CS 4510: Automata and Complexity **Fall 2019**

Lecture 7: Probabilistic Finite Automata

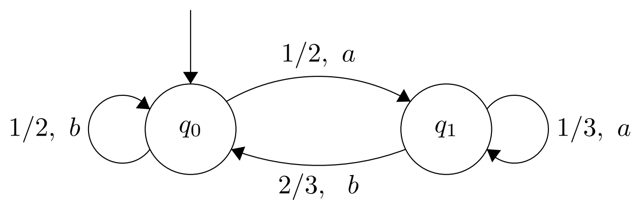
September 23, 2019

Lecturer: Santosh Vempala *Scribe: Xiaofu Niu*

We have studied deterministic finite automata (DFA) and non-deterministic finite automata (NFA). Today we are going to talk about an automata that is the generalization of finite deterministic automata. It is called **probabilistic finite automata**, or probabilistic FA. It is an automata such that in each state on an input alphabet, has a probability of going to any state.

Probabilistic FA Example

Example 1



The above figure shows state diagram for a simple probabilistic FA (PFA). This PFA has two states q_0 and q_2. Note that this PFA takes no input. On the contrary, it outputs an alphabet on each transition. Starting from q_0, this PFA has a probability of 1/2 to go to state q_1 and output letter a, and a probability of 1/2 to stay in state q_0 and output b.

This PFA has multiple possible output at each step with different probability. The list of possible outputs is listed in figure 7.1.

Output	Probability
a	$1/2$
b	$1/2$

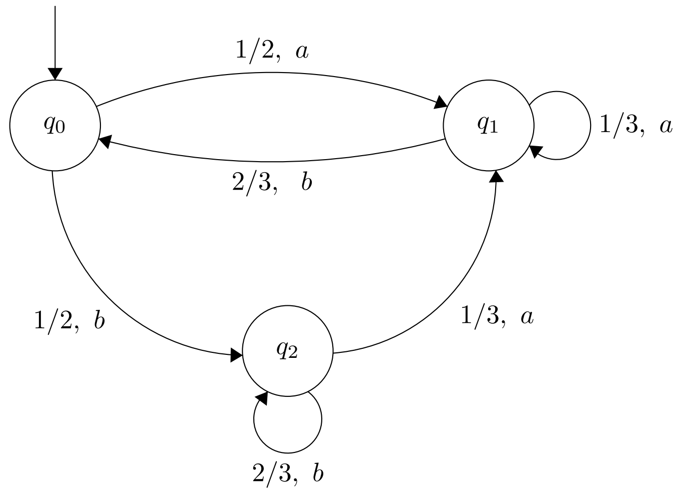
t = 1

Output	Probability
ab	$1/3$
aa	$1/6$
ba	$1/4$
bb	$1/4$

t = 2

Figure 7.1: Possible output

Example 2



The above shows state diagram for another PFA. This PFA has three state. Similar to the first example, we list all possible output and corresponding probabilities in figure 7.2.

Output	Probability
a	1/2
b	1/2

t = 1

Output	Probability
ab	1/3
aa	1/6
ba	1/6
bb	1/3

t = 2

Figure 7.2: Possible output

Transitive Matrix

Based on the previous examples, we can see that a probabilistic FA is similar to a usual automata except that in PFA each pair of transition (q_i, q_j) is assigned certain transition probabilities. We define a vector $\pi^{(t)}$ that describes the distribution of probability at the t^{th} step. The probability of being at state q_i at t^{th} step $P_\pi(q^{(t)} = q_i) = \pi_i^{(t)}(0 \leqslant \pi_i^{(t)} \leqslant 1, \sum_i \pi_i^{(t)} = 1)$. Next we define a transitive matrix. It is a matrix in which each entry $p_{i,j}$ represents the probability of going from state q_i to state q_j. Given the distribution at t^{th} step $\pi^{(t)}$, the distribution at $(t+1)^{th}$ is $\pi^{(t+1)} = P^T \cdot \pi^{(t)}$.

In the example 2, the transitive matrix is

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 2/3 & 1/3 & 0 \\ 0 & 1/3 & 2/3 \end{pmatrix}$$

Given the initial distribution $\pi^{(0)} = (1, 0, 0)^T$, the distribution at $t = 1$ is

$$\pi^{(1)} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 2/3 & 1/3 & 0 \\ 0 & 1/3 & 2/3 \end{pmatrix}^T \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1/2 \\ 1/2 \end{pmatrix}$$

Similarly the distribution at $t = 2$ is

$$\pi^{(2)} = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 2/3 & 1/3 & 0 \\ 0 & 1/3 & 2/3 \end{pmatrix}^T \cdot \begin{pmatrix} 0 \\ 1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

When we increase value of t, we will see the distribution $\pi^{(t)}$ is getting closer to $(1/4, 3/8, 3/8)^T$.

Question: Does $\pi^{(t)}$ converge i.e. $\pi^{(t+1)} = P^T \cdot \pi^{(t)} = \pi^{(t)}(t \to \infty)$?

Claim: $\pi^{(t)}$ converges and follows the rule that

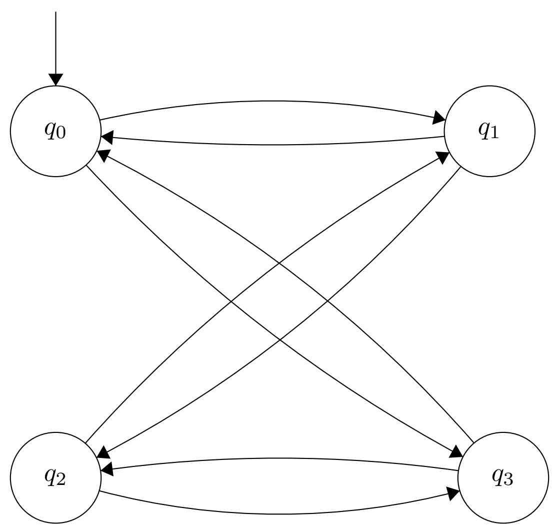
$$P^T \cdot \begin{pmatrix} a \\ b \\ b \end{pmatrix} = \begin{pmatrix} a' \\ b' \\ b' \end{pmatrix}$$

Proof:

Assume the distribution at $t = k$ is $\pi^{(k)} = (a, b, b)^T$. Given transitive matrix P we have $\pi^{(k+1)} = P^T \cdot \pi^{(t)} = (2/3b, 1/2a+2/3b, 1/2a+2/3b)^T$. To evaluate whether $\pi^{(t)}$ converges towards $u = (1/4, 3/8, 3/8)^T$, we calculate the χ^2 distance between $(u, \pi^{(t)})$ and $(u, \pi^{(t+1)})$. It turns out $\chi^2(u, \pi^{(t+1)}) = \frac{1}{9}\chi^2(u, \pi^{(t+1)})$. Clearly as t increases $\pi^{(t)}$ gets closer to u, therefore the distribution $\pi^{(t)}$ converges.

Question: Does $\pi_i^{(t)}$ always converge?

Example 3



The above figure defines another PFA whose $\pi^{(t)}$ does not converge. All transitions in this PFA have the same probability $1/2$. The distribution for this PFA does not converge since it is either $(0, 0, 1/2, 1/2)^T$ or $(1/2, 1/2, 0, 0)^T$.

Perron–Frobenius Theorem

Definition 7.1 *A non-negative matrix square T is called* ***primitive*** *if $\exists k$ such that $\forall i, j$, $(T)^k_{i,j} > 0$.*

Definition 7.2 *A non-negative matrix square T is called* ***irreducible*** *if $\forall i, j$, $\exists k$ such that $(T)^k_{i,j} > 0$.*

Theorem 7.3 ***Perron Theorem***

For a primitive matrix $\mathbf{P}$, $\mathbf{P}^k \cdot \mathbf{x} \longrightarrow \mathbf{v}$ $(x \geqslant 0, \mathbf{x} \neq 0)$ and the following conditions hold

1. $\forall i, j,\ v_{i,j} > 0$
2. $\mathbf{Pv} = \lambda \mathbf{v}$
3. $\forall \mathbf{u},\ \mathbf{Pu} = \alpha \mathbf{u},\ \alpha < \lambda$

Theorem 7.4 ***Frobenius Theorem***
For an irreducible matrix P, $\exists v$ such that the three conditions in Perron theorem hold.

Lemma 7.5 *A matrix is primitive $\iff$ its graph is strongly connected and it is aperiodic.*