+\epsilon)))\geq\lambda\cdot(1-\epsilon/2)$$

we get that

$$\mathbb{P}^{\sigma,\alpha}(\{X_m\in V(x,\delta\cdot(1+\epsilon/2))\})\geq\lambda\cdot(1-\epsilon/2)-\lambda\epsilon/4=\lambda\cdot(1-3\epsilon/4)$$

Now, if $n\geq m+N$, we have $\mathbb{P}^{\sigma,\alpha}(B_n(x,\delta\cdot(1+\epsilon))|\{X_m\in S_i\})\geq 1-\epsilon\cdot\lambda/4$. Taking $m=N$ and $n=2\cdot N$ we get the result, since $(1-\epsilon\cdot\lambda/4)\cdot\lambda\cdot(1-3\epsilon/4)\geq\lambda\cdot(1-\epsilon)$ if $\lambda,\epsilon\in[0;1]$.

□

Theorem 2 (Correctness of $\mathcal{T}_\epsilon$)**.** *For all $\epsilon\in]0;1[$, $\mathcal{T}_\epsilon$ is an ϵ-tester for $\mathcal{L}_\delta^\lambda$.*

Proof.

- Suppose $w\in\mathcal{L}_\delta^\lambda$, $|w|=n$. Let $x=ustat_k(w)$. Then, $\exists\sigma\in HR(\mathcal{S})$ such that $\mathbb{P}^{\sigma,\alpha}(B_n(w,\delta))\geq\lambda$. By lemma 2, with probability at least $2/3$, $\|x-\hat{x}\|\leq\delta\cdot\epsilon/2$, so with probability at least $2/3$, $\exists\sigma\in HR(\mathcal{S})$ such that $\mathbb{P}^{\sigma,\alpha}(B_n(\hat{x},\delta\cdot(1+\epsilon/2)))>\lambda$. Using point 1 of lemma 3, with probability at least $2/3$,

$$MaxReach(\alpha,V(\hat{x},\delta\cdot(1+\epsilon)))\geq\lambda\cdot(1-\epsilon/2)$$

  I.e., $\mathcal{T}_\epsilon$ answers YES with probability at least $2/3$

- Conversely, suppose $d(w,\mathcal{L}_{\delta\cdot(1+\epsilon)}^{\lambda\cdot(1-\epsilon)})>\epsilon$. Let $w'\in\Sigma^n$, and $y=ustat_k(w')$. By contraposition, suppose $MaxReach(\alpha,V(y,\delta\cdot(1+\epsilon)))\geq\lambda\cdot(1-\epsilon/2)$. Then, by the point 2 of lemma 3, $\exists\sigma\in HR(\mathcal{S})$ such that $\mathbb{P}^{\sigma,\alpha}(B_n(y,\delta\cdot(1+\epsilon)))\geq\lambda\cdot(1-\epsilon)$. This implies $w'\in\mathcal{L}_{\delta\cdot(1+\epsilon)}^{\lambda\cdot(1-\epsilon)}$, and by hypothesis, $\mathrm{dist}(w,w')\geq\epsilon$. Using proposition 1, this implies $\|x-y\|\geq\epsilon/7$. Since with probability at least $2/3$, $\|x-\hat{x}\|\leq\epsilon/14$, with probability at least $2/3$ we have $MaxReach(\alpha,V(\hat{x},\delta\cdot(1+\epsilon)))<\lambda\cdot(1-\epsilon/2)$. Hence, $\mathcal{T}_\epsilon$ answers NO with probability at least $2/3$.

□

Using proposition 5 and the results of [7], $\mathcal{T}_\epsilon$ works in time polynomial in $|S|^2\cdot(|\Sigma|\cdot|S|)^k$, which is independent of $|w|$.

4.3. Comparison problems

Two MDPs should be close if they can generate close words with close probabilities. An MDP $\mathcal{S}_1$ should be simulated by an MDP $\mathcal{S}_2$ if $\mathcal{S}_2$ can generate words close to the words generated by $\mathcal{S}_1$, with higher probability. Given $w\in\Sigma^n$ and $\epsilon\geq 0$, let

$$\lambda_1(w,\epsilon)=Sup_{\sigma_1\in HR(\mathcal{S}_1)}\mathbb{P}^{\sigma_1}(B_n(w,\epsilon)),$$
$$\lambda_2(w,\epsilon)=Sup_{\sigma_2\in HR(\mathcal{S}_2)}\mathbb{P}^{\sigma_2}(B_n(w,\epsilon)).$$

The following problem asks wether an MDP can generate words close to w with higher probability than a second MDP, for all w large enough.

Problem 3 (MDP ϵ-simulation).
Input: *$\mathcal{S}_1, \mathcal{S}_2$ with initial distributions α_1 and α_2.* ***Question:*** *$\forall w \in \Sigma^*$ large enough, $\lambda_1(w) \leq \lambda_2(w)$?*

The distance $d_k^{\prec}$ can be seen as a distance of simulation. We prove that if λ_1 and λ_2 are close, then $d_k^{\prec}(\mathcal{S}_1, \mathcal{S}_2)$ is small, and reciprocally.

Given two MDPs $\mathcal{S}_1$ and $\mathcal{S}_2$ with respective initial distributions α_1 and α_2, let N_1, N_2 be as for lemma 3, for $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. If $i \in \{1, 2\}$, $w \in \Sigma^n$ and $\epsilon > 0$, let $\lambda_i(w, \epsilon) = Sup_{\sigma_i \in HR(\mathcal{S}_i)} \mathbb{P}_i^{\sigma_i, \alpha_i}(B_n(w, \epsilon))$. The following proposition is the analogous of the tester 1 for comparing MDPs. It gives an approximate additive solution for problem 3.

Proposition 8. *Let $n \geq Max(N_1, N_2)$.*

- *If $d_k^{\prec}(\mathcal{S}_1, \mathcal{S}_2) \leq \epsilon/2$, then for all $w \in \Sigma^*$, $\lambda_1(w, 0) \leq \lambda_2(w, \epsilon) + \epsilon$.*
- *Conversely, suppose $d_k^{\prec}(\mathcal{S}_1, \mathcal{S}_2) > 3 \cdot \epsilon/2$. Then we can find $w \in \Sigma^*$ such that $\lambda_1(w, \epsilon/2) > \lambda_2(w, \epsilon) + \epsilon/4$*

Proof. We write α_1, α_2 for the respective initial distributions on $\mathcal{S}_1$ and $\mathcal{S}_2$.

Suppose first $d_k^{\prec}(\mathcal{S}_1, \mathcal{S}_2) \leq \epsilon/2$. recall that if $x \in \mathbb{R}^{\Sigma^k}$ and $\delta \geq 0$, $V(x, \delta) = \bigcup \{S_i | d_k(x, \mathcal{S}_i) \leq \delta\}$. Let $x \in \mathbb{R}^{\Sigma^k}$. Let $\lambda = MaxReach_1(\alpha_1, V_1(x, 0))$. Using the definition of $d_k^{\prec}$ we know that $\lambda \leq MaxReach_2(\alpha_2, V_2(x, \epsilon/2)) + \epsilon/2$. Since $n \geq Max(N_1, N_2)$, there exists a policy σ_2 on $\mathcal{S}_2$ such that $\mathbb{P}_2^{\sigma_2, \alpha_2}(B_n(x, \epsilon)) \geq \lambda$, and for all policy σ_1 on $\mathcal{S}_1$, $\mathbb{P}_1^{\sigma_1, \alpha_1}(B_n(x, 0)) \leq \lambda + \epsilon$. Finally, for all $\sigma_1 \in HR(\mathcal{S}_1)$ and $\sigma_2 \in HR(\mathcal{S}_2)$, we have $\mathbb{P}_1^{\sigma_1, \alpha_1}(B_n(x, 0)) \leq \mathbb{P}_2^{\sigma_2}(B_n(x, \epsilon)) + \epsilon$.

For the second point. Suppose $d_k^{\prec}(\mathcal{S}_1, \mathcal{S}_2) > 3 \cdot \epsilon/2$. Let $\gamma_1(x) = MaxReach_1(\alpha_1, V_1(x, 0))$ and $\gamma_2(x) = MaxReach_2(\alpha_2, V_2(x, 3 \cdot \epsilon/2))$. Then, there exists $x_0 \in \mathbb{R}^{\Sigma^k}$ such that $\gamma_1(x_0) > \gamma_2(x_0) + 3 \cdot \epsilon/2$. Since $n \geq N$, there exists a policy σ_1 on $\mathcal{S}_1$ such that $\mathbb{P}_1^{\sigma_1, \alpha_1}(B_n(x_0, \epsilon/2)) > \gamma_1(x_0) \cdot (1 - \epsilon)$, whereas for all σ_2 on $\mathcal{S}_2$, $\mathbb{P}_2^{\sigma_2, \alpha_2}(B_n(x_0, \epsilon)) \leq \gamma_2(x_0) \cdot (1 + \epsilon)$. Since $\gamma_1(x_0)$ and $\gamma_2(x_0)$ are in $[0; 1]$ and $\gamma_1(x_0) > \gamma_2(x_0) + 3 \cdot \epsilon/2$, we have $\gamma_1(x_0) \cdot (1 - \epsilon) \geq \gamma_2(x_0) \cdot (1 + \epsilon) + \epsilon/4$ if $\epsilon \leq 1/2$, which proves the result, by taking $w \in \Sigma^*$ such that $ustat_k(w) = x_0$.

□

Corollary 2. *We can decide approximately problem 3 in PTIME.*

5. Problems related to Probabilistic Automata

In this section we study some problems related to Probabilistic Automata. They are the analogous of the problems studied in the previous section concerning MDPs. We will see that these are much more difficult to solve, even approximately. We will deal with three problems concerning PAs: the language emptiness problem, the membership problem of a given word to the language of a PA, and the comparison problem between two PAs.

5.1. The n-emptiness Problem for PAs

If $\mathcal{A}$ is a probabilistic automaton and $w \in \Sigma^*$, $\mathbb{P}_{\mathcal{A}}(w)$ is the probability to arrive to an accepting state when w is read on $\mathcal{A}$. Recall first the undecidability of the *emptiness promise problem for a PA*: Given a PA, given $1/2 > \epsilon > 0$, even if we know that we are in one of the two cases, it is undecidable to decide if there exists $w \in \Sigma^*$ such that $\mathbb{P}(w) > 1 - \epsilon$, or for all $w \in \Sigma^*$, $\mathbb{P}(w) < \epsilon$ (Corollary 3.4 of [17]). If the length of the word is fixed, the associated promise problem is NP-complete:

Problem 4 (n-Emptyness).
Input: *$\mathcal{A}$, $\epsilon \in]0; 1/2[$, $n \in \mathbb{N}$.*
Question: *Decide if there is a word $w \in \Sigma^n$ such that $\mathbb{P}_{\mathcal{A}}(w) \geq 1 - \epsilon$, or if for all word $w \in \Sigma^n$, $\mathbb{P}_{\mathcal{A}}(w) \leq \epsilon$.*

Proposition 9. *For any $\epsilon \in]0; 1/2[$, the n-Emptiness Problem is NP complete.*

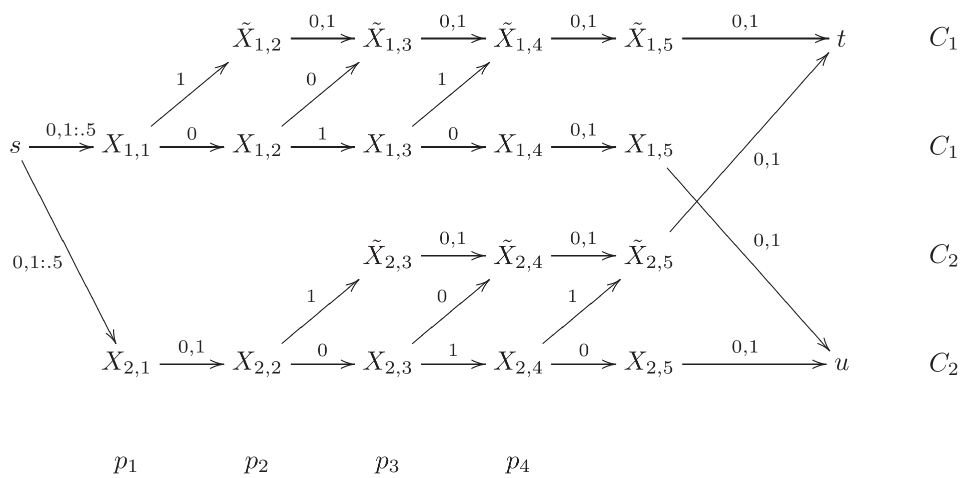
Proof. The fact that n-emptiness is is NP is clear: given a word and an automaton we can compute in PTIME the probability that the automaton accepts the word. We give a PTIME reduction from 3-SAT to this problem when $\epsilon = 1/4$. We can prove the result for any $\epsilon \in]0; 1/2[$ in an analogous way.
Let I be an instance of 3-SAT. I is a conjunction $C_1 \wedge C_2 \wedge ... \wedge C_l$ of 3-clauses, on the boolean variables $p_1, ..., p_m$. We build a probabilistic automaton $\mathcal{A}$ from this instance. We will prove the following:

There exists a word $w \in \Sigma^m$ such that $\mathbb{P}_{\mathcal{A}}(0w0) > 1 - 1/(2 \cdot l)$ iff I is satisfiable.

The structure of $\mathcal{A}$ is layered:

- The alphabet of the automaton is $\Sigma = \{0, 1\}$. $\mathcal{A}$ has $3 + (2 * l) * (m + 1)$ states: s, t, u, $X_{i,j}$ for $i \in [1, l], j \in [1, m + 1]$, and $\tilde{X}_{i,j}$ for $i \in [1, l], j \in [1, m]$. State s is the initial state, and t is the only terminal state.
- There are two vertical layers for each 3-clauses, and as many horizontal steps as there are many variables.
- From state $X_{i,j}, i \in [1, l], j \in [1, m]$, the transitions depend on the quantification of variable p_j in the clause C_i:
  - If p_i appears positively in C_i, then $\mathcal{A}$ in state $X_{i,j}$ can read 1 and move deterministally to state $\tilde{X}_{i,j+1}$, and read 0 and move deterministally to state $X_{i,j+1}$.
  - If p_i appears negatively in C_i, then $\mathcal{A}$ in state $X_{i,j}$ can read 0 and move deterministally to state $\tilde{X}_{i,j+1}$, and read 1 and move deterministally to state $X_{i,j+1}$.
  - If p_i does not appear in C_i, then $\mathcal{A}$ in state $X_{i,j}$ can read 0 or 1 and move deterministally to state $X_{i,j+1}$.
- From state s, $\mathcal{A}$ can read 0 or 1 and move probabilistically to one of the states $X_{1,1}, X_{2,1}, ..., X_{l,1}$ with uniform probability $\frac{1}{l}$ between these states. This is the only probabilistic transition.
- From state $\tilde{X}_{i,j}, i \in [1, l], j \in [1, m]$, $\mathcal{A}$ can read 0 or 1 and move deterministically to state $\tilde{X}_{i,j+1}$.
- From state $\tilde{X}_{i,m+1}, i \in [1, l]$, $\mathcal{A}$ can read 0 or 1 and move deterministically to state t.
- From state $X_{i,m+1}, i \in [1, l]$, $\mathcal{A}$ can read 0 or 1 and move deterministically to state u.

Intuitively, being on one of the states $\tilde{X}_{i,j}, i \in [1, l], j \in [1, m]$ means that the valuation satisfies the clause C_i. We give an example in the case where $l = 2$, $m = 4$ and $I = (p1 \vee \neg p_2 \vee p_3) \wedge (p_2 \vee \neg p_3 \vee p_4)$:

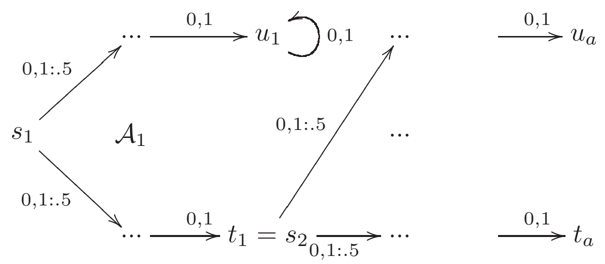


Then, there exists a word $w \in \Sigma^m$ such that $\mathbb{P}_{\mathcal{A}}(0w0) > 1 - 1/(2 \cdot l)$ iff I is satisfiable. The first remark is that if a valuation $(p_1 = a_1, p_2 = a_2, ..., p_m = a_m),\ a_i \in \{0,1\}$, satisfies a clause $C_i,\ i \in [1;l]$, then, after having read the word $0a_1a_2...a_m$, we are on state $\tilde{X}_{i,m}$ with probability at least $1/l$.

Now, suppose that there is a valuation $(p_1 = a_1, p_2 = a_2, ..., p_m = a_m),\ a_i \in \{0,1\}$ which satisfies I. Then, after having read the word $0a_1a_2...a_m$, we are with probability one in one of the states $\tilde{X}_{i,m}, i \in [1;l]$. Hence, with probability one, after having read $0a_1a_2...a_m0$, we are in state t, that is $0a_1a_2...a_m0$ is accepted with probability one.

Conversely, suppose that after having read a word $0a_1a_2...a_m$ we are in one of the states $X_{i,m},\ i \in [1;l]$, with positive probability. Then we are in this state with probability at least $1/l$, by the construction of $\mathcal{A}$. This proves that if after having read a word $0a_1a_2...a_m0$ we are on state u with positive probability, then we are in fact on state u with probability at least $1/l$, and $0a_1a_2...a_m0$ is not accepted with probability at least $1/l$. We get that if there exists a word $0a_1a_2...a_m0$ which is accepted with probability greater than $1 - 1/l$, the valuation $(p_1 = a_1, p_2 = a_2, ..., p_m = a_m),\ a_i \in \{0,1\}$ satisfies all the clauses $C_i,\ i \in [1;l]$, and I is satisfiable.

In order to get a larger gap, we consider a sequence of the preceding construction: The probabilistic automaton associated to I is now a sequence of a number a of the probabilistic automata we just built with $a \in \mathbb{N}$ such that $(1 - 1/l)^a \leq \epsilon$, for instance $a \geq -l \cdot ln(\epsilon)$. Hence the reduction is still polynomial. We show the construction for the last example. In the sequence, the accepting state t_k of the k-th copy of $\mathcal{A}$ plays the role of the initial state of the $k+1$-th copy, whereas u_i is a dead non accepting state:



Finally on the new automaton $\mathcal{A}'$ we have:

- There exists a word $w \in \Sigma^m$ such that $\mathbb{P}(0w0) \geq 1 - \epsilon$ iff I is satisfiable
- I is not satisfiable iff for all words $w \in \Sigma^m$, $\mathbb{P}(0w0) \leq \epsilon$

□

5.2. Membership problems for PAs

We consider now membership problems. An automaton $\mathcal{A}$, and a threshold $\lambda \in]0;1[$, are fixed.

Problem 5 (Membership problem)**.**
Input: $w \in \Sigma^*$.
Question: *Do we have* $\mathbb{P}_{\mathcal{A}}(w) \geq \lambda$*?*

The associated language is:

$$\mathcal{L}^\lambda = \{w \in \Sigma^* | \mathbb{P}_{\mathcal{A}}(w) \geq \lambda\}.$$

Problem 5 is clearly computable in time $O(|\mathcal{A}| \cdot |w|)$. In this section we prove that the language $\mathcal{L}^\lambda$ is not constant time testable, which contrasts with the results of [12] and our results in the context of MDPs. In [12], the authors consider the membership problem of a word w to the language of a non-deterministic automata. Using a geometric construction, they prove that this problem is testable in time independent of w.

Lemma 4. *Let $\mathcal{T}$ be a randomized $O(1)$-algorithm with inputs in Σ^* which works in time $N \in \mathbb{N}$ and samples subwords of length at most $k \in \mathbb{N}$. Let $w, w' \in \Sigma^*$ be such that $\|ustat_k(w) - ustat_k(w')\| \leq \frac{\epsilon}{2 \cdot N \cdot |\Sigma|^{k \cdot N}}$. Then*

$$|\mathbb{P}_{\mathcal{T}}(YES|w) - \mathbb{P}_{\mathcal{T}}(YES|w')| + |\mathbb{P}_{\mathcal{T}}(NO|w) - \mathbb{P}_{\mathcal{T}}(NO|w')| \leq \epsilon$$

Proof. Since $\mathcal{T}$ samples the subwords of the input with uniform probability, we can suppose that $\mathcal{T}$ samples once and for all, with uniform probability, a number N of subwords of the input, each one of length k. This gives a sequence $(u_1, ..., u_N) \in (\Sigma^k)^N$. Now, a word $w \in \Sigma^*$ induces the distribution $ustat_k(w)$ on Σ^k: if $\rho \in \Sigma^k$, $ustat_k(w)[\rho]$ is the probability to sample ρ on w in one step. A sequence $(v_1, ..., v_N) \in (\Sigma^k)^N$ is sampled in N steps with probability

$$\mathbb{P}_w((v_1, ..., v_N)) = \prod_{i=1}^{N} ustat_k(w)[v_i] \tag{3}$$

For $v = (v_1, ..., v_N) \in (\Sigma^k)^N$, let $\mathbb{P}_{\mathcal{T}}(YES|(v_1, ..., v_N))$ and $\mathbb{P}_{\mathcal{T}}(NO|(v_1, ..., v_N))$ be the probabilities that $\mathcal{T}$ answers YES or NO once v has been sampled. Then,

$$\mathbb{P}_{\mathcal{T}}(YES|w) = \textstyle\sum_{(v_1,...,v_N)\in(\Sigma^k)^N} \mathbb{P}_w((v_1, ..., v_N)) \cdot \mathbb{P}_{\mathcal{T}}(YES|(v_1, ..., v_N))$$
$$\mathbb{P}_{\mathcal{T}}(YES|w') = \textstyle\sum_{(v_1,...,v_N)\in(\Sigma^k)^N} \mathbb{P}_{w'}((v_1, ..., v_N)) \cdot \mathbb{P}_{\mathcal{T}}(YES|(v_1, ..., v_N))$$

But $\|ustat_k(w) - ustat_k(w')\| = \sum_{\rho\in\Sigma^k} |\beta_w(\rho) - \beta_{w'}(\rho)|$. Since $\|ustat_k(w) - ustat_k(w')\| \leq \frac{\epsilon}{2\cdot N\cdot|\Sigma|^{N\cdot k}}$ by hypothesis, and since the $\beta_w(\rho), \beta_{w'}(\rho)$ are in $[0; 1]$, we get:

$$\text{For all } (v_1, ..., v_N) \in (\Sigma^k)^N, \ \ |\prod_{i=1}^{N} \delta_w(v_i) - \prod_{i=1}^{N} \delta_{w'}(v_i)| \leq \frac{\epsilon}{2 \cdot |\Sigma|^{N\cdot k}}$$

Using equation 3, we get:

$$\text{For all } (v_1, ..., v_N) \in (\Sigma^k)^N, \ \ \mathbb{P}_w((v_1, ..., v_N)) - \mathbb{P}_{w'}((v_1, ..., v_N))| \leq \frac{\epsilon}{2 \cdot |\Sigma|^{N\cdot k}}$$

This proves $|\mathbb{P}_{\mathcal{T}}(YES|w) - \mathbb{P}_{\mathcal{T}}(YES|w')| \leq \epsilon/2$. The same holds for the answer NO. □

Thus, a constant time tester gives close results on inputs with close statistics, *independently of the size of the input*. It implies that if a property can be tested in constant time, it should be testable by considering inputs of bounded size. In order to try to test the language $\mathcal{L}^{\lambda}$, as we did in the context of MDPs, we may allow a relaxation on the threshold: if $\epsilon \geq 0$, a word $w \in \Sigma^n$ is ϵ-close to $\mathcal{L}^{\lambda}$ if there exists $w' \in \Sigma^n$ such that $d(w, w') \leq \epsilon$, and $w' \in \mathcal{L}^{\lambda\cdot(1-\epsilon)}$. As a consequence of the last lemma, we prove that there exists some PAs for which we cannot construct a constant time tester for the associated membership problem, even with the relaxation on the threshold.

Theorem 3. *The membership problem for PAs is not constant time testable.*

Proof. We have to show that we can find some PAs for which we cannot have a constant time tester. Consider the promise problem Q: given a PA $\mathcal{A}$ and $\delta > 0$, decide if there exists $w \in \Sigma^*$ such that $\mathbb{P}_{\mathcal{A}}(w) \geq 1 - \delta$, or if for all $w \in \Sigma^*$ we have $\mathbb{P}_{\mathcal{A}}(w) \leq \delta$.

Consider a PA $\mathcal{A}$ and $\delta \in]0; 1/4[$ for which emptiness problem Q is undecidable. Suppose now that there exists a constant time ϵ-tester $\mathcal{T}$ for the $1/2$-membership problem, with $\epsilon < 1/8$. Namely, on input $w \in \Sigma^*$, $\mathcal{T}$ satisfies:

- If $d(w, \mathcal{L}^{\frac{1}{2}}) = 0$, (i.e. if $w \in \mathcal{L}^{\frac{1}{2}}$), then $\mathcal{T}$ answers YES with probability at least $3/4$.
- If $d(w, \mathcal{L}^{\frac{1}{2}\cdot(1-\epsilon)}) \geq \epsilon$, (i.e. if w is far from $\mathcal{L}^{\frac{1}{2}}$), then $\mathcal{T}$ answers NO with probability at least $3/4$.

We reduce problem Q on $\mathcal{A}$ and δ to the existence of such a tester, and conclude that no such tester can exist. Let Δ be the set of k-statistic vectors in $\mathbb{R}^{\Sigma^k}$, Δ is a convex subset of $\mathbb{R}^k$. Let $\Gamma = x_1, x_2, ..., x_l$ be the vertices of a grid on Δ such that any points in Δ is at most $\epsilon/2$-far from a point in Γ. Clearly, we can take $l = |\Gamma| \leq (1/\epsilon)^k$. Now, let $n \in \mathbb{N}$ be large enough such that for all $x \in \Gamma$ there exists a word $w \in \Sigma^n$ with $\|ustat_k(w) - x\| \leq \epsilon/2$.
Our decision procedure works as follows: for all $x \in \Gamma$, chose $w_x \in \Sigma^n$ such that $\|ustat_k(w_x) - x\| \leq \epsilon/2$. For each one of these w_x, run $10 \cdot l$ times the tester $\mathcal{T}$ on this input. Let $r_x = 1$ if a majority of the i runs have given answer YES, and let $r_x = 0$ if not. Finally, let $r = 1$ if one of the r_x is one, and $r = 0$ if not. We prove that:

- If there exists $w \in \Sigma^*$ of length greater than n such that $\mathbb{P}_{\mathcal{A}}(w) \geq 1 - \delta$, then $r = 1$ with probability at least $3/4$.

- If for all $w \in \Sigma^*$ of length greater than n we have $\mathbb{P}_{\mathcal{A}}(w) \leq \delta$, then $r = 0$ with probability at least $3/4$.

Suppose first that there exists $w \in \Sigma^*$ with length greater than n such that $\mathbb{P}_{\mathcal{A}}(w) \geq 1 - \delta$. Let $x \in \Gamma$ be such that $\|x - ustat_k(w)\| \leq \epsilon/2$. Then $|ustat_k(w) - ustat_k(w_x)| \leq \epsilon$. But $\mathbb{P}(w) \geq 1 - \delta > 1/2$, so on input w, $\mathcal{T}$ answers YES with probability at least $3/4$. This proves that on input w_x, $\mathcal{T}$ answers YES with probability at least $3/4 - \epsilon$, that is at least $5/8$. Now, by a simple majority lemma, it is easy to see that if we run $10 \cdot l$ times an algorithm which answers YES with probability greater than $5/8$ and we check if we have a majority of YES, we will get a positive answer with probability at least $3/4$. Hence r_x is one with probability at least $3/4$, and by construction $r = 1$ with probability at least $3/4$.

Conversely, suppose that no word $w \in \Sigma^*$ of length greater than n is accepted with probability greater than δ. Then for any $x \in \Gamma$, there is no $w \in \Sigma^*$ such that $d(w, w_x) \leq \epsilon$ and $\mathbb{P}_{\mathcal{A}}(w) > 1/2$. Hence for all $x \in \Gamma$, on input w_x, the tester $\mathcal{T}$ will answer NO with probability at least $3/4$. Now, again by a majority lemma, we can see that if we run $10 \cdot l$ times such an algorithm on the same input w_x and we check if we get a majority of YES, we will have a positive answer with probability at most $1/(4 \cdot l)$. Thus, r_x is 1 with probability at most $1/(4 \cdot l)$. Now, by construction and using a union bound, $r = 1$ with probability at most $l \cdot 1/(4 \cdot l)$, that is with probability at most $1/4$. Finally, $r = 0$ with probability at least $3/4$.

We have proved that if $\mathcal{T}$ exists, then we can find $T \in \mathbb{N}$ and an algorithm $\mathcal{T}'$ which works in time T on input $\mathcal{A}$ and which gives answers YES or NO such that:

- If there exists $w \in \Sigma^*$ of length greater than n such that $\mathbb{P}_{\mathcal{A}}(w) \geq 1 - \delta$, then $\mathcal{T}'$ answers YES with probability at least $3/4$.
- If for all $w \in \Sigma^*$ of length greater than n we have $\mathbb{P}_{\mathcal{A}}(w) \leq \delta$, then $\mathcal{T}'$ answers NO with probability at least $3/4$.

This contradicts the undecidability of the emptiness promise problem for PAs. □

5.3. Comparison problems

In [28], the author proves that we can decide whether two PAs accept the same words with the same probabilities in PTIME. However, the following problem, which can be seen as a relaxation of the equivalence problem, is undecidable.

Problem 6 (Approximate equivalence)**.**
Input*: Two PAs $\mathcal{A}_1$ and $\mathcal{A}_2$, $\epsilon \in [0; 1]$.*
Question*: for all $w \in \Sigma^*$, $|\mathbb{P}_1(w) - \mathbb{P}_2(w)| \leq \epsilon$?*

Even if we introduce a relaxation on the input, as in [12], the problem is still undecidable.

Problem 7 (Approximate simulation on close inputs)**.**
Input*: Two PAs $\mathcal{A}_1$ and $\mathcal{A}_2$, $\epsilon \in [0; 1]$.*
Question*: for all $w \in \Sigma^*$, there exists a word $w' \in \Sigma^*$ such that $d(w, w') \leq \epsilon$ and $|\mathbb{P}_1(w) - \mathbb{P}_2(w')| \leq \epsilon$?*

The undecidability of problem 6 is proved in [17], we give the reduction for sake of completeness. We reduce the undecidable emptiness problem for a PA to each one of these problems. Let $\mathcal{A}_\emptyset$ be a probabilistic automaton with no accepting state. Hence $\mathcal{A}_\emptyset$ accepts no word with probability greater than zero, and a given automaton $\mathcal{A}$ will be ϵ-close to $\mathcal{A}_\emptyset$ (for problems 6) iff $\mathcal{A}$ does not accept any word with probability greater than ϵ. Considering problem 7, the automaton $\mathcal{A}$ is ϵ-simulated by $\mathcal{A}_\emptyset$ iff $\mathcal{A}$ does not accept any word with probability greater to ϵ. Since by the first remark of this paragraph it is undecidable to decide if all the finite words are accepted by $\mathcal{A}$ with probability at most ϵ, problem 6 and 7 are undecidable.

6. Ultimate properties

In this section, we present a class of properties on infinite words which satisfy two conditions:

- The properties are *ultimate*, that is if w is a trace and w' is obtained by changing only a finite number of letters in w, then w and w' should satisfy the same properties.
- They are not sensitive to time translation. That is, if w' is a suffix of w, then w and w' satisfy the same properties.

Two purely probabilistic processes will be *ultimately equivalent* if they satisfy the same ultimate properties with the same probabilities. Two MDPs will be ultimately equivalent if for any policy on one of the MDPs there exists a policy on the second MDP such that the induced purely probabilistic processes are ultimately equivalent.

First we will prove that two probabilistic processes are ultimately equivalent iff they are *trace equivalent*, iff they have the same statistic (Theorem 5, 6). Second, we will prove that two weakly communicating MDPs are ultimately equivalent iff their statistic polytopes coincide, that is iff they are distance 0 (Theorem 7).

In order to compare different MDPs, we consider the traces of their runs as before. Let $(\Sigma^\omega, \mathcal{F}, \mathbb{P}^{\sigma,\alpha})$ be the probability space where $\mathcal{F}$ is the σ-field generated on Σ^ω by the cones $C_{a_1a_2\ldots a_l}$ of infinite words starting with letters $a_1, a_2, ..., a_l$. We write $\mathbb{P}^{\sigma,\alpha}$ indifferently for the probability distribution on Ω, as in the previous section, and for the probability distribution induced by Tr on Σ^ω: If $\Gamma \in \mathcal{F}$, $\mathbb{P}^{\sigma,\alpha}(\Gamma)$ is $\mathbb{P}^{\sigma,\alpha}(Tr^{-1}(\Gamma))$. We still write X_n for the random variable on Ω which associates to a run its state at time n. The random variable Y_n, defined on Ω or Σ^ω, associates to a run or a trace its action label at time n.

A *property* is a set $\Gamma \in \mathcal{F}$. Let $T : \Sigma^\omega \longrightarrow \Sigma^\omega$ be defined as $T(w_0w_1...) = (w_1w_2...)$. An *ultimate property* is a property Γ such that $T^{-1}(\Gamma) = \Gamma$. We write $\mathcal{G}$ for the class of ultimate properties. $\mathcal{G}$ is a σ-field, sometimes called the *invariant σ-field* ([1]). Clearly, $\mathcal{G}$ satisfies the two requirements formulated above. A large class of natural properties belong to our class of ultimate properties, we give some examples.

Example 4 (Some ultimate and non-ultimate properties)**.**

- *If $A \subseteq \Sigma$, then $\{A\ i.o.\} = \bigcap_{n\in\mathbb{N}} \bigcup_{m\geq n}\{Y_m \in A\}$ is in $\mathcal{G}$. Here $i.o.$ denotes* infinitely often.
- *If $A \subseteq \Sigma$, then $\{A\ ult.\} = \bigcup_{n\in\mathbb{N}} \bigcap_{m\geq n}\{Y_m \in A\}$ is in $\mathcal{G}$. Here $ult.$ denotes* ultimately.
- *If $A \subseteq \Sigma$, then $\bigcup_{n\in\mathbb{N}} \bigcap_{m\geq n}\{Y_{2\cdot m} \in A\}$ is in $\mathcal{F}_\infty$. However it is not in $\mathcal{G}$: indeed, it is sensitive to translation.*
- *Let $a \in \Sigma$, and $S^a_n = \dfrac{\sum_{i=0}^n I_{Y_i=a}}{n}$. Then $\{lim_{n\to\infty} S^a_n\ exists\}$, i.e. the set of runs r for which the sequence $S^a_n(r)$ converges as n goes to infinity, is in $\mathcal{G}$.*
- *Let $k \in \mathbb{N}$ and $x \in \mathbb{R}^{\Sigma^k}$. Then $\{w | ustat_k(w) = x\} \in \mathcal{G}$.*

For $n \geq 0$, let $\mathcal{F}_n = \mathcal{B}(Y_n, Y_{n+1}...)$ be the smallest σ-field on Σ^ω with respect to which all the $Y_i, i \geq n$ are measurable. Let $\mathcal{F}_\infty = \bigcap_{n\in\mathbb{N}} \mathcal{F}_n$, called the *tail σ-field* of $Y_n, n \geq 0$. Intuitively, an event Γ is in $\mathcal{F}_\infty$ iff changing a finite number of letters of an outcome $w \in \Sigma^\omega$ does not affect the occurrence of the event $r \in \Gamma$. Notice that $\mathcal{G}$ and $\mathcal{F}_\infty$, as σ-fields, are closed under union and intersection. The following result of [1] shows that ultimate properties form a particular class of properties of the tail σ-field.

Proposition 10. *T maps $\mathcal{F}_\infty$ one-to-one onto itself, and $\mathcal{G} = \{\Gamma \in \mathcal{F}_\infty | T \cdot \Gamma = \Gamma\}$.*

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $\Gamma \in \mathcal{F}$. We say that Γ is a $\mathbb{P}$-*atomic* set of $\mathcal{F}$ if $\mathbb{P}(\Gamma) > 0$, and for all $\Gamma' \in \mathcal{F}$ such that $\Gamma' \subseteq \Gamma$ we have $\mathbb{P}(\Gamma') = 0$ or $\mathbb{P}(\Gamma') = \mathbb{P}(\Gamma)$. If $\mathcal{F}$ can be decomposed as a finite union of $\mathbb{P}$-atomic sets for $\mathcal{F}$ and a $\mathbb{P}$-negligible set, we shall say that $\mathcal{F}$ is *finite*. The following theorem is a generalization of the classical Kolmogorov's $0-1$ law to the context of Markov chains.

Theorem 4. *If σ has finite memory $i \in \mathbb{N}$, then the tail σ-field of $Y_n, n \geq 0$, is finite, and the number of $\mathbb{P}^{\sigma,\alpha}$-atomic sets of the invariant σ-field $\mathcal{G}$ does not exceed $|S|^i$.*

Theorem. If σ has finite memory $i \in \mathbb{N}$, then the tail σ-field of $Y_n, n \geq 0$, is finite, and the number of $\mathbb{P}^{\sigma,\alpha}$-atomic sets of the invariant σ-field $\mathcal{G}$ does not exceed $|S|^i$.

Proof. We use our construction of the power MDP, and the results of Blackwell et al [2].
First we prove the result when σ is memoryless. In this case, $(X_n, Y_n)_{n\geq 0}$ is a Markov chain. Suppose first that $(X_n, Y_n)_{n\geq 0}$ is irreducible, and has period $d \geq 1$. We write $\mathbb{P}$ for the probability distribution on runs induced by the chain. We prove that in that case $\mathcal{G}$ is trivial: it can be decomposed as the union of *one* $\mathbb{P}$-atomic set and a set of $\mathbb{P}$-measure zero. Let $\{I_c, c \in [0; d-1]\}$ be a partition of the state space of the chain into its cyclically moving subclasses.That is, I_0 is the set of states in $S \times \Sigma$ accessible from the initial distribution of $(X_n, Y_n)_{n\geq 0}$ in $l \cdot d$ steps for $l \in \mathbb{N}$, and if $c \in [0; d-1]$, I_c is the set of states accessible in $l \cdot d + c$ steps, for an $l \in \mathbb{N}$. In this context, [2] proves that the tail σ-field of the chain is equivalent to the σ-field generated by the sets $\{(X_0, Y_0) \in I_c\}$ for $c \in [0; d-1]$, hence is finite. Let Γ be in $\mathcal{G}_0$, the invariant σ-field of $(X_n, Y_n)_{n\geq 0}$, such that $\mathbb{P}(\Gamma) > 0$. Then, there exists $c \in [0; d-1]$ such that $\{(X_0, Y_0) \in I_c\} \subseteq \Gamma$. But since Γ satisfies by definition $T^{-1}(\Gamma) = \Gamma$, we have that $\{(X_0, Y_0) \in I_{c-1}\} \subseteq \Gamma$, where $c-1 = d-1$ in the case $c = 0$. Since by proposition 10 we have $T \cdot \Gamma = \Gamma$, we get that $\{(X_0, Y_0) \in I_{c+1}\} \subseteq \Gamma$, where $c + 1 = 0$ in the case $c = d - 1$. This proves that any $\Gamma \in \mathcal{G}_0$ with positive measure contains all the generators of $\mathcal{F}_\infty$. This proves that $\mathcal{G}_0$ is trivial. The invariant σ-field $\mathcal{G}$ of $Y_n, n \in \mathbb{N}$ is a projection of $\mathcal{G}_0$: the probability of $\Gamma \in \mathcal{G}$ is $\mathbb{P}(Tr^{-1}(\Gamma))$. This proves that $\mathcal{G}$ is trivial.

Suppose now that $(X_n, Y_n)_{n\geq 0}$ is not irreducible. Let $S \times \Sigma = U_0 \cup U_1 \cup ... \cup U_l$ be the decomposition of the state space of $(X_n, Y_n)_{n\geq 0}$ into its irreducible components. U_0 is the set of transient states, and the U_i, for $i \in [1; l]$, are the state space of the irreducible components. Notice that $l \leq |S|$. Indeed, if $s \in S$ and $a, a' \in \Sigma$, if a and a' are in the support of $\sigma(s)$, then (s, a) and (s, a') are in the same irreducible component of the chain. By the previous result, once the chain has entered one of the $U_i, i \in [1; l]$, it behaves as an irreducible Markov chain and the associated invariant field $\mathcal{G}$ is trivial. This proves that $\mathcal{G}_0$ is equivalent to the tail σ-field generated by the invariant σ-fields $\mathcal{G}_0^i$ associated to the $S_i, i \in [1; l]$. Since each $\mathcal{G}_0^i$ is trivial, this proves that the number of atomic sets of $\mathcal{G}_0$ is at most $|S|$.As before, this implies that the number of atomic sets of $\mathcal{G}$ is at most $|S|$.

Suppose finally that σ has finite memory $i \in \mathbb{N}^*$. Then σ induces a Markov chain on $\mathcal{S}^i$, that we write $\mathcal{S}^{i,\sigma}$. We write $\mathcal{S}^\sigma$ for the probabilistic process induced on $\mathcal{S}$ by σ. Then, the tail σ-fields on $\mathcal{S}^\sigma$ and $\mathcal{S}^{i,\sigma}$ coincide, and the invariant σ-fields on $\mathcal{S}^\sigma$ and $\mathcal{S}^{i,\sigma}$ coincide, and we can use the past results to prove that the number of atomic subsets of the invariant σ-field $\mathcal{G}$ of $\mathcal{S}^{i,\sigma}$ does not exceed $|S|^i$. □

Theorem 4 is a generalization of results of [2]. Notice that if σ has infinite memory, $\mathcal{F}_\infty$ is not finite in general, and can be $\mathbb{P}^{\sigma,\alpha}$-completely non-atomic, that is it may contain no atomic subsets. The proof of theorem 4 gives directly the following corollary.

Corollary 3 ($\mathcal{G}$ is always trivial for an irreducible chain)**.** *Let $\mathcal{S}$ be an irreducible LTMC. Then for all $\Gamma \in \mathcal{G}$, $\mathbb{P}^\alpha(\Gamma) \in \{0, 1\}$, and $\mathbb{P}^\alpha(\Gamma)$ is independent of α.*

6.1. Markov chains with labelled transitions

In this subsection we present results concerning ultimate properties associated to homogeneous Markov chains. A memoryless policy σ on an MDP $\mathcal{S} = (S, \Sigma, P)$ induces on the state space $S \times \Sigma$ a Labelled Transition Markov Chain $\mathcal{S}^\sigma$ (written LTMC). This chain $(X_n, Y_n)_{n\in\mathbb{N}}$, with $(X_n, Y_n) \in S \times \Sigma$ for all $n \in \mathbb{N}$, has transition probabilities:

$$\mathbb{P}[(X_{n+1}, Y_{n+1}) = (x', y')|(X_n, Y_n) = (x, y)] = P(x'|x, y) \cdot \sigma(x')(y')$$

Let $\mathcal{S} = (X_n, Y_n)_{n\in\mathbb{N}}$ be an irreducible LTMC, with initial distribution α. Let $k \in \mathbb{N}$. Since the chain is supposed irreducible, there exists a vector in $\mathbb{R}^{(S\times\Sigma)^k}$, which we call $ustat_k(\mathcal{S})$, or the k-gram of the chain, such that with probability one the k-gram of the trace of a run on $\mathcal{S}$ converges to $ustat_k(\mathcal{S})$. Moreover, $ustat_k(\mathcal{S})$ is independent of the initial distribution α ([26], chapter 1). In the following $\mathcal{S}_1, \alpha_1$ and $\mathcal{S}_2, \alpha_2$ are two LTMCs. We write $\mathbb{P}_1^{\alpha_1}$ and $\mathbb{P}_2^{\alpha_2}$ for the associated probability distributions on Σ^ω.

Definition 11 (Ultimate equivalence)**.** *$\mathcal{S}_1, \alpha_1$ and $\mathcal{S}_2, \alpha_2$ are said to be* ultimately equivalent, *written $\mathcal{S}_1, \alpha_1 \sim_u \mathcal{S}_2, \alpha_2$, if for all $\Gamma \in \mathcal{G}$, $\mathbb{P}_1^{\alpha_1}(\Gamma) = \mathbb{P}_2^{\alpha_2}(\Gamma)$*

If $\mathcal{S}_1, \alpha_1 \sim_u \mathcal{S}_2, \alpha_2$, then for any $\Gamma \in \mathcal{G}$, Γ is $\mathbb{P}_1$-atomic iff Γ is $\mathbb{P}_2$-atomic. If $\mathcal{S}$ is an irreducible LTMC with initial distribution α, then for all $\Gamma \in \mathcal{G}$, $\mathbb{P}^\alpha(\Gamma) \in \{0,1\}$, and $\mathbb{P}^\alpha(\Gamma) \in \{0,1\}$ is independent of α. (by corollary 3). As a consequence, given two irreducible LTMCs $\mathcal{S}_1$ and $\mathcal{S}_2$, we can use the notation $\mathcal{S}_1 \sim_u \mathcal{S}_2$ to say that they are ultimately equivalent.

Definition 12 (Trace equivalence for LTMCs)**.** *$\mathcal{S}_1$ and $\mathcal{S}_2$ are* trace equivalent *if there exists two initial distributions α_1 and α_2, on $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively, such that:*

For all $w \in \Sigma^*$, $\mathbb{P}_1^{\alpha_1}(C_w) = \mathbb{P}_2^{\alpha_2}(C_w)$.

The following theorem shows that it is enough to know the k-grams of an irreducible chain for a bounded number of k's, to characterize completely the ultimate properties of the chain.

Theorem 5. *Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two irreducible LTMCs. Then the following are equivalent:*

1. $\forall k \in [1; (|S_1| + |S_2|)^2]$, $ustat_k(\mathcal{S}_1) = ustat_k(\mathcal{S}_2)$
2. *For all $k \in \mathbb{N}$, $ustat_k(\mathcal{S}_1) = ustat_k(\mathcal{S}_2)$*
3. $\mathcal{S}_1 \sim_u \mathcal{S}_2$
4. *$\mathcal{S}_1$ and $\mathcal{S}_2$ are trace equivalent.*

Proof. $1 \Leftrightarrow 2$ is the analogous of lemma 6.
We prove that $2 \Rightarrow 4 \Rightarrow 3 \Rightarrow 2$.
Suppose 2. $\mathcal{S}_1$ and $\mathcal{S}_2$ are irreducible Markov chains, so there exists unique stationary distributions α_1 and α_2 on $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively. (α_1, α_2 are distributions on $|S_1 \times \Sigma|$ and $|S_2 \times \Sigma|$). Let $w \in \Sigma^*$. Then

$\mathbb{P}_1^{\alpha_1}(C_w) = \sum_{t \in (S_1 \times \Sigma)} \alpha_1(t) \cdot \mathbb{P}_1(w|(X_0, Y_0) = t)$
$= ustat_{|w|}(\mathcal{S}_1)(w)$ by definition.
$= ustat_{|w|}(\mathcal{S}_2)(w)$ by 2
$= \sum_{t \in (S_2 \times \Sigma)} \alpha_2(t) \cdot \mathbb{P}_1(w|(X_0, Y_0) = t)$
$= \mathbb{P}_2^{\alpha_2}(C_w)$

Hence 4.
Suppose 4. Using corollary 3, we know that given $\Gamma \in \mathcal{G}$ and α_1, α_2 initial distributions on $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively, $\mathbb{P}_1^{\alpha_1}(\Gamma)$ and $\mathbb{P}_2^{\alpha_2}(\Gamma)$ are independent of α_1 and α_2. Suppose that we chose α_1 and α_2 to be the initial distributions on $\mathcal{S}_1$ and $\mathcal{S}_2$ respectively such that $\mathcal{S}_1, \alpha_1$ and $\mathcal{S}_2, \alpha_2$ are trace equivalent. Clearly, two trace equivalent Markov chains associate the same probabilities to any property, and in particular to any ultimate property. This proves 3.
$3 \Rightarrow 2$ is simple, since we have seen in example 4 that given $x \in \mathbb{R}^{\Sigma^k}$, the property that $ustat_k(w) = x$ is a property in $\mathcal{G}$. ☐

Given $\mathcal{S}$ a general LTMC on state spaces S with initial distribution α, let $S = S_0 \cup S_1 \cup ... \cup S_l$ be its decomposition into irreducible components: S_0 is the set of transient states, and the $S_i, i \in [1; l]$ are the irreducible components of the chain. Each S_i gives an irreducible LTMC $\mathcal{S}_i$. Given $i \in [1; l]$, let $Reach(S_i)$ be the set of infinite runs on $\mathcal{S}$ which enter S_i eventually (and then never leave it), and $p_i = \mathbb{P}^\alpha(Reach(S_i))$. The p_i sum to one. Let $\lambda_{S_i} : \mathcal{G} \to \{0,1\}$ be such that for $\Gamma \in \mathcal{G}$, $\lambda_{S_i}(\Gamma)$ is the probability that a run executed on $\mathcal{S}_i$ is in Γ. By corollary 3, the λ_{S_i} are well defined and take values in $\{0,1\}$.

Lemma 5. *Let $\mathcal{S}, \alpha$ be an LTMC and $\Gamma \in \mathcal{G}$. Then*

$$\mathbb{P}^\alpha(\Gamma) = \sum_{i=1}^{l} \mathbb{P}^\alpha(Reach(S_i)) \cdot \lambda_{S_i}(\Gamma)$$

Proof. For all $n \in \mathbb{N}$ we have $\mathbb{P}^\alpha(\Gamma) = \sum_{i=0}^{l} \mathbb{P}^\alpha(\Gamma \wedge X_n \in S_i)$, hence

$$\mathbb{P}^\alpha(\Gamma) = \sum_{i=0}^{l} \mathbb{P}^\alpha(\Gamma | X_n \in S_i) \cdot \mathbb{P}^\alpha(X_n \in S_i).$$

But for all $i \in [1;l]$, $\mathbb{P}^\alpha(X_n \in S_i)$ goes to $\mathbb{P}^\alpha(Reach(S_i))$ as n goes to infinity, and corollary 3 proves that $\mathbb{P}^\alpha(\Gamma | X_n \in S_i) = \lambda_{S_i}(\Gamma)$. Since by definition $\mathbb{P}^\alpha(X_n \in S_0)$ goes to zero as n goes to infinity, this implies that $\mathbb{P}^\alpha(\Gamma) = \sum_{i=1}^{l} \mathbb{P}^\alpha(Reach(S_i)) \cdot \lambda_i(\Gamma)$. ☐

Let $\mathcal{S}_1, \alpha_1$ and $\mathcal{S}_2, \alpha_2$ be two LTMCs on state spaces S_1 and S_2. Write $S_1 = S_0^1 \cup S_1^1 \cup ... \cup S_{l_1}^1$ and $S_2 = S_0^2 \cup S_1^2 \cup ... \cup S_{l_2}^2$ for the decompositions into irreducible components, and $Reach_j(S_i^j)$ for the set of infinite runs on $\mathcal{S}_j$ which enter S_i^j eventually. Let $p_i^j = \mathbb{P}_j(Reach_j(S_i^j))$. The irreducible components S_i^j can be seen as irreducible LTMC $\mathcal{S}_i^j$, and $\sim_u$ is an equivalence relation on $\{\mathcal{S}_i^j, j \in \{0,1\}, i \in [1; l_j]\}$. Write $\{T_1, ..., T_l\}$ for the equivalence classes of $\sim_u$. If $i \in [1;l]$, T_i is a union of S_j^1 and $S_{j'}^2$. The next theorem summarizes our results on LTMCs.

Theorem 6. *Let $\mathcal{S}_1, \alpha_1$ and $\mathcal{S}_2, \alpha_2$ be two LTMCs. Then the following are equivalent:*

1. $\mathcal{S}_1, \alpha_1 \sim_u \mathcal{S}_2, \alpha_2$
2. $\forall i \in [1;l]\ \mathbb{P}_1^{\alpha_1}(Reach_1(T_i)) = \mathbb{P}_2^{\alpha_2}(Reach_2(T_i))$

Proof. For $i \in [1;l]$, let λ_{T_i} be identified with one of the function λ_{S_j}, with $S_j \in T_i$ (By definition of $\sim_u$, all the λ_{S_j} coincide, for $S_j \in T_i$).

- Suppose 2. Let $\Gamma \in \mathcal{G}$. Using lemma 5, we have $\mathbb{P}_1^{\alpha_1}(\Gamma) = \sum_{i=1}^{l_1} \mathbb{P}_1^{\alpha_1}(Reach(S_i)) \cdot \lambda_{S_i}^1(\Gamma)$. But $\sum_{i=1}^{l_1} \mathbb{P}_1^{\alpha_1}(Reach(S_i)) \cdot \lambda_{S_i}^1(\Gamma) = \sum_{i=1}^{l} \mathbb{P}_1^{\alpha_1}(Reach(T_i)) \cdot \lambda_{T_i}(\Gamma)$. By 2, this implies $\mathbb{P}_1^{\alpha_1}(\Gamma) = \sum_{i=1}^{l_1} \mathbb{P}_2^{\alpha_2}(Reach(T_i)) \cdot \lambda_{T_i}(\Gamma)$, i.e. $\mathbb{P}_1^{\alpha_1}(\Gamma) = \mathbb{P}_2^{\alpha_2}(\Gamma)$, so $\mathcal{S}_1, \alpha_1 \sim_u \mathcal{S}_2, \alpha_2$.
- Conversely, suppose that there exists $i_0 \in [1;l]$ such that $\mathbb{P}_1^{\alpha_1}(Reach_1(T_{i_0})) \neq \mathbb{P}_2^{\alpha_2}(Reach_2(T_{i_0}))$. Given $i \neq j$ in $[1;l]$, we know that the irreducible components in T_i are not ultimately equivalent to the irreducible components in T_j. Hence, since $\mathcal{G}$ is closed under complementation, for $i \neq j$ in $[1;l]$, there exists $\Gamma_{i,j} \in \mathcal{G}$ such that $\lambda_{T_i}(\Gamma_{i,j}) = 1$ and $\lambda_{T_j}(\Gamma_{i,j}) = 0$.
  Now, let $\Gamma = \bigcap_{j=1}^{l} \Gamma_{i_0,j}$. Then, $\lambda_{T_i}(\Gamma) = 1$ if $i = i_0$, and zero elsewhere. using lemma 5, we have that $\mathbb{P}_1^{\alpha_1}(\Gamma) = \sum_{i=1}^{l} \mathbb{P}_1^{\alpha_1}(Reach_1(T_i)) \cdot \lambda_{T_i}(\Gamma)$, hence $\mathbb{P}_1^{\alpha_1}(\Gamma) = \mathbb{P}_1^{\alpha_1}(Reach(T_{i_0}))$. In the same way, $\mathbb{P}_2^{\alpha_2}(\Gamma) = \mathbb{P}_2^{\alpha_2}(Reach_2(T_{i_0}))$. By hypothesis, we get $\mathbb{P}_2^{\alpha_2}(\Gamma) \neq \mathbb{P}_1^{\alpha_1}(\Gamma)$, thus $\mathcal{S}_1, \alpha_1 \not\sim_u \mathcal{S}_2, \alpha_2$.

☐

6.2. Ultimate simulation and equivalence for MDPs

In this section we compare the long term behaviors of MDPs, by comparing the Markovian processes induced on their state spaces by policies. Given a memoryless policy σ on an MDP $\mathcal{S} = (S, \Sigma, P)$, it induces an LTMC $\mathcal{S}^\sigma$ on S. If σ has memory $i \in \mathbb{N}$, we can see σ as a memoryless policy on $\mathcal{S}^i$, and we write also $\mathcal{S}^\sigma$ for the LTMC it induces on $\mathcal{S}^i$.

The class of weakly communicating MDPs will play the role of the irreducible Markov chains of the last subsection. If we consider only i-memory policies, the ultimate simulation relation between two weakly communicating MDPs would depend on the initial distributions of the MDPs. To tackle this problem, if $i \in \mathbb{N}$ we define the class $UR(i)(\mathcal{S})$ of the *ultimately memory i policies* on $\mathcal{S}$, introduced in [16] in the case $i = 0$. A policy σ is in $UR(i)(\mathcal{S})$ if there exists a policy $\sigma^\infty \in MR(i)(\mathcal{S})$, called the *tail* of σ, and a random stopping time τ on Ω, called the *switching time* of σ, such that:

- If $r \in \Omega$, if $n \geq \tau(r)$ then $\sigma(r_{|n}) = \sigma^\infty(r_{|n})$.
- $\mathbb{P}^\sigma(\{r | \tau(r) < \infty\}) = 1$.

In other words, σ is in $UR(i)(\mathcal{S})$ if with probability one, after a finite number of steps, σ behaves as a policy of memory at most i. We prove a generalization of theorem 1 for policies in $UR(i)$: if $\mathcal{S}$ is weakly communicating, then for all $k, i \in \mathbb{N}$ and all initial distribution α on $\mathcal{S}$, $\mathcal{H}_k^{UR(i)}(\alpha)(\mathcal{S}) = \overline{\mathcal{H}_k^{MR(i)}}(\mathcal{S})$.

Proposition 11. *Suppose that $\mathcal{S}$ is weakly communicating. Then for all $k, i \in \mathbb{N}$ and all initial distribution α on $\mathcal{S}$, we have $\mathcal{H}_k^{UR(i)}(\alpha)(\mathcal{S}) = \overline{\mathcal{H}_k^{MR(i)}(\alpha)(\mathcal{S})}$).*

Proof. Let α_0 be an initial distribution on S.
First, given an initial distribution α on $\mathcal{S}$, the inclusion $\mathcal{H}_k^{MR(i)}(\alpha)(\mathcal{S})) \subset \overline{\mathcal{H}_k^{UR(i)}(\alpha)}(\mathcal{S})$ is direct. Then, by a direct generalization of the results of [9], we know that if $\alpha_1 \in \Delta(S)$, then $\overline{\mathcal{H}_k^{MR(i)}(\alpha_1)(\mathcal{S}))} = \bigcup_{\alpha \in \Delta(S)} \mathcal{H}_k^{MR(i)}(\alpha)(\mathcal{S}))$. By proposition 12, we know that $\mathcal{H}_k^{UR(i)}(\alpha)(\mathcal{S})$ is independent of α. This implies that $\overline{\mathcal{H}_k^{MR(i)}(\alpha_0)(\mathcal{S}))} \subseteq \mathcal{H}_k^{UR(i)}(\alpha_0)(\mathcal{S})$.
Conversely, let $x \in \mathcal{H}_k^{UR(i)}(\alpha_0)(\mathcal{S})$, and $\sigma \in UR(i)(\mathcal{S})$ such that $x = x^{\infty}_{\sigma,\alpha_0,k}$. Let σ^∞ be the tail of σ. Let S_i be an irreducible component of $\mathcal{S}^{\sigma^\infty}$ such that $\mathbb{P}^{\sigma,\alpha_0}(Reach(S_i)) > 0$. Let $\alpha_1 \in \Delta(S)$ with its support on S_i. Then we have: $x^{\infty}_{\sigma,\alpha_0,k} = x^{\infty}_{\sigma^\infty,\alpha_1,k}$. That is, $x \in \mathcal{H}_k^{MR(i)}(\alpha_1)(\mathcal{S}))$. This proves the result, since $\overline{\mathcal{H}_k^{MR(i)}(\alpha_0)(\mathcal{S}))} = \bigcup_{\alpha \in \Delta(S)} \mathcal{H}_k^{MR(i)}(\alpha)(\mathcal{S}))$ □

We write $\mathcal{S}^\sigma$ for the probabilistic process with labelled transitions (which is not any more a Markov chain), induced by σ on $\mathcal{S}$. If α is an initial distribution on $\mathcal{S}$, $\mathbb{P}^{\sigma,\alpha}$ is the associated probability distribution on the set of runs. As in definition 11, two processes $\mathcal{S}_1^{\sigma_1}, \alpha_1$ and $\mathcal{S}_2^{\sigma_2}, \alpha_2$ are said to be *ultimately equivalent* if for all $\Gamma \in \mathcal{G}$ we have $\mathbb{P}^{\sigma_1,\alpha_1}(\Gamma) = \mathbb{P}^{\sigma_2,\alpha_2}(\Gamma)$.

The analogous of theorem 6 holds: $\mathcal{S}_1^{\sigma_1}, \alpha_1 \sim_u \mathcal{S}_2^{\sigma_2}, \alpha_2$ iff for all $i \in [1;l]$, $\mathbb{P}_1^{\sigma_1,\alpha_1}(Reach_1(T_i)) = \mathbb{P}_2^{\sigma_2,\alpha_2}(Reach_2(T_i))$. Here the T_i are equivalence classes on the set of irreducible components of $\mathcal{S}_1^{\sigma_1^\infty}$ and $\mathcal{S}_2^{\sigma_2^\infty}$

Definition 13 (Simulation between MDPs). $\mathcal{S}_1, \alpha_1$ *is said to be* i-memory ultimately simulated *by* $\mathcal{S}_2, \alpha_2$*, written* $\mathcal{S}_1, \alpha_1 \prec_u^i \mathcal{S}_2, \alpha_2$*, if for all* $\sigma_1 \in UR(i)(\mathcal{S}_1)$*, there exists* $\sigma_2 \in UR(i)(\mathcal{S}_2)$ *such that* $\mathcal{S}_1^{\sigma_1}, \alpha_1 \sim_u \mathcal{S}_2^{\sigma_2}, \alpha_2$.

$\mathcal{S}_1, \alpha_1$ and $\mathcal{S}_2, \alpha_2$ are said to be i*-memory ultimately equivalent*, written $\mathcal{S}_1, \alpha_1 \sim_u^i \mathcal{S}_2, \alpha_2$, if $\mathcal{S}_1, \alpha_1 \prec_u^i \mathcal{S}_2, \alpha_2$ and $\mathcal{S}_2, \alpha_2 \prec_u^i \mathcal{S}_1, \alpha_1$. As for irreducible LTMC, the simulation relation between weakly communicating MDPs does not depend on the initial distributions.

Proposition 12. *Suppose* $\mathcal{S}_1$ *and* $\mathcal{S}_2$ *weakly communicating. Let* α_1, α_1' *and* α_2, α_2' *be initial distributions on* S_1 *and* S_2*. Then* $\mathcal{S}_1, \alpha_1 \prec_u^i \mathcal{S}_2, \alpha_2$ *iff* $\mathcal{S}_1, \alpha_1' \prec_u^i \mathcal{S}_2, \alpha_2'$.

Proof. This proposition is a consequence of the fact that if an MDP $\mathcal{S}$ is weakly communicating and α, α' are initial distributions on $\mathcal{S}$, then for all $\sigma \in UR(i)(\mathcal{S})$, there exists $\sigma' \in UR(i)(\mathcal{S})$ such that $\mathcal{S}^\sigma, \alpha \sim_u \mathcal{S}^{\sigma'}, \alpha'$. To prove this, let $\mathcal{S}$ be a weakly communicating MDP, α, α' two initial distributions on $\mathcal{S}$, and $\sigma \in UR(i)(\mathcal{S})$. Let $\sigma^\infty \in MR(i)(\mathcal{S})$ be the tail of σ, and let $S = S_0 \cup S_1 \cup ... \cup S_l$ be the decomposition of S into the maximal irreducible components of the Markov chain $\mathcal{S}^{\sigma^\infty}$ induced by σ^∞ on $\mathcal{S}$. The set S_0 is the set of transient states of this chains, and the $S_i, i \in [1;l]$ are the classes of recurrent states. For $i \in [1;l]$, let $\lambda_i = \mathbb{P}^{\sigma,\alpha}(Reach(S_i))$. Since $\mathcal{S}$ is weakly communicating, using the results of [7], for all $i \in [1;l]$ there exists $\sigma_i \in SR(\mathcal{S})$ such that $\mathbb{P}^{\sigma_i,\alpha'}(Reach(S_i)) = 1$. The idea for the construction of our σ' is, for all $i \in [1;l]$, to take σ' to behave like σ_i with probability λ_i, until we have reach a state in S_i, and then to behave like σ^∞. Then the invariant σ-fields of the two processes will be equivalent, using corollary 3. The switching time of σ' is finite with probability one, since with probability one, using policy σ_i, we reach S_i in finite time. □

This allows the notation $\mathcal{S}_1 \prec_u^i \mathcal{S}_2$ if $\mathcal{S}_1$ and $\mathcal{S}_2$ are weakly communicating.

The following lemma is reminiscent of the algorithm of Tzeng, [28]. If $\mathcal{S}, \alpha$ is an MDP and $\sigma \in HR(\mathcal{S})$, if there exists $x \in \mathbb{R}^{\Sigma^k}$ such that with probability one the k-gram of the trace of the prefix of a run on the process induced by σ and α converges to x, then we write $ustat_k(\mathcal{S}^\sigma, \alpha) = x$.

Lemma 6. *Let* $\mathcal{S}_1$ *be an irreducible LTMC, and* $\mathcal{S}_2, \alpha$ *be an MDP. Suppose*

$$\forall k \in [1; (|S_1| + |S_2|)^2], \exists \sigma_k \in SR(\mathcal{S}_2) \mid ustat_k(\mathcal{S}_1) = ustat_k(\mathcal{S}_2^{\sigma_k}, \alpha).$$

Then

$$\exists \sigma \in SR(\mathcal{S}) \mid \forall k \in \mathbb{N},\ ustat_k(\mathcal{S}_1) = ustat_k(\mathcal{S}_2^\sigma, \alpha).$$

Proof. The proof of this lemma is reminiscent of Tzeng [28].
Let $S_1\{s_1, ..., s_{n_1}\}$ and $S_2 = \{s'_1, ..., s'_{n_2}\}$ be the state spaces of $\mathcal{S}_1$ and $\mathcal{S}_2$.

If $a \in \Sigma$ and $\sigma \in SR(\mathcal{S}_2)$, we define the matrices $M^1_a = [\mathbb{P}_1(s_j|s_i, a)]_{i,j\in[1;n_1]}$ and $M^{2,\sigma}_a = [P(s'_j|s'_i, a) \cdot \sigma(s'_i)(a)]_{i,j\in[1;n_2]}$. Remark that $P(s'_j|s'_i, a) \cdot \sigma(s'_i)(a)$ is the probability to go from state s'_i to state s'_j on $\mathcal{S}_2$ if we use the policy σ. If $w = a_1a_2...a_l \in \Sigma^*$, $M^i_w = \prod_{i=1}^{l} M^i_{a_i}$. If $w \in \Sigma^*$ and $s \in S_1$, $\mathbb{P}_1(w|s)$ is the probability on $\mathcal{S}_1$ to read the word w in $|w|$ steps when the system is initiated on state s. If $\sigma \in SR(\mathcal{S})$, $\mathbb{P}^\sigma_2(w|s)$ is the analogous on $\mathcal{S}_2$ where σ resolves the non determinism. Let $E^1 \in \mathbb{R}^{n_1}$ and $E^2 \in \mathbb{R}^{n_2}$ be the column vectors with only ones on their component. If $s \in S_1$ let $I^1_s \in \mathbb{R}^{n_1}$ be the line vector with zeros on its components, except on component corresponding to state s where there is a one, and $I^2_{s'} \in \mathbb{R}^{n_2}$ is the analogous with $s' \in S_2$. If $w \in \Sigma^*$ and $\sigma \in SR(\mathcal{S}_2)$, we have:

$$\mathbb{P}_1(w|s) = I^1_s \cdot M^1_w \cdot E^1$$
$$\mathbb{P}^\sigma_2(w|s) = I^2_s \cdot M^{2,\sigma}_w \cdot E^2$$

Let $\mu_1 \in \Delta(S_1)$ be the limit distribution on the states of the irreducible Markov chain $\mathcal{S}_1$ (If $\mathcal{S}_1$ is periodic, take the average of the limit distributions). If $\sigma \in SR(\mathcal{S}_2)$ and $ustat_k(\mathcal{S}^\sigma_2, \alpha)$ is well defined, let μ^σ_2 be the limit distribution on the states of the chain $\mathcal{S}^\sigma_2$. In that case, if $w \in \Sigma^k$, we have:

$$ustat_k(\mathcal{S}_1)[w] = \sum_{s\in S_1} \mu_1(s) \cdot \mathbb{P}_1(w|s)$$
$$ustat_k(\mathcal{S}^\sigma_2, \alpha)[w] = \sum_{s\in S_2} \mu^\sigma_2(s) \cdot \mathbb{P}^\sigma_2(w|s).$$

Using the previous equations for $\mathbb{P}_1(w|s)$ and $\mathbb{P}^\sigma_2(w|s)$, $ustat_k(\mathcal{S}_1)[w] = ustat_k(\mathcal{S}^\sigma_2, \alpha)[w]$ can be rewritten:

$$(\mu_1, \mu_2) \cdot \begin{pmatrix} M^1_w, 0 \\ 0, M^2_w \end{pmatrix} \cdot \begin{pmatrix} E_1 \\ -E_2 \end{pmatrix} = 0$$

Now, for $a \in \Sigma$, the matrix $\begin{pmatrix} M^1_a, 0 \\ 0, M^2_a \end{pmatrix}$ has size $|S_1| + |S_2|$. Hence the dimension of the vector space generated by the $\begin{pmatrix} M^1_w, 0 \\ 0, M^2_w \end{pmatrix}$,$w \in \Sigma^*$, is at most $(|S_1| + |S_2|)^2$. This proves that if we can find a memoryless policy σ on $\mathcal{S}$ such that for all $k \in [1; (|S_1| + |S_2|)^2]$ we have $ustat_k(\mathcal{S}_1) = ustat_k(\mathcal{S}^\sigma_2)$, then this is true for all $k \in \mathbb{N}$.

□

The following theorem summarizes the different notions presented in this paper: polytopes, distance, ultimate properties.

Theorem 7. *Let $\mathcal{S}_1$ and $\mathcal{S}_2$ be two weakly communicating MDPs. Then the following are equivalent:*

- $\mathcal{S}_1 \prec^i_u \mathcal{S}_2$
- *For all* $k \in \mathbb{N}$, $\Pi^i_k(\mathcal{S}_1) \subseteq \Pi^i_k(\mathcal{S}_2)$.
- *For all* $k \in [1; (|S_1| + |S_2|)^i]$, $\Pi^i_k(\mathcal{S}_1) \subseteq \Pi^i_k(\mathcal{S}_2)$.
- $d^{\prec}_{(|S_1|+|S_2|)^i}(\mathcal{S}_1, \mathcal{S}_2) = 0$.

Proof. We prove these equivalences in the case $i = 0$. For $i > 0$ we just have to consider the i-th powers of the considered systems.
Suppose 1. Let $k \in \mathbb{N}$ and $y \in \Pi^0_k(\mathcal{S}_1)$. Then $\exists \sigma_1 \in SR(\mathcal{S})$ and an initial distribution α_1 on $\mathcal{S}_1$, such that $\pi(\hat{x}^T_k)$ converges to y as T goes to infinity, $\mathbb{P}^{\sigma_1,\alpha_1}$ a.s.. Now, let

$$\Gamma = \{r \in \Omega | lim_{T\to\infty} ustat_k(Tr(r_{|T})) = y\}$$

Then $\Gamma \in \mathcal{G}$, and $\mathbb{P}^{\sigma_1,\alpha_1}(\Gamma) = 1$. Since $\mathcal{S}_1 \prec_u^0 \mathcal{S}_2$, we can find $\sigma_2 \in UR(0)(\mathcal{S}_2)$ and an initial distribution α_2 on $\mathcal{S}_2$ such that $\mathcal{S}^{\sigma_1}, \alpha_1 \sim_u \mathcal{S}^{\sigma_2}, \alpha_2$. Hence $\mathbb{P}^{\sigma_2,\alpha_2}(\Gamma) = 1$. This proves that $\mathbb{P}^{\sigma_2,\alpha_2}$ a.s., $\pi(\hat{x}_k^T)$ converges to y, as T goes to infinity. Thus, $y \in \pi(H^{UR(0)}(\alpha_2)(\mathcal{S}_2))$. By proposition 11, $y \in \pi(\overline{H^{MR(0)}(\alpha_2)(\mathcal{S}_2)})$, i.e. $y \in \Pi_k^0(\mathcal{S}_2)$. So $1 \Rightarrow 2$.

Conversely, suppose 2: for all $k \in \mathbb{N}$, $\Pi_k^0(\mathcal{S}_1) \subseteq \Pi_k^0(\mathcal{S}_2)$. Let α_1 and α_2 be initial distributions on $\mathcal{S}_1$ and $\mathcal{S}_2$, and $\sigma_1 \in UR(0)(\mathcal{S}_1)$. We have to prove that there exists $\sigma_2 \in UR(0)(\mathcal{S}_2)$ such that $\mathcal{S}_1^{\sigma_1}, \alpha_1 \sim_u \mathcal{S}_2^{\sigma_2}, \alpha_2$. Let σ_1^∞ be the tail of σ_1, ans write $S_1 = S_0^1 \cup S_1^1 \cup ... \cup S_{l_1}^1$ for the decomposition into irreducible components of the Markov chain induced by σ_1^∞ on $\mathcal{S}_1$. The MDP $\mathcal{S}_2$ is weakly communicating, and for all $k \in \mathbb{N}$ we have $\Pi_k^0(\mathcal{S}_1) \subseteq \Pi_k^0(\mathcal{S}_2)$. By proposition 11, we know that $\Pi_k^0(\mathcal{S}_2) \subseteq \pi(H_k^{UR(0)}(\alpha)(\mathcal{S}_2))$. Thus, for all $j \in [1; l_1]$ and $k \in \mathbb{N}$, there exists $\sigma_2^{k,j} \in UR(0)(\mathcal{S}_2)$ such that $ustat_k(S_j^1) = ustat_k(\mathcal{S}_2^{\sigma_2^{k,j}}, \alpha_2)$. By lemma 6, for all $j \in [1; l_1]$, there exists a policy $\sigma_2^j \in UR(0)(\mathcal{S}_2)$ such that for all $k \in \mathbb{N}$, $ustat_k(\mathcal{S}_j^1) = ustat_k(\mathcal{S}_2^{\sigma_2^j}, \alpha_2)$. If $j \in [1; l_1]$, let $\sigma_2^{j,\infty}$ be the tail of σ_2^j. Let S_2^j be an irreducible component of the chain induced by $\sigma_2^{j,\infty}$ on $\mathcal{S}_2$, such that $\mathbb{P}^{\sigma_2^j,\alpha_2}(Reach(S_2^j)) > 0$. Then, for all $j \in [1; l_1]$ and $k \in \mathbb{N}$, $ustat_k(\mathcal{S}_j^1) = ustat_k(\mathcal{S}_2^i)$. Using theorem 5, we have that for all $j \in [1; l_1]$, $\mathcal{S}_j^1 \sim_u \mathcal{S}_2^i$. So far, we have built policies on $\mathcal{S}_2$ which are ultimately memoryless, and which induce ultimate Markov chains which are equivalent to the Markov chains induced on $\mathcal{S}_1$ by σ_1. All we have to do now is to build a policy σ_2 on $\mathcal{S}_2$ which takes the behaviors of the policies σ_2^j with the good probabilities. For all $j \in [1; l_1]$, let $p_j = \mathbb{P}^{\sigma_1,\alpha_1}(Reach(S_j^1))$. Then σ_2 is as follows: at the first step, chose with probability distribution $p_j, j \in [1; l_1]$, the policy among the σ_2^i that it will mimic. Once this choice has been made, continue with to mimic the same policy forever. Then, $\mathcal{S}_1^{\sigma_1}, \alpha_1 \sim_u \mathcal{S}_2^{\sigma_2}, \alpha_2$. So we get $2 \Rightarrow 1$.

$2 \Leftrightarrow 3$ follows from lemma 6.

$4 \Leftrightarrow 3$ is the definition of the distance d. $\square$

Two weakly communicating MDPs are equivalent according to the relation $\sim_u^i$ induced by $\prec_u^i$ iff their polytopes coincide. The study of the ultimate properties of a general MDP can be done by studying the ultimate properties of its maximal end components, and the probabilities to reach these end components. The maximal end components play the role for MDPs that the irreducible components play for LTMCs.

7. Conclusion

We introduced Property and Equivalence Testing to approximate classical hard problems on the long term behavior of MDPs, and characterized Equivalent systems with the class of ultimate properties. These methods do not generalize to Probabilistic Automata. Potential applications are the approximate verification of quantitative properties of large probabilistic systems and future work will study how these methods may work with compact representations and with partially observed MDPs.

References

[1] A.F. Abrahamse. The tail field of a Markov Chain. *The Annals of Mathematical Statistics*, 40(1):127–136, 1969.

[2] D. Blackwell and D. Freedman. The tail σ-field of a Markov chain and a theorem of Orey. *Ann. Math. Stat*, 35:1291–1295, 1964.

[3] V.D. Blondel. Undecidable Problems for Probabilistic Automata of Fixed Dimension. *Theory of Computing Systems*, 36(3):231–245, 2008.

[4] V.D. Blondel and J.N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica-Oxford-*, 36:1249–1274, 2000.

[5] C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *JACM*, 42(4):857–907, 1995.

[6] L. de Alfaro. Formal Verification of Probabilistic Systems. PhD Thesis. *Technical Report: CS-TR-98-1601, Stanford University*, 1998.

[7] L. de Alfaro. Computing Minimum and Maximum Reachability Times in Probabilistic Systems. *LNCS*, pages 66–81, 1999.

[8] M. de Rougemont and M. Tracol. Statistic Analysis for Probabilistic Processes. In *LICS '09: Proceedings of the 24nd Annual IEEE Symposium on Logic in Computer Science*, 2009.

[9] C. Derman. *Finite State Markovian Decision Processes*. Academic Press, Inc. Orlando, FL, USA, 1970.

[10] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004.

[11] L. Doyen, T.A. Henzinger, J. Raskin, T. Harju, and J. Karhumaki. Equivalence of Labeled Markov Chains. *International Journal of Foundations of Computer Science*, 19(3):549–563, 2008.

[12] E. Fischer, F. Magniez, and M. de Rougemont. Approximate Satisfiability and Equivalence. In *LICS '06: Proceedings of the 21nd Annual IEEE Symposium on Logic in Computer Science.*, pages 421–430, 2006. To appear in SIAM Journal of Computing.

[13] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *JACM*, 45(4):653–750, 1998.

[14] LCM Kallenberg. Linear programming and finite Markovian control problems. PhD Thesis. 1983.

[15] S. Khot and A. Naor. Linear Equations Modulo 2 and the L1 Diameter of Convex Bodies. In *FOCS '07: Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science. IEEE Computer Society.*, pages 318–328, 2007.

[16] D. Krass and OJ Vrieze. Achieving Target State-Action Frequencies in Multichain Average-Reward Markov Decision Processes. *Mathematics of Operations Research*, 27(3):545, 2002.

[17] O. Madani, S. Hanks, and A. Condon. On the Undecidability of Probabilistic Planning and Infinite-Horizon Partially Observable Markov Decision Problems. In *Proceedings of the National Conference on Artificial Intelligence*, pages 541–548, 1999.

[18] S. Mannor and J.N. Tsitsiklis. On the Empirical State-Action Frequencies in Markov Decision Processes Under General Policies. *Mathematics of Operations Research*, 30(3):545, 2005.

[19] D. Ornstein and B. Weiss. How sampling reveals a process. *The Annals of Probability*, 18(3):905–930, 1990.

[20] C.H. Papadimitriou. *Computational complexity*. Addison-Wesley Reading, MA, 1994.

[21] CH Papadimitriou and JN Tsitsiklis. The complexity of Markov decision processes. *Mathematics of operations research*, 12(3):441–450, 1987.

[22] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc. New York, NY, USA, 1994.

[23] M.O. Rabin. Probabilistic Automata. *Sequential Machines: Selected Papers*, 1964.

[24] R. Rubinfeld and M. Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25:23–32, 1996.

[25] R. Segala and N. Lynch. Probabilistic simulations for probabilistic processes. *Nordic Journal of Computing*, 2(2):250–273, 1995.

[26] P. Shields. *The ergodic theory of discrete sample paths*. The American Mathematical Society, Graduate Studies in Mathematics, volume GS-13, 1996.

[27] M. Tracol. Fast convergence to state-action frequency polytopes for MDPs. *Operations Research Letters*, 37(2):123–126, 2009.

[28] W.G. Tzeng. A Polynomial-Time Algorithm for the Equivalence of Probabilistic Automata. *SIAM Journal on Computing*, 21:216, 1992.

[29] F. van Breugel and J. Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 331(1):115–142, 2005.

[30] M.Y. Vardi. Automatic verification of probabilistic concurrent finite state programs. In *FOCS '84: Proceedings of the 25th Annual Symposium on Foundations of Computer Science. IEEE Computer Society.*, pages 327–338, 1985.