Tema 2: Autómatas finitos probabilisticos

Contenido

- Repaso
- Autómata probabilísticos
- Modelos Ocultos de Markov
  - ¿Cuál es la probabilidad $x \subseteq$ L(AFP)?
  - ¿Qué secuencia de estados es más probable dada x para L(AFP)?
  - ¿Cómo definir los elementos de L(AFP) para que generalice ejemplos del cadenas?

Conceptos

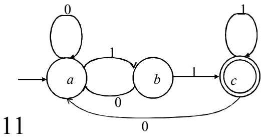
- Espacio de muestreo y ley de Probabilidad
- Eventos: simples y complejos
- Probabilidad condicional
- Probabilidad total
- Red de Bayes
- Independencia e independencia condicional

Lenguaje estocástico

- Un lenguaje estocástico es una distribución probabilística para las cadenas de Σ^*
  - P(w) donde $w \in \Sigma^*$
- $\Omega = \Sigma^*$
- AFP nos van a ayudar a definir lenguages estocásticos

AF determinísticos

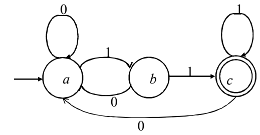
- AF determinísticos para reconocer cadenas



- 0010011
- aabaabc

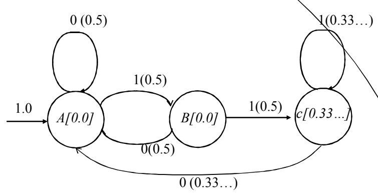
AF

- AF no determinísticos para generar cadenas



- a b a a
- 0 1 0 ??
- ¿Necesitamos una estrategía?

Generando cadenas



Generando cadenas

1. Inicializar la máquina: Escoger un estado usando la distribución I
2. Sea q el estado actual, decidir si se para la generación con F(q) o si avanzamos con *(q,a,q')* de tal forma que se genera "a" y el nuevo estado actual es q'
3. Si se decidió parar, finalizar, caso contrario Regresar a 2

Nuevas preguntas

- Agregamos a nuestra lista
  - ¿Cuál es la probabilidad $x \subseteq L(AFP)$?
  - ¿Qué secuencia de estados es más probable dada *x* para L(*AFP*)?
  - ¿Cómo definir los elementos de L(*AFP*) para que generalice ejemplos del cadenas?

Autómatas finitos probabilísticos

- AF Probabilístico No-determinístico (AFP)
- AF Probabilístico No-determinístico con transiciones Λ (AFP-Λ)
- AF Probabilístico determinístico (AFPD)

AFP

- Sexteta donde:
  - Q es un conjunto finito (de estados)
  - Σ es un alfabeto (finito) de símbolos de entrada
  - δ es un conjunto de transiciones de $Q \times \Sigma \times Q$
  - I es $Q \rightarrow R^+$, probabilidad inicial
  - P es $\delta \rightarrow R^+$, probabilidad de transición
  - F es $Q \rightarrow R^+$, probabilidad final

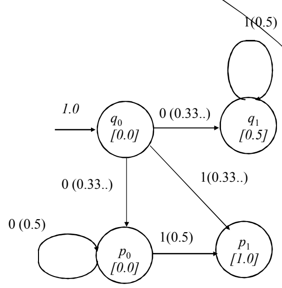
Restricciones en I,P,F

$$\boxed{\sum_{q\in Q} I(q) = 1}$$

$$\boxed{\forall q \in Q, F(q) + \sum_{a\in\Sigma, q'\in Q} P(q,a,q') = 1}$$

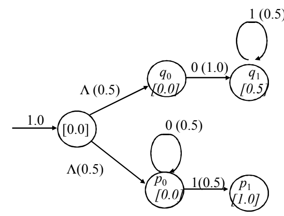
Un viejo conocido

- $L_1 \cup L_2 = \mathbf{01^* + 0^*1}$



AFP-Λ

- Una sexteta (Q, Σ, δ,I,P,F) pero con:
  - δ de esta forma $Q \times \Sigma \cup \{\Lambda\} \times Q$



AFPD

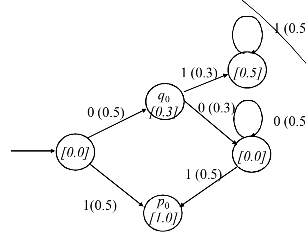
- Una sexteta (Q, Σ, δ,I,P,F) pero con:
  - para solamente un estado q_0 se cumple $I(q_0)=1$

$$I(q_0) = 1.0$$

  - para las transiciones se cumple

$$\forall q \in Q, a \in \Sigma, \left|\{q' : (q,a,q') \in \delta\}\right| \leq 1$$

Ejemplo AFPD



P(x) dado un *AFP*

- Dada una cadena x, con |x| = k, $P(x_1 x_2 \ldots x_k)$
  - Comenzar en el estado s0
  - Generar a x_0
  - llegar a s_1
  - Generar a x_1
  - Llegar a s_2
  - …
  - Llegar a s_k
  - Parar

P(x) dado un *AFP*

- Dada una cadena x, con |x| = k, $P(x_1 x_2 \ldots x_k)$
  - $I(s_0)$
  - $P(s_0, x_0, s_1)$
  - $P(s_1, x_1, s_2)$
  - …
  - $P(s_{k-1}, x_k, s_k)$
  - $F(s_k)$

P(x) dado un *AFP*

- Dada una cadena x, con |x| = k, $P(x_1 x_2 \ldots x_k)$

$$I(S_0)\left(\prod_{j=1}^{k} P(s_{j-1}, x_j, s_j)\right)F(s_k)$$

Ruta ψ/*Path*

- Secuencia de estados y cadena,
  - $s_0, x_0, s_1, x_1, \ldots x_{k-1}, x_k, s_k$
- ¿Qué pasa si se trata de un DAFP?
  - Solamente existe una ruta
- ¿Qué pasa si se trata de un AFP?
  - Potencialmente existen varias rutas
- ¿Qué pasa si se trata de un AFP-Λ?
  - La ruta puede ser más grande que |x|!

P(x) dado un *AFP*

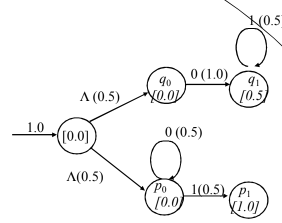
- Dada una cadena x, con |x| = k, $P(x_1 x_2 \ldots x_k)$
  - Para una ruta ψ

$$P(\psi) = I(S_0)\left(\prod_{j=1}^{k} P(s_{j-1}, x_j, s_j)\right)F(s_k)$$

  - Ruta válida, sí $P(\psi)>0$
  - Sea $\Psi(x)$ el conjunto de todas las rutas válidas de x
- Entonces:

$$P(x) = \sum_{\psi \in \Psi(x)} P(\theta)$$

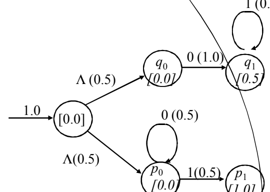
Ejemplo: P(01)



- $\psi_1 = s_0 \Lambda q_0 0 q_1 1 q_1$
- $\psi_2 = s_0 \Lambda p_0 0 p_0 1 p_1$

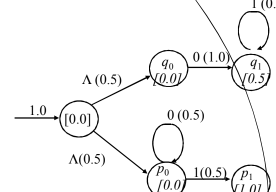
Ejemplo: P(01)

- $\psi_1 = I(s_0)P(s_0,\Lambda,q_0)P(q_0,0,q_1)P(q_1,1,q_1)F(q_1)$
- $\psi_1 = 1.0*0.5*1.0*0.5*0.5$



Ejemplo: P(01)

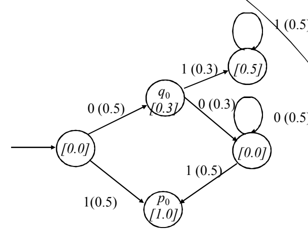
- $\psi_2 = I(s_0)P(s_0,\Lambda,p_0)P(p_0,0,p_0)P(p_0,1,p_1)F(p_1)$
- $\psi_2 = 1.0*0.5*0.5*0.5*1.0$



Ejemplo: P(01)

- θ_1=1.0*0.5*1.0*0.5*0.5=0.125
- θ_2=1.0*0.5*0.5*0.5*1.0=0.125
- P(01)=0.125+0.125=.25

Ejemplo: P(01) para AFDP



- P(01)=P(θ)=1.0*0.5*0.3*0.5=0.075

Equivalencias

- Dos PFA son equivalentes si generan la misma distribución
- ¿Son equivalentes los AFP, AFP-Λ y AFPD?
  - Existen distribuciones que pueden ser generadas por AFP pero no por AFPD
  - Para todo AFP-Λ, existe un AFP con la misma distribución

Consistencia

- ¿Realmente todas las probabilidades suman 1?

  Un PFA es consistente si todos sus estados son *útiles*

- Un estado es útil si aparece en una ruta valida

Ejemplo: HMM

- HMM son modelos AFP que factoriza la probabilidad P(δ) en dos partes
  - Probabilidad de transición: A:QxQ
  - Probabilidad de emisión: B:QxΣ
- Además, no hay estados finales

¿Por qué factorizar?

- Suponga |Q|=10 y |Σ|=10
- P(QxQx Σ)=10*10*10=1000
- A(QxQ)=10*10=100
- B(Qx Σ)=10*10=100

Equivalencia

- $P(q_0,a,q_1)=A(q_1|q_0)B(a|q_0)$

Dr. Ivan Meza, IIMAS, UNAM, 2010