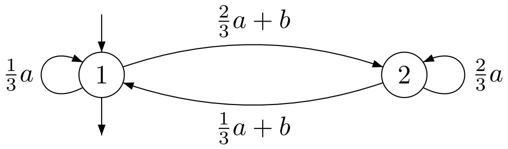
Probabilistic Aspects of Computer Science: TD9
Probabilistic Automata

Emile Contal
`http://econtal.perso.math.cnrs.fr/teaching`

November 14, 2013

Exercise 1. We consider the probabilistic automaton $\mathcal{A}$ over alphabet $A = \{a, b\}$ depicted below:



1. Describe the probability associated with word ab^n for every natural number n.
2. What are the possible probabilities that can be associated to a word by $\mathcal{A}$?
3. Determine the (finite) class $\mathcal{L} = \{L_{>\theta}(\mathcal{A}) \mid \theta \in [0, 1]\}$.

Exercise 2. Prove or disprove that the following languages are stochastic languages.

1. $\{w \in \{a, b\}^\star \mid |w|_a = |w|_b\}$, $\{w \in \{a, b, c\}^\star \mid |w|_a = |w|_b = |w|_c\}$ and $\{a^n b^n c^n \mid n > 0\}$
2. $\{a^n b^m \mid n \neq m\}$
3. $\{a^m b a^{m_1} b a^{m_2} b \cdots a^{m_k} b a^* \mid \exists j \leq k \quad m_1 + m_2 + \cdots + m_j = m\}$
4. language of palindromes over $\{a, b\}$, i.e., words w such that $\overline{w} = w$
5. $\{a^n b^m a^n b^m \mid n, m \geq 1\}$

Exercise 3 (Isolated cutpoints)**.** Let $\mathcal{A} = (Q, A, \{\mathbf{P}_a\}_{a \in A}, \pi_0, F)$ be a probabilistic automaton and let $\theta \in [0, 1]$. We say that θ is an *isolated cut point* of $\mathcal{A}$ if there is $\delta > 0$ such that for all $w \in A^*$, we have $|\mathbf{Pr}_{\mathcal{A}}(w) - \theta| \geq \delta$. In the following, we will consider that $Q = \{1, \ldots, n\}$ with 1 the unique initial state (i.e., $\pi_0(1) = 1$).

1. Let $L = L_{>\theta}(\mathcal{A})$. We consider the Myhill-Nerode congruence $\equiv_L \subseteq A^* \times A^*$ given as follows: for $u, v \in A^*$,
$$u \equiv_L v \text{ iff. } \forall w \in A^* \quad (uw \in L \iff vw \in L)$$
As we know, $\equiv_L$ has finite index (i.e., admits a finite set of equivalence classes) iff. L is regular. We also define for $u \in A^*$, $\xi^u = (\mathbf{P}_u[1, 1], \ldots, \mathbf{P}_u[1, n])$. Show that if θ is an isolated cut point,
$$u \not\equiv_L v \Longrightarrow \|\xi^u - \xi^v\|_1 \geq 4\delta$$
2. Deduce that if θ is an isolated cut point of $\mathcal{A}$, the language $L_{>\theta}(\mathcal{A})$ is regular.
3. We assume that the following variant of PCP is undecidable: Given an alphabet A and two morphisms $\varphi_1, \varphi_2 \colon A \to 1\{0, 1\}^*$, is $\{\varphi_1(w) \wedge \varphi_2(w) \mid w \in A^*\}$ finite?[1] Let φ_1 and φ_2 an instance of this problem. Show that there exists $\delta > 0$ such that for all $w \in A^+$, $|bin(0.\overline{\varphi_1(w)}) - bin(0.\overline{\varphi_2(w)})| \geq \delta$ iff. the set $\{\varphi_1(w) \wedge \varphi_2(w) \mid w \in A^*\}$ is finite[2]. Deduce that the following problem is undecidable: given a probabilistic automaton $\mathcal{A}$ and $\theta \in (0, 1)$, is θ an isolated cut point of $\mathcal{A}$?

[1] We denote $u \wedge v$ the longest common suffix of the words u and v. You may find the proof of undecidability in V. Blondel and V. Canterini: *Undecidable Problems for Probabilistic Automata of Fixed Dimension* in Theory of Computing Systems, 2001.

[2] We denote by $bin(0.w)$ the unique rational number between 0 and 1 having as binary decomposition $0.w$, where a is interpreted as 0 and b as 1.