

ASRSV – curs 10 - cont.

Metode si tehnici de recunoastere HMM

*L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proc. of the IEEE, Vol.77, No.2, pp.257--286, 1989.

Available from

<http://ieeexplore.ieee.org/iel5/5/698/00018626.pdf?arnumber=18626>

Elementele unui MMA: $\lambda = (A, B, \Pi)$; se definesc următoarele elemente :

N - numărul de stări reprezentate prin $S = \{s_1, s_2, \dots, s_N\}$, o stare la un moment dat este cotate cu q_t ($q_t \in S$).

M - numărul de simboluri observabile în fiecare stare. Multimea de observatii este $V = \{v_1, v_2, \dots, v_M\}$

Un element O_t din V desemneaza un simbol observat la momentul t .

A - o matrice de probabilități de tranziție între stările lantului : a_{ij} - reprezentând probabilitatea cu care modelul evoluează din starea i spre j , în cazul general :

$$a_{ij} = A(i, j) = P(q_{t+1} = s_j \mid q_t = s_i) \quad \forall i, j \in [1, N], \quad \forall t \in [1, T], \quad a_{ij} \geq 0 \text{ si } \forall i, j, \sum a_{ij} = 1.$$

B – o matrice de probabilități de observare a simbolurilor în fiecare stare a modelului : $b_j(k)$ care reprezintă probabilitatea ca să observăm simbolul v_k când modelul se află în starea j :

$$b_j(k) = P(O_t = v_k \mid q_t = s_j), \quad 1 \leq j \leq N \quad 1 \leq k \leq M$$
$$b_j(k) \geq 0, \quad \forall j, k \text{ si } \sum b_j(k) = 1$$

P - o multime a densității de probabilitate initiale : $\Pi = \{\pi_i\}$, $i = \overline{1, \dots, N}$, unde π_i reprezintă probabilitatea ca starea de pornire a modelului să fie i , astfel :

$$\pi_i = P(q_1 = s_i) \quad i = 1, \dots, N \text{ si } \pi_i \geq 0, \quad \sum \pi_i = 1.$$

Problemele de rezolvat la MMA

Sunt trei probleme de bază a unui MMA care trebuie solutionate :

1. Evaluarea probabilităților de observare

Fiind data succesiunea de observatii: $O = O_1, O_2, \dots, O_T$ si modelul $\lambda = (A, B, \Pi)$, se pune problema determinării probabilităților $P(O | \lambda)$, ca succesiunea observatiilor să fie generată de model.

-Solutia - Algoritmul Forward-backward

2. Problema descoperirii secvenței de stări ascunse

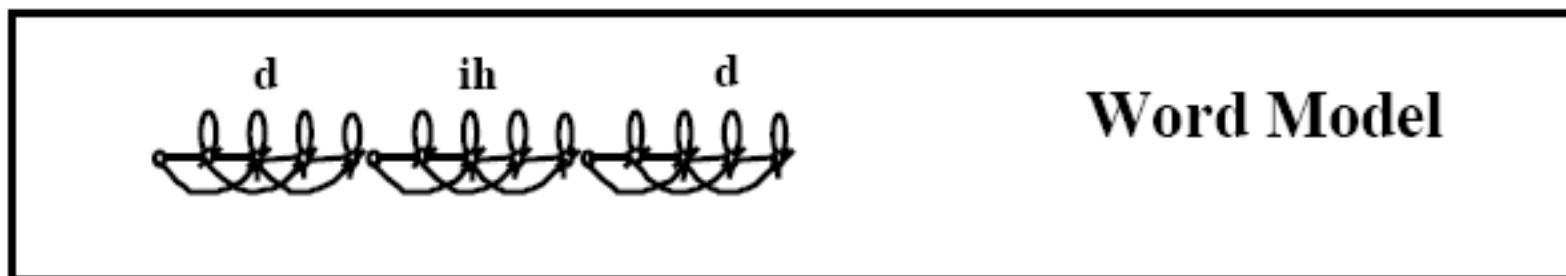
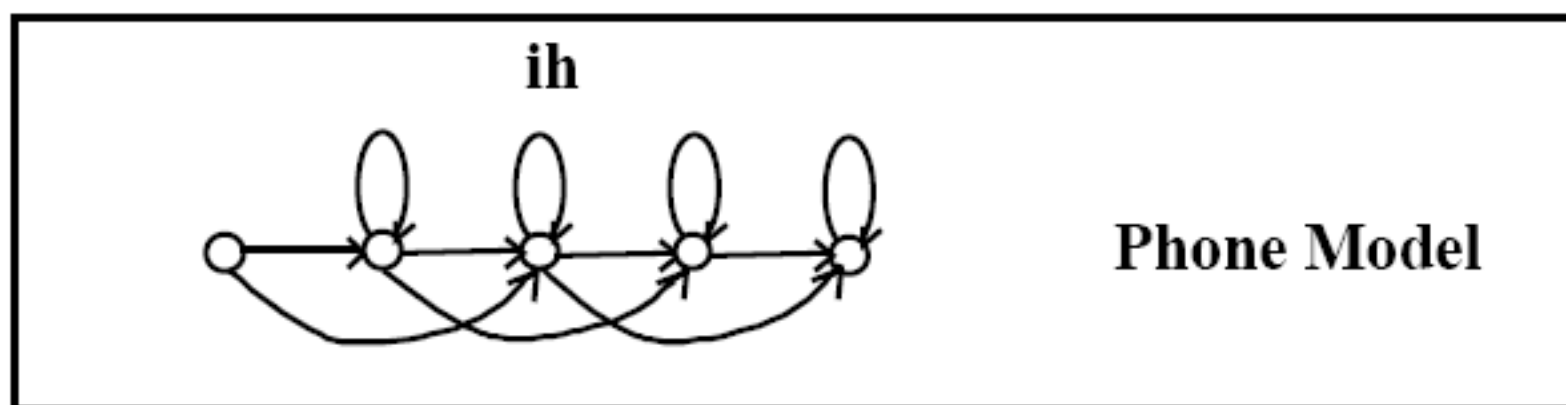
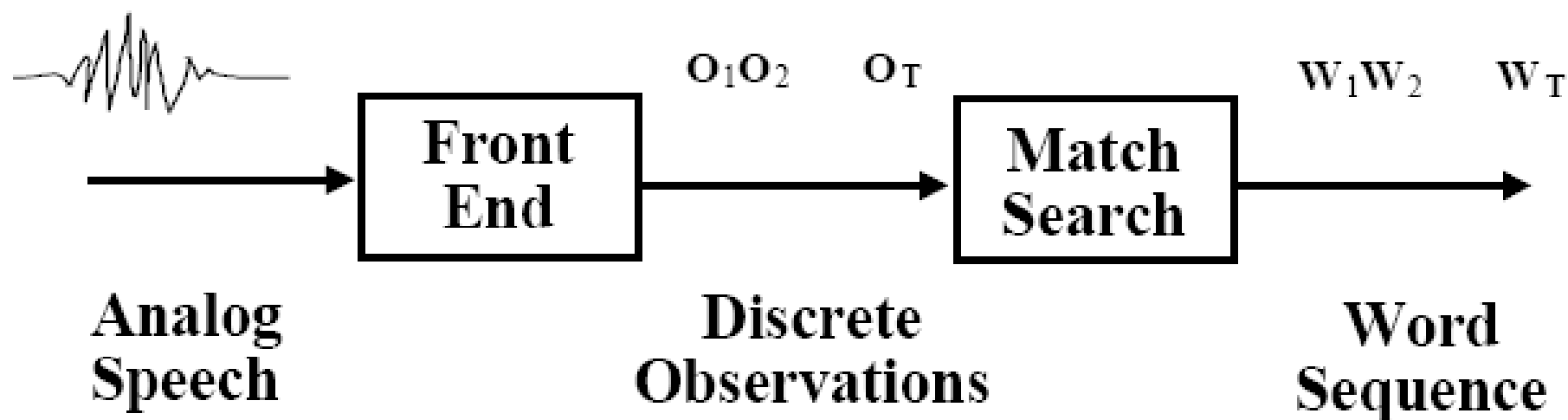
- Fiind dată succesiunea observatiilor $O = O_1, O_2, \dots, O_T$ si modelul λ : să se determine secvența de stări, $Q = q_1, q_2, \dots, q_T$, cea mai probabilă din care au rezultat observatiile O_t .
- Se obtin informatii legate de partea ascunsă a modelului (succesiunea de stări).
- Se foloseste un criteriu optimal pentru descoperirea secvenței de stări

-Solutia – Algoritmul Viterbi

3. Problema antrenării

- Cum trebuie modificati parametrii modelului $\lambda = (A, B, \Pi)$ pentru a maximiza $P(O | \lambda)$ - reprezintă problema antrenării modelului.

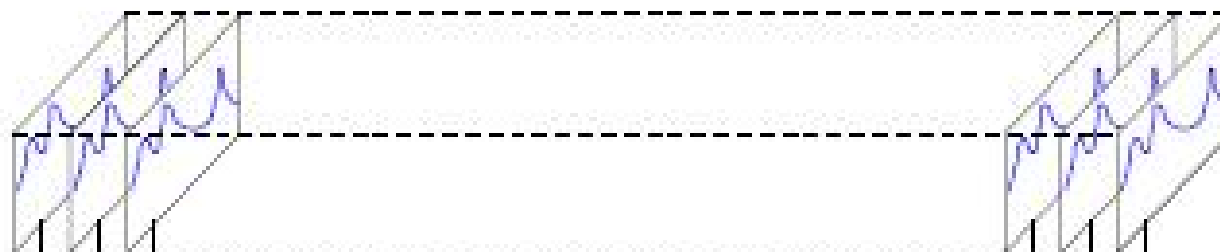
-Solutia – Algoritmul Baum-Welch



Speech signal



Spectral analysis

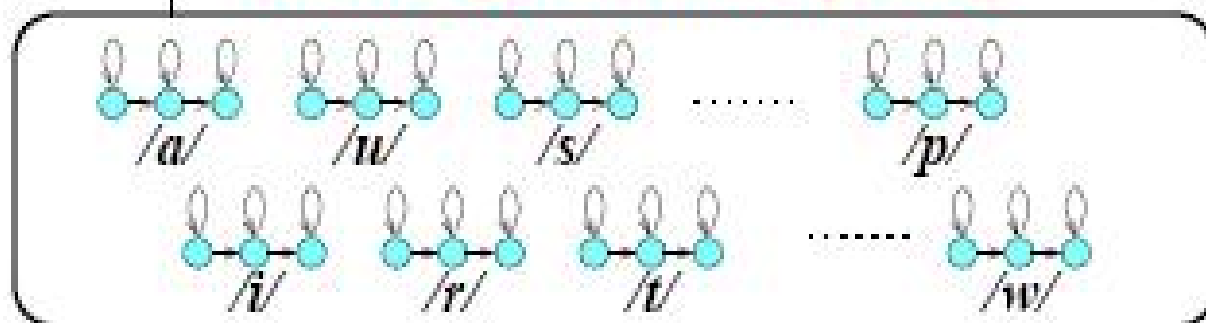


Feature vector
sequence



$$p(X|\text{sayonara}) = p(X_1|s) p(X_2|a) p(X_3|y) p(X_4|o) p(X_5|n) p(X_6|a) p(X_7|r) p(X_8|a)$$

Acoustic (phone) model [HMM]



1. Evaluarea probabilității de observare (totale)

Există mai multe maniere de a evalua probabilitatea de observare, din care se prezintă două mai importante :

a) Evaluarea directă Probabilitatea ca succesiunea de observații O , să fie dată de modelul λ este egală cu suma probabilității tuturor cailor de stări posibile Q , a probabilităților conjuncte a lui O și Q :

$$P(O|\lambda) = \sum_Q P(O, Q|\lambda) = \sum_Q P(O|Q, \lambda) \cdot P(Q|\lambda)$$

sau $P(Q|\lambda) = \prod (q_1 a_{q_1 q_2} a_{q_2 q_3} a_{q_2 q_3} \dots a_{q_{T-1} q_T})$ și

$$P(O|Q, \lambda) = b_{q_1}(O_1) \cdot b_{q_2}(O_2) \cdot \dots \cdot b_{q_T}(O_T), \quad \text{de unde :}$$

$$P(O|\lambda) = \sum_{q_1, q_2, \dots, q_T} \prod q_1 \cdot b_{q_1}(O_1) \cdot b_{q_2}(O_2) \cdot \dots \cdot b_{q_T}(O_T)$$

În această abordare este nevoie de $NT-1$ adunări și $(2T-1)$ înmulțiri (N numărul de cai posibile de lungime T), adică $2TN$ calcule care este un volum mare (de ex. $N=5$, $T=100$ avem $\approx 10^{72}$).

Pentru această problemă mai este o variantă pentru evaluare numită forward-backward (înainte - înapoi).

b) Evaluarea prin functiile forward-backward

In acest caz putem considera că observatia are două etape :

- mai întâi observatiile de inceput $O(1:t)$ ajungând până la starea q_i la momentul t , apoi
- emisia sfârșitului de observare $O(t+1:T)$, știind că se pleacă din starea q_i în momentul t . In acest caz evaluarea observatiei este egală cu :

$$P(O|\lambda) = \sum_{q_i} \alpha(t, q_i) \cdot \beta(t, q_i)$$

unde $\alpha(t, q_i)$ este probabilitatea de a emite inceputul $O(1:t)$ si de a ajunge la q_i in momentul t si $\beta(t, q_i)$ este probabilitatea de a emite secventa finală $O(t+1:T)$ știind ca se pleaca din q_i la momentul t . Calculul lui α se face cu t crescator, iar a lui β cu t descrescator , de unde si numele algoritmului.

Calculul lui α .

Se notează secvența $\alpha(t, q_i)$ prin $\alpha_t(i)$:

$$\alpha_t(i) = P(o_1 o_2 \dots o_t, q_t = s_i | \lambda)$$

care se calculează recursiv după cum urmează :

- inițializare

$$\alpha_1(i) = \pi_i b_i(o_1) \quad \text{unde } i=1, \dots, N$$

- inducție (recurența)

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(o_{t+1}), \quad 1 \leq t \leq T-1 \quad \text{și} \quad 1 \leq j \leq N$$

- terminare

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i)$$

Aceasta indică faptul că pentru a emite începutul observațiilor $O(1:t+1)$ și a ajunge în starea s_j în momentul $t+1$, trebuie neapărat să fim în una din stările s_i la momentul t . Volumul de calcul cerut este $N(N+1)(T-1)+N$ înmulțiri și $N(N-1)(T-1)$ adunări (de ex. $N=5$ și $T=1000$ avem ≈ 3000 calcule)⁸

Calculul lui β .

Notam $\beta(t, qi) = \beta_t(i)$.

$$\beta_t(i) = P(O_{t+1} O_{t+2} \dots O_T | q_T = si | \lambda)$$

Se deduce β_t din β_{t+1} in mod inductiv astfel :

- initializare

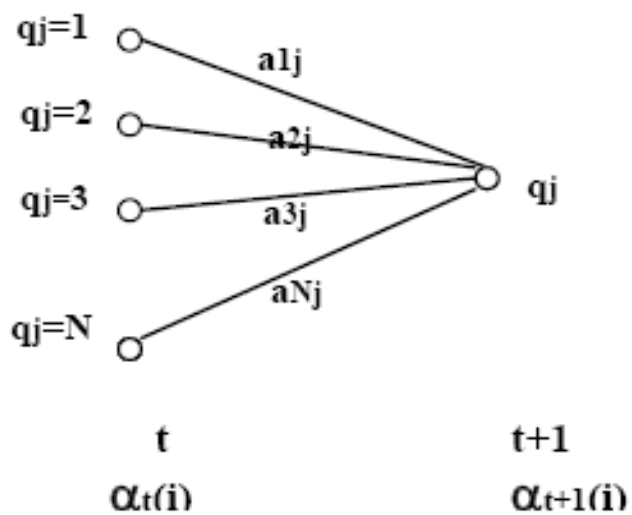
$$\beta_T(i) = 1, 1 \leq i \leq N$$

- inductie

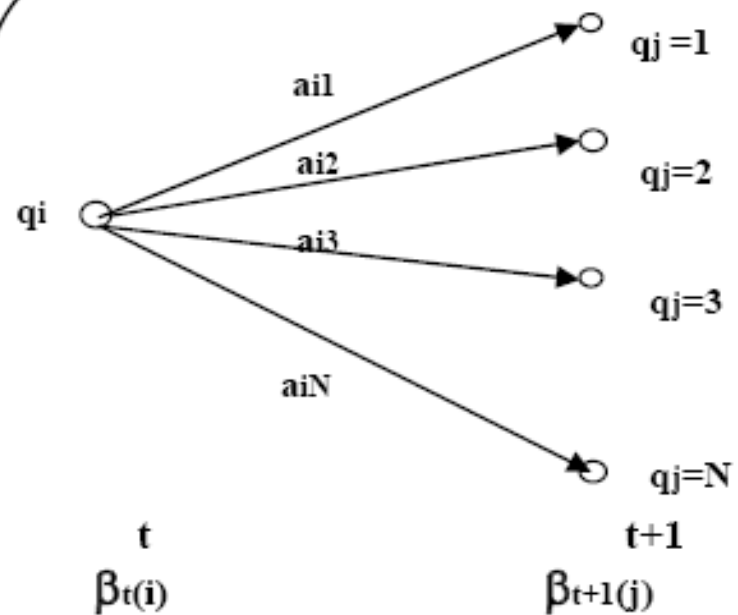
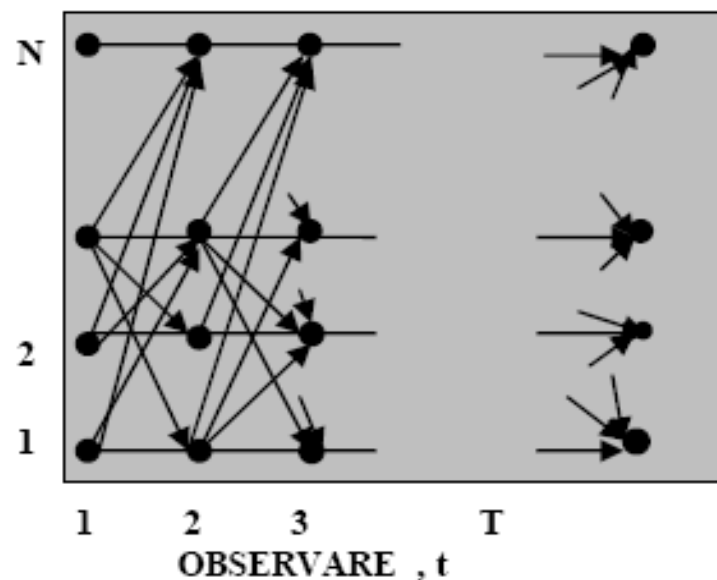
$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(O_{t+1}) \beta_{t+1}(j), T-1 \geq t \geq 1 \text{ si } 1 \leq i \leq N$$

- terminare - calculul probabilității de observare

$$P(O | \lambda) = \sum_{i=1}^N \Pi_i b_i(O_{t+1}) \beta_1(i)$$



STARE



Secventa de operatii
ceruta de calculul variabilei
backward $\beta_t(i)$.

Secventa de operatii ceruta de calculul
variabilei forward $\alpha_{t+1}(j)$.

Pentru calculul lui β se fac N^2T calcule. Totusi adesea se folosesc valorile obtinute pentru 2 cazuri particulare : $t=0$ si $t=T$ ce dau :

$$P(O|\lambda) = \sum_{q_i} \alpha_T(i) = \sum_{q_i} \pi_i \beta_0(i).$$

2. Calculul secvenței optime

O sarcină importantă a modulelor ce folosesc MMA este secvența corespunzătoare observației, astfel spus de a găsi în model cea mai bună secvență de stări care să maximizeze mărimea următoare, numită secvență de stări Viterbi :

$$P(Q | O, \lambda) = P(Q, O | \lambda)$$

Algoritmul VITERBI

Pentru a găsi în model cel mai bun drum $Q=(q_1, q_2, \dots, q_T)$ pentru succesiunea de observații $O=(O_1, O_2, \dots, O_T)$ se definește $\delta_t(i)$ care este probabilitatea celui mai bun drum care duce la starea s_i în momentul t trecând prin primele t observații:

$$\delta_t(i) = \max_{q_1, q_2, \dots, q_{t-1}} P(q_1, q_2, \dots, q_{t-1}, q_t = s_i, O_1, O_2, \dots, O_t | \lambda)$$

Prin inducție , se calculează

$$\delta_{t+1}(j) = [\max_i \delta_t(i) a_{ij}] b_j(O_{t+1})$$

pastrând în calcule succesiunea de stări care da cel mai bun drum ducând la starea s_i într-un tablou Φ .

Algoritmul este urmatorul :

- initializare

$$\delta_1(i) = \pi_i \cdot b_i(O_1) \quad , \quad i=1 \dots N$$
$$\Psi_1(i) = 0$$

- recurenta

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(O_t) \quad , \quad 1 \leq j \leq N \quad \text{si } t=2 \dots N$$

- terminare

$$P^* = \max[\delta_T(i)] \qquad q^*_T = \operatorname{argmax}_{1 \leq i \leq N} [\delta_T(i)]$$

- determinarea secventei de stari (backtracking)

$$q^*_t = \Psi_{t+1}(q^*_{t+1}) \quad , T-1 \leq t \leq 1$$

unde functia argmax permite memorarea indicelui $i=1, \dots, N$, cu care se atinge maximul valorilor $\delta_{t-1}(i) a_{ij}$. Volumul de calcul este $O(N \cdot N \cdot T)$ operatii. ¹³

3. Algoritmul pentru antrenare

Problema antrenării modelului este cea mai dificilă din cele trei, deoarece nu este o metoda analitică pentru determinarea modelului. În acest caz se pune problema ajustării parametrilor modelului (A, B, Π) pentru a maximiza probabilitatea secvențelor de observare date de model $P(O|\lambda)$.

Se folosește o metoda iterativă ca de ex. metoda BAUM-WELCH sau tehnicile gradientului pentru optimizare.

Ideea antrenării este utilizarea unor proceduri de reestimare care să ajusteze modelul puțin câte puțin trecând prin etapele:

- alegerea unui set inițial de parametri λ_0 ;
- calculul lui λ_1 din λ_0
- repetarea procesului până ajungem la un criteriu de terminare

Plecând de la λ_n , λ_{n+1} trebuie să satisfacă :

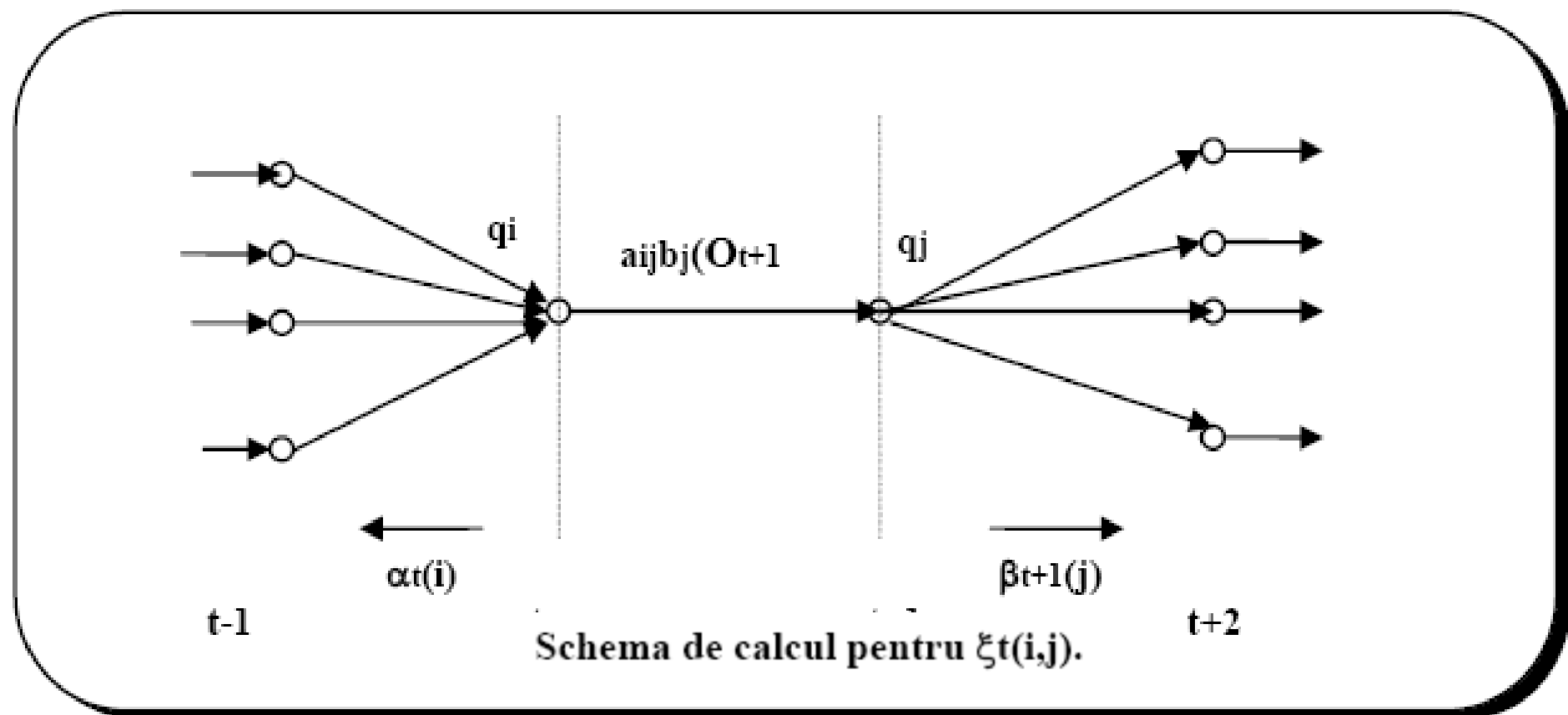
$$\prod_r P(O^r | \lambda_{n+1}) \geq \prod_r (P(O^r | \lambda_n))$$

ceea ce arată că λ_{n+1} trebuie să îmbunătățească probabilitatea de emisie a observațiilor, ceea ce revine la definirea unei funcții F care să :

$$\lambda_{n+1} = F(\lambda_n)$$

Sunt mai multe modalitati de a defini functia F , cea mai simpla este de a face statistica cu folosirea tranzitiilor si distributiilor:

- revenind a calcula frecventa de utilizare, plecând de la multimea de antrenare. Daca multimea este însemnata, statistica furnizeaza o buna aproximatie a probabilitatilor a posteriori folosite ca parametri ai modelului pentru iteratia urmatoare.



Formulele pentru reestimare.

Se definește :

$$\xi_t(i,j)=P(q_t=s_i, q_{t+1}=s_j|O,\lambda)$$

și se deduce :

$$\xi_t(i,j)=\frac{P(q_t = s_i, q_{t+1} = s_j, O|\lambda)}{P(O|\lambda)}$$

Dezvoltând numărătorul și introducând α, β se ajunge la :

$$\xi_t(i,j)=\frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{P(O|\lambda)}$$

Definind : $\gamma_t(i)=P(q_t=s_i|O,\lambda)$, rezulta :

$$\gamma_t(i) = \sum_{j=1}^N P(q_t = s_i, q_{t+1} = s_j|O,\lambda) = \sum_{j=1}^N \xi_t(i,j)$$

, unde $\sum_i \gamma_t(i)$ da informatii despre numarul tranzitiilor posibile prin starea s_i sau $\sum_t \xi_t(i,j)$ da informatii despre numarul tranzitiilor posibile plecând de la starea s_i și la starea s_j .

Formele de reestimare posibile pentru Π, A, B pot fi :

$\overline{\pi_i}$ - care reprezintă de câte ori este posibil de a fi în starea i și în momentul 1

$\overline{a_{ij}}$ - numărul de tranziții posibile de la i la j / numărul de tranziții plecând de la i

$\overline{b_j(k)}$ - numărul de posibilități de a fi în j observând v_k / numărul de posibilități de a fi în j

Fie :

$$\overline{\pi} = \gamma_1(i)$$

$$\overline{a_{ij}} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i, j)}, \quad \overline{b_j(k)} = \frac{\sum_{t=1, O_t=v_k}^T \gamma_t(i)}{\sum_{t=1}^T \gamma_t}$$

Constrângerile impuse spre respectare :

$$\sum_{i=1}^N \overline{\pi_i} = 1 \quad \text{și} \quad \sum_{j=1}^N \overline{b_j} = 1$$

Aceste relații rezultă din teorema Baum-Welch

Algoritmul de reestimare Baum-Welch.

1. Fixăm valorile initiale:
 $a_{ij}^0, b_j^0(k), \pi_i^0 \quad 1 \leq i, j \leq N, \quad 1 \leq k \leq N$
2. Se calculeaza cu ajutorul functiilor forward-backward :
 $\xi(i, j), \gamma_t(i) \quad 1 \leq i, j \leq N, \quad 1 \leq t \leq T-1$ si $\bar{\lambda}$ folosind formulele de reestimare.
3. Se repeta 2 până se atinge o limita impusa.

Observatii.

- Alegerea modelului initial afecteaza rezultatul.

Toate valorile din A si B egale cu 0 la început ramân 0 si la sfârșitul antrenarii

- Algoritmul converge spre valori ale parametrilor care formeaza un punct critic $P(O|\lambda)$. Acest punct poate corespunde la un maxim local sau la un punct de inflexiune, de unde si buna alegere a modelului.

- Numarul de iteratii se fixeaza empiric

- Se poate aplica metoda la orice tip de semnal . Ajunge sa avem primitive si a se da o ordine a observatiilor (secventiere), ceea ce este reprezentat de $1 \leq t \leq T$.

EXEMPLU

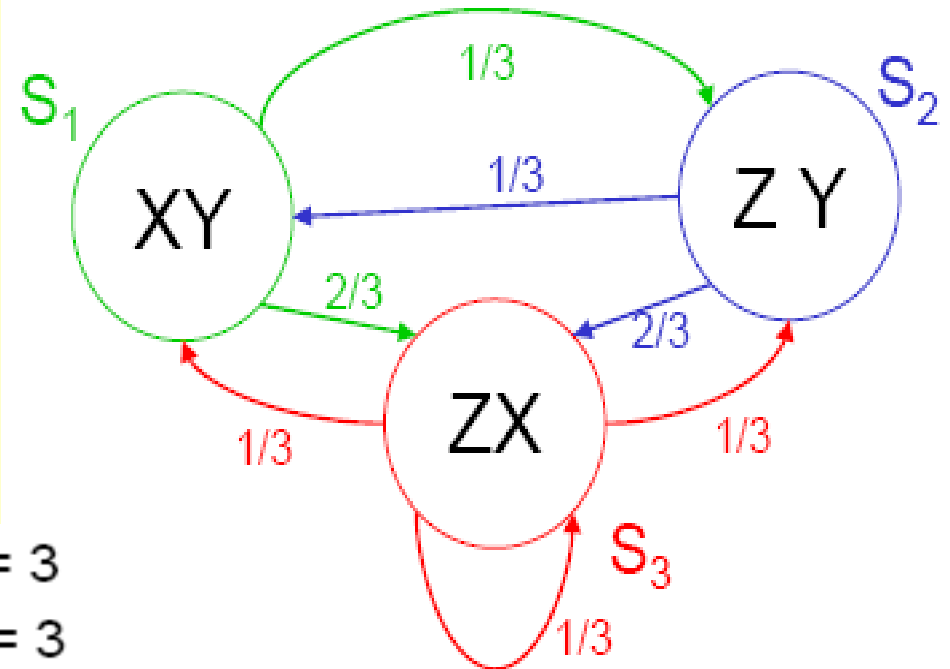
Se porneste aleator din S1 sau S2

Se alege in fiecare stare unul dintre simbolurile de iesire, aleator

Generam secventa de observatii:

Probabilitatea initiala este 50% ptr.

S1 sau S2



$$N = 3$$

$$M = 3$$

$$\pi_1 = 1/2$$

$$\pi_2 = 1/2$$

$$\pi_3 = 0$$

$$a_{11} = 0$$

$$a_{12} = 1/3$$

$$a_{13} = 2/3$$

$$a_{12} = 1/3$$

$$a_{22} = 0$$

$$a_{13} = 2/3$$

$$a_{13} = 1/3$$

$$a_{32} = 1/3$$

$$a_{13} = 1/3$$

$$b_1(X) = 1/2$$

$$b_1(Y) = 1/2$$

$$b_1(Z) = 0$$

$$b_2(X) = 0$$

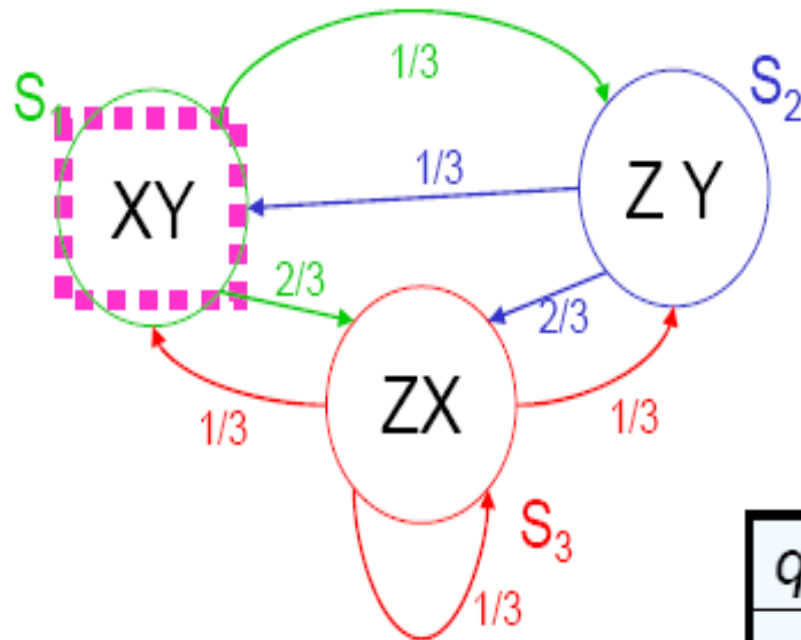
$$b_2(Y) = 1/2$$

$$b_2(Z) = 1/2$$

$$b_3(X) = 1/2$$

$$b_3(Y) = 0$$

$$b_3(Z) = 1/2$$



50% probabilitatea de a observa in S1, X sau Y

$$N = 3$$

$$M = 3$$

$$\pi_1 = \frac{1}{2}$$

$$\pi_2 = \frac{1}{2}$$

$$\pi_3 = 0$$

$$a_{11} = 0$$

$$a_{12} = \frac{1}{3}$$

$$a_{13} = \frac{2}{3}$$

$$a_{12} = \frac{1}{3}$$

$$a_{22} = 0$$

$$a_{13} = \frac{2}{3}$$

$$a_{13} = \frac{1}{3}$$

$$a_{32} = \frac{1}{3}$$

$$a_{13} = \frac{1}{3}$$

$$b_1(X) = \frac{1}{2}$$

$$b_1(Y) = \frac{1}{2}$$

$$b_1(Z) = 0$$

$$b_2(X) = 0$$

$$b_2(Y) = \frac{1}{2}$$

$$b_2(Z) = \frac{1}{2}$$

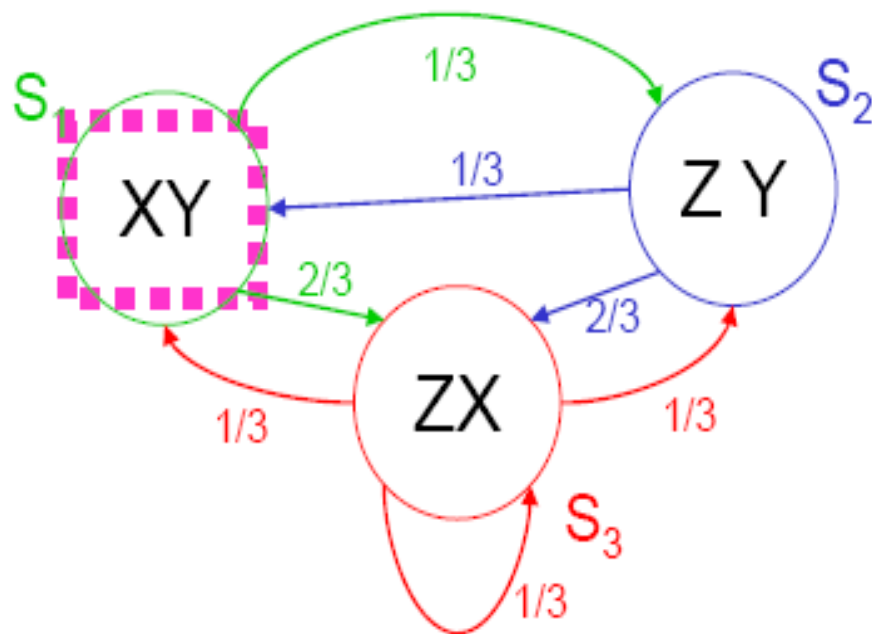
$$b_3(X) = \frac{1}{2}$$

$$b_3(Y) = 0$$

$$b_3(Z) = \frac{1}{2}$$

$q_0 =$	S_1	$O_0 =$	
$q_1 =$	—	$O_1 =$	—
$q_2 =$	—	$O_2 =$	—

$q_0 =$	S_1	$O_0 =$	X
$q_1 =$		$O_1 =$	—
$q_2 =$	—	$O_2 =$	—



$$N = 3$$

$$M = 3$$

$$\pi_1 = \frac{1}{2}$$

$$\pi_2 = \frac{1}{2}$$

$$\pi_3 = 0$$

$$a_{11} = 0$$

$$a_{12} = \frac{1}{3}$$

$$a_{13} = \frac{2}{3}$$

$$a_{12} = \frac{1}{3}$$

$$a_{22} = 0$$

$$a_{13} = \frac{2}{3}$$

$$a_{13} = \frac{1}{3}$$

$$a_{32} = \frac{1}{3}$$

$$a_{13} = \frac{1}{3}$$

$$b_1(X) = \frac{1}{2}$$

$$b_1(Y) = \frac{1}{2}$$

$$b_1(Z) = 0$$

$$b_2(X) = 0$$

$$b_2(Y) = \frac{1}{2}$$

$$b_2(Z) = \frac{1}{2}$$

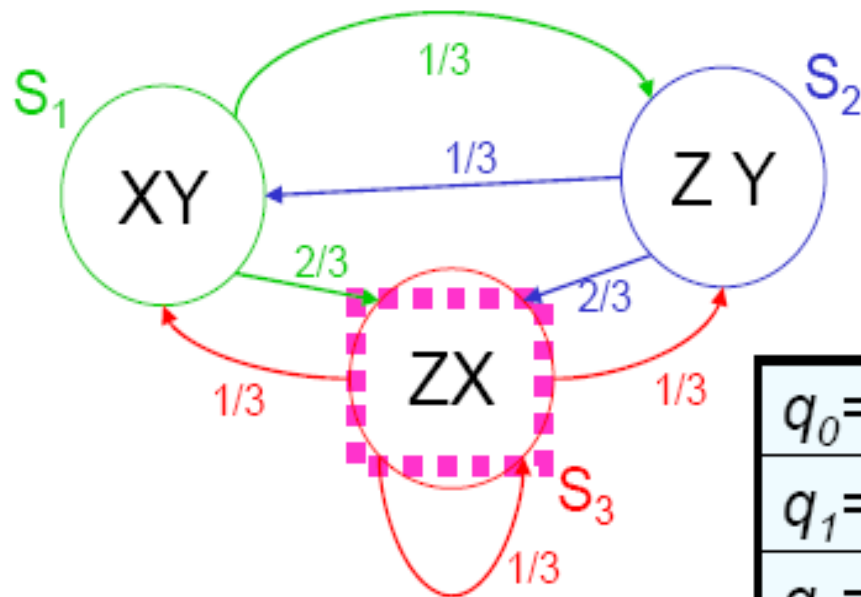
$$b_3(X) = \frac{1}{2}$$

$$b_3(Y) = 0$$

$$b_3(Z) = \frac{1}{2}$$

Trec in S3 cu probabilitatea 2/3 sau in S2 cu probabilitatea 1/3

$q_0 =$		$O_0 =$	X
$q_1 =$		$O_1 =$	—
$q_2 =$	—	$O_2 =$	—



$$N = 3$$

$$M = 3$$

$$\pi_1 = \frac{1}{2}$$

$$\pi_2 = \frac{1}{2}$$

$$\pi_3 = 0$$

$$a_{11} = 0$$

$$a_{12} = \frac{1}{3}$$

$$a_{13} = \frac{1}{3}$$

$$a_{12} = \frac{1}{3}$$

$$a_{22} = 0$$

$$a_{32} = \frac{1}{3}$$

$$a_{13} = \frac{2}{3}$$

$$a_{13} = \frac{2}{3}$$

$$a_{13} = \frac{1}{3}$$

$$b_1(X) = \frac{1}{2}$$

$$b_2(X) = 0$$

$$b_3(X) = \frac{1}{2}$$

$$b_1(Y) = \frac{1}{2}$$

$$b_2(Y) = \frac{1}{2}$$

$$b_3(Y) = 0$$

$$b_1(Z) = 0$$

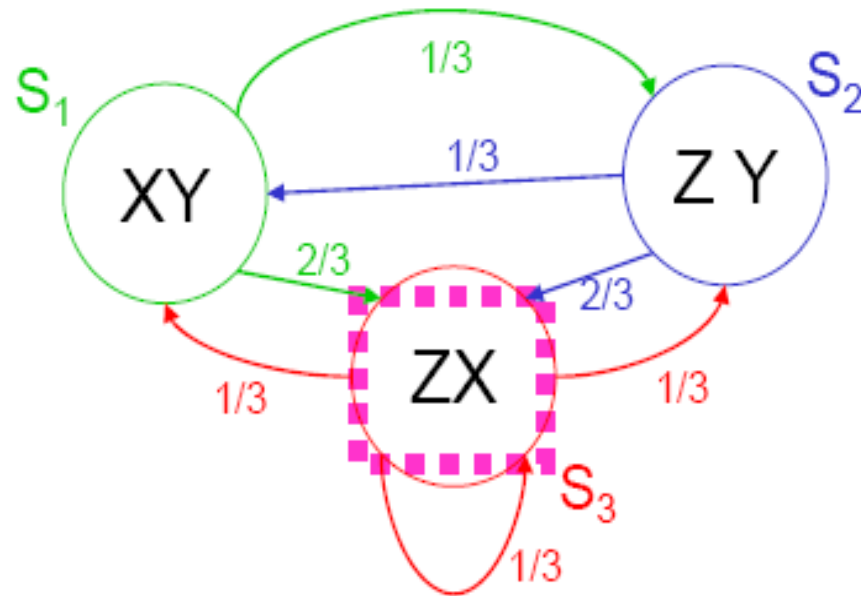
$$b_2(Z) = \frac{1}{2}$$

$$b_3(Z) = \frac{1}{2}$$

$q_0 =$	S_1	$O_0 =$	X
$q_1 =$	S_3	$O_1 =$	—
$q_2 =$	—	$O_2 =$	—

50% probabilitatea de a observa in S3, X sau Z, fie X

$q_0 =$	S_1	$O_0 =$	X
$q_1 =$	S_3	$O_1 =$	X
$q_2 =$	—	$O_2 =$	—



$$N = 3$$

$$M = 3$$

$$\pi_1 = \frac{1}{2}$$

$$\pi_2 = \frac{1}{2}$$

$$\pi_3 = 0$$

$$a_{11} = 0$$

$$a_{12} = \frac{1}{3}$$

$$a_{13} = \frac{2}{3}$$

$$a_{12} = \frac{1}{3}$$

$$a_{22} = 0$$

$$a_{13} = \frac{2}{3}$$

$$a_{13} = \frac{1}{3}$$

$$a_{32} = \frac{1}{3}$$

$$a_{13} = \frac{1}{3}$$

$$b_1(X) = \frac{1}{2}$$

$$b_1(Y) = \frac{1}{2}$$

$$b_1(Z) = 0$$

$$b_2(X) = 0$$

$$b_2(Y) = \frac{1}{2}$$

$$b_2(Z) = \frac{1}{2}$$

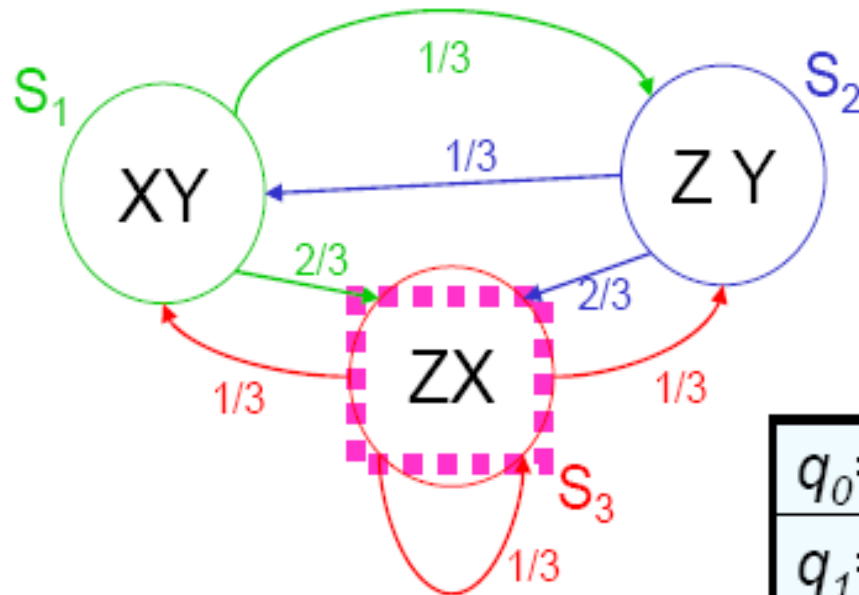
$$b_3(X) = \frac{1}{2}$$

$$b_3(Y) = 0$$

$$b_3(Z) = \frac{1}{2}$$

Trecerea in oricare din cele 3 stari este echiprobabila 1/3, fie S3

$q_0 =$	S_1	$O_0 =$	X
$q_1 =$	S_3	$O_1 =$	X
$q_2 =$	S_3	$O_2 =$	—



$$N = 3$$

$$M = 3$$

$$\pi_1 = \frac{1}{2}$$

$$\pi_2 = \frac{1}{2}$$

$$\pi_3 = 0$$

$$a_{11} = 0$$

$$a_{12} = \frac{1}{3}$$

$$a_{13} = \frac{2}{3}$$

$$a_{12} = \frac{1}{3}$$

$$a_{22} = 0$$

$$a_{13} = \frac{2}{3}$$

$$a_{13} = \frac{1}{3}$$

$$a_{32} = \frac{1}{3}$$

$$a_{13} = \frac{1}{3}$$

$$b_1(X) = \frac{1}{2}$$

$$b_1(Y) = \frac{1}{2}$$

$$b_1(Z) = 0$$

$$b_2(X) = 0$$

$$b_2(Y) = \frac{1}{2}$$

$$b_2(Z) = \frac{1}{2}$$

$$b_3(X) = \frac{1}{2}$$

$$b_3(Y) = 0$$

$$b_3(Z) = \frac{1}{2}$$

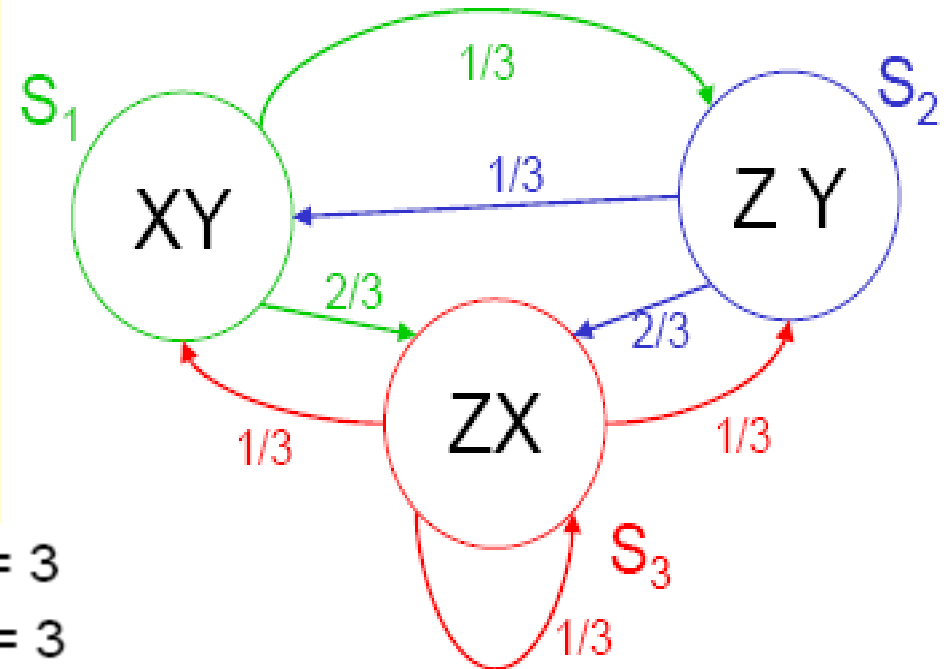
$q_0 =$	S_1	$O_0 =$	X
$q_1 =$	S_3	$O_1 =$	X
$q_2 =$	S_3	$O_2 =$	—

50% probabilitatea de a observa in S3, X sau Z, fie Z

$q_0 =$	S_1	$O_0 =$	X
$q_1 =$	S_3	$O_1 =$	X
$q_2 =$	S_3	$O_2 =$	Z

ESTIMAREA STARILOR

$q_0 =$	?	$O_0 =$	X
$q_1 =$	?	$O_1 =$	X
$q_2 =$	?	$O_2 =$	Z



$$N = 3$$

$$M = 3$$

$$\pi_1 = 1/2$$

$$\pi_2 = 1/2$$

$$\pi_3 = 0$$

$$a_{11} = 0$$

$$a_{12} = 1/3$$

$$a_{13} = 2/3$$

$$a_{12} = 1/3$$

$$a_{22} = 0$$

$$a_{13} = 2/3$$

$$a_{13} = 1/3$$

$$a_{32} = 1/3$$

$$a_{13} = 1/3$$

$$b_1(X) = 1/2$$

$$b_1(Y) = 1/2$$

$$b_1(Z) = 0$$

$$b_2(X) = 0$$

$$b_2(Y) = 1/2$$

$$b_2(Z) = 1/2$$

$$b_3(X) = 1/2$$

$$b_3(Y) = 0$$

$$b_3(Z) = 1/2$$

PROBABILITATEA SERIILOR DE OBSERVATII

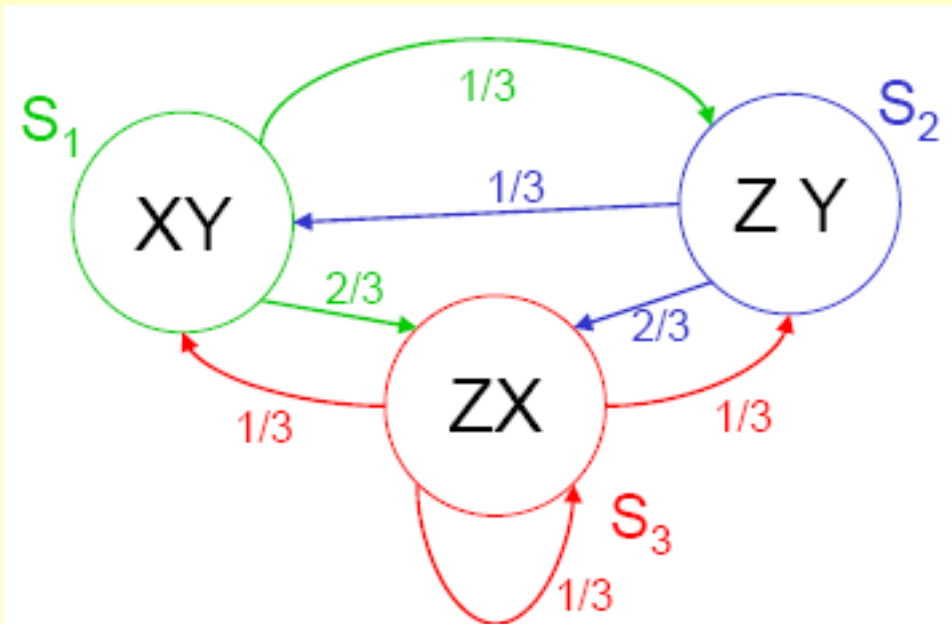
Ce reprezinta:

$$P(\mathbf{O}) = P(O_1 O_2 O_3) = P(O_1 = X \wedge O_2 = X \wedge O_3 = Z)?$$

$$P(\mathbf{O}) = \sum_{Q \in \text{Paths of length 3}} P(\mathbf{O} \wedge \mathbf{Q}) = \sum_{Q \in \text{Paths of length 3}} P(\mathbf{O} | \mathbf{Q}) P(\mathbf{Q})$$

Cum sa calculam $P(Q)$ pentru o cale arbitrara Q ?

Cum sa calculam $P(O|Q)$ pentru o cale arbitrara Q ?

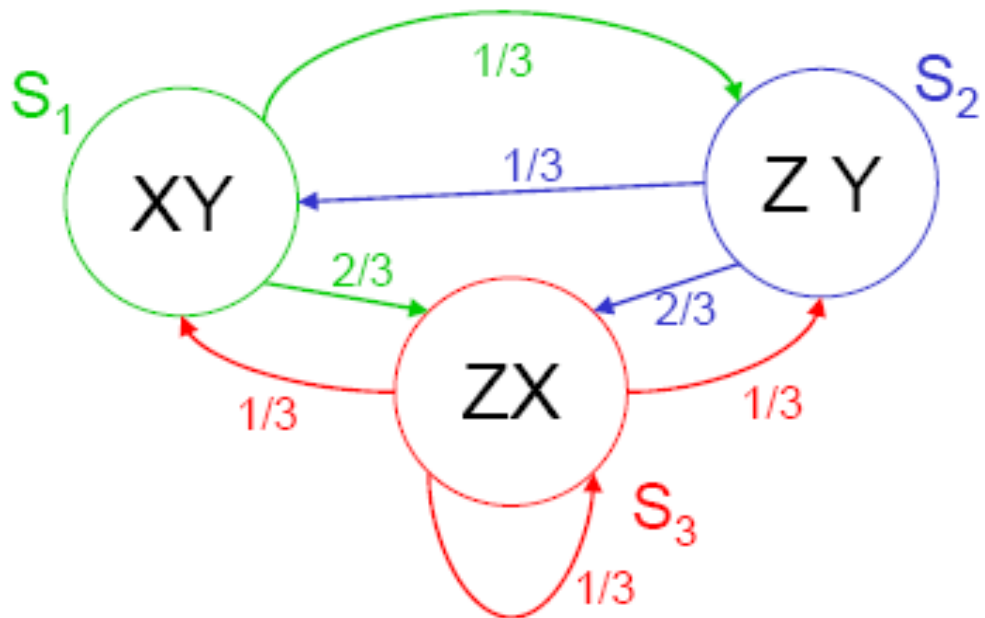


PROBABILITATEA SERIILOR DE OBSERVATII

$$P(\mathbf{O}) = P(O_1 O_2 O_3) = P(O_1 = X \wedge O_2 = X \wedge O_3 = Z)?$$

$$P(\mathbf{O}) = \sum_{Q \in \text{Paths of length 3}} P(\mathbf{O} \wedge \mathbf{Q}) = \sum_{Q \in \text{Paths of length 3}} P(\mathbf{O} | \mathbf{Q}) P(\mathbf{Q})$$

Cum sa calculam $P(Q)$ pentru o cale arbitrara Q ?



$$\begin{aligned} P(Q) &= P(q_1, q_2, q_3) \\ &= P(q_1) P(q_2, q_3 | q_1) \\ &= P(q_1) P(q_2 | q_1) P(q_3 | q_2, q_1) \\ &= P(q_1) P(q_2 | q_1) P(q_3 | q_2) \end{aligned}$$

Ex. in cazul: $Q = S_1 S_3 S_3$:

$$= 1/2 * 2/3 * 1/3 = 1/9$$

Cum sa calculam $P(O|Q)$ pentru o cale arbitrara Q ?

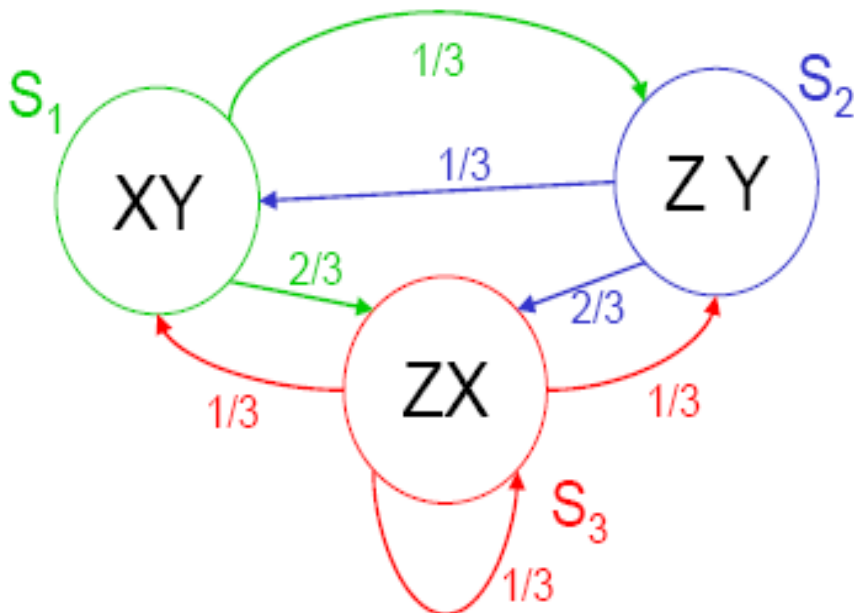
PROBABILITATEA SERIILOR DE OBSERVATII

Ce reprezinta:

$$P(\mathbf{O}) = P(O_1 O_2 O_3) = P(O_1 = X \wedge O_2 = X \wedge O_3 = Z)?$$

$$P(\mathbf{O}) = \sum_{Q \in \text{Paths of length 3}} P(\mathbf{O} \wedge \mathbf{Q}) = \sum_{Q \in \text{Paths of length 3}} P(\mathbf{O} | \mathbf{Q}) P(\mathbf{Q})$$

Cum sa calculam $P(\mathbf{O} | \mathbf{Q})$ pentru o cale arbitrara $\mathbf{Q}$?



$$P(\mathbf{O} | \mathbf{Q}) = P(O_1 O_2 O_3 | q_1 q_2 q_3) \\ = P(O_1 | q_1) P(O_2 | q_2) P(O_3 | q_3)$$

$$\text{Ex. in cazul } Q = S_1 S_3 S_3 \\ = P(X | S_1) P(X | S_3) P(Z | S_3) = \\ = 1/2 * 1/2 * 1/2 = 1/8$$

In exemplu sunt necesare $3^3 = 27$ de calcule de $P(\mathbf{O})$ si $P(\mathbf{O} | \mathbf{Q})$, dar la o secventa de 20 de observatii sunt necesare 3^{20} de calcule de $P(\mathbf{O})$ si $P(\mathbf{O} | \mathbf{Q})$ - ineficient

Probabilitatea unei serii de observatii : non-exponential-cost-style

Fiind date observatiile $O_1 O_2 \dots O_T$

Definim

$$\alpha_t(i) = P(O_1 O_2 \dots O_t \wedge q_t = S_i \mid \lambda), \text{ unde } 1 \leq t \leq T$$

$\alpha_t(i)$ = Probabilitatea ca, într-un proces aleatoriu,
Am văzut primele t observații
Am fi ajuns în S_i ca a t -a stare vizitata.

In exemplu, ce reprezinta $\alpha_2(3)$?

$\alpha_t(i)$: este usor de definit recursiv:

$$\alpha_t(i) = P(O_1 O_2 \dots O_T \wedge q_t = S_i \mid \lambda)$$

$$\alpha_t(i) = P(O_1 O_2 \dots O_T \wedge q_t = S_i \mid \lambda)$$

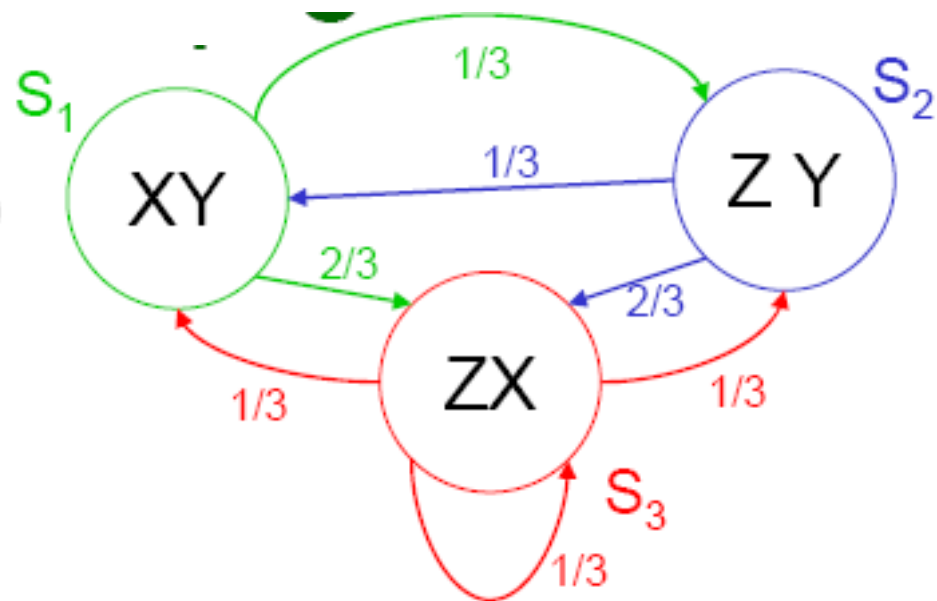
$$\begin{aligned}\alpha_1(i) &= P(O_1 \wedge q_1 = S_i) \\ &= P(q_1 = S_i)P(O_1 | q_1 = S_i) \\ &= \end{aligned}$$

$$\begin{aligned}\alpha_{t+1}(j) &= P(O_1 O_2 \dots O_t O_{t+1} \wedge q_{t+1} = S_j) \\ &= \sum_{i=1}^N P(O_1 O_2 \dots O_t \wedge q_t = S_i \wedge O_{t+1} \wedge q_{t+1} = S_j) \\ &= \sum_{i=1}^N P(O_{t+1}, q_{t+1} = S_j | O_1 O_2 \dots O_t \wedge q_t = S_i) P(O_1 O_2 \dots O_t \wedge q_t = S_i) \\ &= \sum_i P(O_{t+1}, q_{t+1} = S_j | q_t = S_i) \alpha_t(i) \\ &= \sum_i P(q_{t+1} = S_j | q_t = S_i) P(O_{t+1} | q_{t+1} = S_j) \alpha_t(i) \\ &= \sum_i a_{ij} b_j(O_{t+1}) \alpha_t(i)\end{aligned}$$

$$\alpha_t(i) = P(O_1 O_2 \dots O_t \wedge q_t = S_i | \lambda)$$

$$\alpha_1(i) = b_i(O_1) \pi_i$$

$$\alpha_{t+1}(j) = \sum_i a_{ij} b_j(O_{t+1}) \alpha_t(i)$$



Se observa : O1 O2 O3 = X X Z

$$\alpha_1(1) = \frac{1}{4} \quad \alpha_1(2) = 0 \quad \alpha_1(3) = 0$$

$$\alpha_2(1) = 0 \quad \alpha_2(2) = 0 \quad \alpha_2(3) = \frac{1}{12}$$

$$\alpha_3(1) = 0 \quad \alpha_3(2) = \frac{1}{72} \quad \alpha_3(3) = \frac{1}{72}$$

Care este cea mai probabila cale (CMPC) care a generat observatiile : $O_1O_2...O_t$?

Ce reprezinta $\underset{Q}{\operatorname{argmax}} P(Q|O_1O_2...O_T)$?

$$\begin{aligned} & \underset{Q}{\operatorname{argmax}} P(Q|O_1O_2...O_T) \\ &= \underset{Q}{\operatorname{argmax}} \frac{P(O_1O_2...O_T|Q)P(Q)}{P(O_1O_2...O_T)} \\ &= \underset{Q}{\operatorname{argmax}} P(O_1O_2...O_T|Q)P(Q) \end{aligned}$$

Un mod eficient de a calcula CMPC; se calculeaza:

$$\delta_t(i) = \max_{q_1 q_2 \dots q_{t-1}} P(q_1 q_2 \dots q_{t-1} \wedge q_t = S_i \wedge O_1 \dots O_t)$$

= CMPC de lungime t-1 cu sansa maxima de a indeplini toate conditiile:

De a fi urmata si a sfarsi in starea S_i si a produce iesirile $O_1 \dots O_t$

CMPC_t(i) = aceasta cale
Astfel : $\delta_t(i) = \text{Prob}(\text{CMPC}_t(i))$