

Modele Markov Ascunse

De la Teorie la Aplicații

Alexandru Sorici, Tudor Berariu

Asociația Română pentru Inteligență Artificială

5 noiembrie 2012

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

3 Implementarea MMA

- Problema Evaluării: Algoritmul Forward-Backward
- Problema Interpretării: Algoritmul Viterbi
- Problema Estimării: Algoritmul Baum-Welch

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

3 Implementarea MMA

- Problema Evaluării: Algoritmul Forward-Backward
- Problema Interpretării: Algoritmul Viterbi
- Problema Estimării: Algoritmul Baum-Welch

4 Demo: Recunoașterea Simbolurilor

Outline

- 1 Aplicații în Învățarea Automată pentru MMA
 - Învățarea Automată
 - MMA în Învățarea Automată
- 2 Teoria MMA
 - Cele Trei Probleme ale MMA
 - Fundamente Matematice
- 3 Implementarea MMA
 - Problema Evaluării: Algoritmul Forward-Backward
 - Problema Interpretării: Algoritmul Viterbi
 - Problema Estimării: Algoritmul Baum-Welch
- 4 Demo: Recunoașterea Simbolurilor
- 5 Tipuri de MMA

Outline

- 1 Aplicații în Învățarea Automată pentru MMA
 - Învățarea Automată
 - MMA în Învățarea Automată
- 2 Teoria MMA
 - Cele Trei Probleme ale MMA
 - Fundamente Matematice
- 3 Implementarea MMA
 - Problema Evaluării: Algoritmul Forward-Backward
 - Problema Interpretării: Algoritmul Viterbi
 - Problema Estimării: Algoritmul Baum-Welch
- 4 Demo: Recunoașterea Simbolurilor
- 5 Tipuri de MMA
- 6 Discuții și Concluzii

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

3 Implementarea MMA

- Problema Evaluării: Algoritmul Forward-Backward
- Problema Interpretării: Algoritmul Viterbi
- Problema Estimării: Algoritmul Baum-Welch

4 Demo: Recunoașterea Simbolurilor

5 Tipuri de MMA

6 Discuții și Concluzii

What is Machine Learning?

Machine Learning

A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .

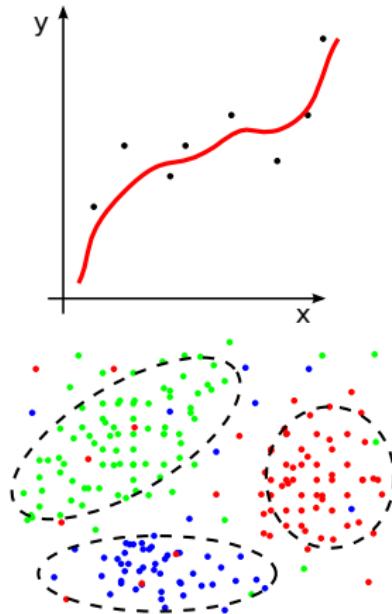
Machine Learning Applications

- Computer Vision: Google Car
- Machine Translation
- Speech Recognition
- Recommender Systems
- Intelligent Advertising

Machine Learning Classification

Types of Machine Learning Problems

- Regression
- Classification
- Reinforcement Learning
- supervised learning (eg. ..)
- unsupervised



Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

3 Implementarea MMA

- Problema Evaluării: Algoritmul Forward-Backward
- Problema Interpretării: Algoritmul Viterbi
- Problema Estimării: Algoritmul Baum-Welch

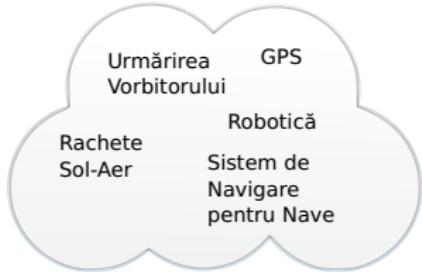
4 Demo: Recunoașterea Simbolurilor

5 Tipuri de MMA

6 Discuții și Concluzii

Probleme cu Secvențe Temporale (I)

URMĂRIREA OBIECTELOR



RECUNOAȘTEREA VORBIRII

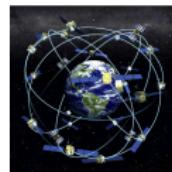
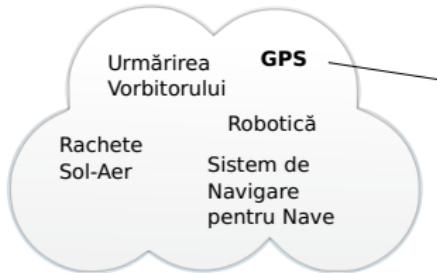


RECUNOAȘTEREA GESTURILOR



Probleme cu Secvențe Temporale (I)

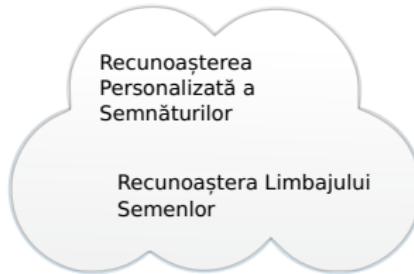
URMĂRIREA OBIECTELOR



RECUNOAȘTEREA VORBIRII

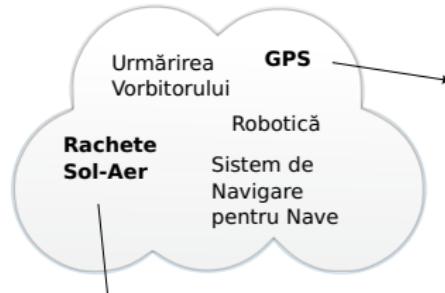


RECUNOAȘTEREA GESTURILOR



Probleme cu Secvențe Temporale (I)

URMĂRIREA OBIECTELOR



RECUNOAȘTEREA VORBIRII

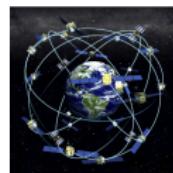
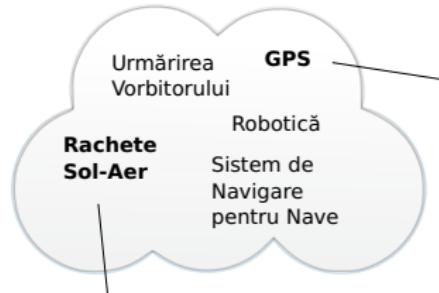


RECUNOAȘTEREA GESTURILOR



Probleme cu Secvențe Temporale (I)

URMĂRIREA OBIECTELOR



RECUNOAȘTEREA VORBIRII

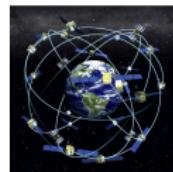
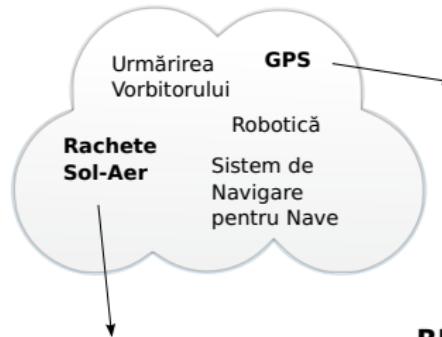


RECUNOAȘTEREA GESTURILOR



Probleme cu Secvențe Temporale (I)

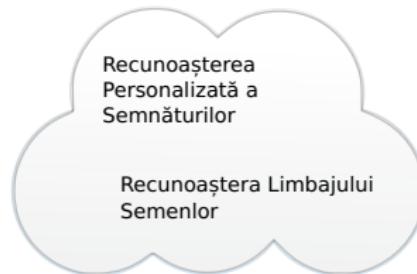
URMĂRIREA OBIECTELOR



RECUNOAȘTEREA VORBIRII

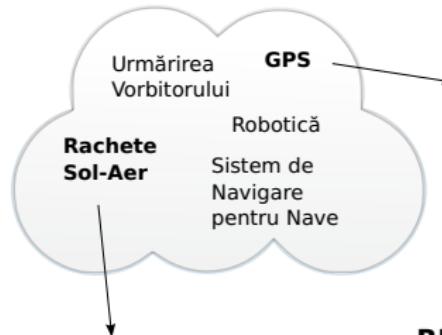


RECUNOAȘTEREA GESTURILOR



Probleme cu Secvențe Temporale (I)

URMĂRIREA OBIECTELOR



RECUNOAȘTEREA VORBIRII



RECUNOAȘTEREA GESTURILOR



Probleme cu Secvențe Temporale (II)

BIOINFORMATICA

Secvențierea Proteinelor

Modelarea unei Rețele
Regulatoare Genetice

ECONOMIE

Predictia Valorilor
Bursiere

Econometrie

- estimate a country's economic indicators across time -

Probleme cu Secvențe Temporale (II)

BIOINFORMATICA

Secvențierea Proteinelor

Modelarea unei Rețele
Regulatoare Genetice



ECONOMIE

Predictia Valorilor
Bursiere

Econometrie
- estimate a country's economic indicators across time -

Raționament Probabilistic Temporal - Modele

Să ne gândim la unele din problemele anterioare ...

Raționament Probabilistic Temporal - Modele

Să ne gândim la unele din problemele anterioare ...

Cum modelăm astfel de situații dinamice?

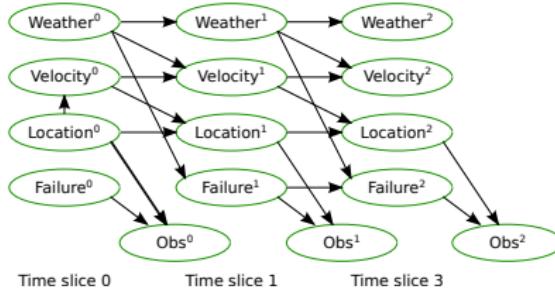
Raționament Probabilistic Temporal - Modele

Să ne gândim la unele din problemele anterioare ...

Cum modelăm astfel de situații dinamice?

Stări și Observații

- Procesul de schimbare este văzut ca o serie de **snapshot-uri**
- Fiecare snapshot conține un set de variabile aleatoare
 - \mathbf{O}_t - setul tuturor variabilelor de măsurare (*observable*) la momentul t
 - \mathbf{Q}_t - setul tuturor variabilelor de stare (*neobservable / ascunse*) la momentul t



Exemplu: problemă de localizare a unui vehicul [KF09]

Raționament Probabilistic Temporal - Presupuneri

Să ne gândim la unele din problemele anterioare ...

Raționament Probabilistic Temporal - Presupuneri

Să ne gândim la unele din problemele anterioare ...

Ce **presupuneri** (la o adică) facem?

Raționament Probabilistic Temporal - Presupuneri

Să ne gândim la unele din problemele anterioare ...

Ce **presupuneri** (la o adică) facem?

Proces staționar

Procesul de schimbare este guvernat de legi care **nu se schimba in timp**.

Urmare: trebuie să specificăm relațiile între variabile doar pentru un snapshot *reprezentativ*.

Raționament Probabilistic Temporal - Presupuneri

Să ne gândim la unele din problemele anterioare ...

Ce **presupuneri** (la o adică) facem?

Proces staționar

Procesul de schimbare este guvernat de legi care **nu se schimba in timp**.

Urmare: trebuie să specificăm relațiile între variabile doar pentru un snapshot *reprezentativ*.

Presupunerea Markov

Starea curentă a unui proces de schimbare depinde doar de o **istorie finită** de stări anterioare.

Urmare: avem un număr **limitat** de "parinți" pentru variabilele din fiecare snapshot.

Raționament Probabilistic Temporal - Inferență

Care sunt principalele inferențe ce se doresc făcute?

Raționament Probabilistic Temporal - Inferență

Care sunt principalele inferențe ce se doresc făcute?

Filtrare (monitorizare)

Sarcina de a calcula **starea de fapt** - distribuția posterioară de probabilitate a **stării curente**, date fiind toate observațiile de până acum.

Raționament Probabilistic Temporal - Inferență

Care sunt principalele inferențe ce se doresc făcute?

Filtrare (monitorizare)

Sarcina de a calcula **starea de fapt** - distribuția posterioară de probabilitate a **stării curente**, date fiind toate observațiile de până acum.

Evaluare

Sarcina de a calcula **probabilitatea (likelihood)** a observațiilor făcute până în prezent.

Raționament Probabilistic Temporal - Inferență

Predictie

Sarcina de a calcula distribuția posterioară de probabilitate peste o **stare viitoare**, date fiind toate observațiile de până acum.

Raționament Probabilistic Temporal - Inferență

Predictie

Sarcina de a calcula distribuția posterioară de probabilitate peste o **stare viitoare**, date fiind toate observațiile de până acum.

Netezire (hindsight)

Sarcina de a calcula distribuția posterioară de probabilitate peste o **stare anterioară**, date fiind toate observațiile de până acum.

Furnizează o estimare mai bună asupra stării respective, decât a fost posibil la momentul respectiv.

Raționament Probabilistic Temporal - Inferență

Cea mai probabilă explicație

Dându-se o secvență de observații, se cere găsirea celei mai probabile secvenței de stări care a generat acele observații.

Raționament Probabilistic Temporal - Inferență

Cea mai probabilă explicație

Dându-se o secvență de observații, se cere găsirea celei mai probabile secvenței de stări care a generat acele observații.

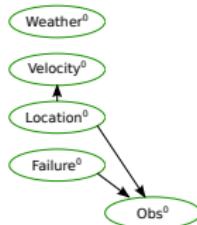
Învățare

Dându-se un set de secvențe de observații, găsește o metodă de a învăța modelele de tranzitie și senzoriale / de măsurare pe baza acestor observații.

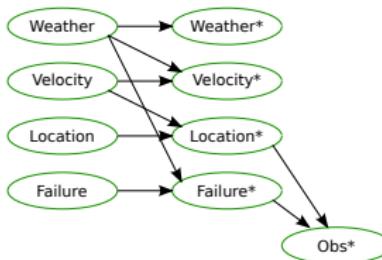
Raționament Probabilistic Temporal - Metode Cunoscute

Rețele Bayesiene Dinamice (RBD)

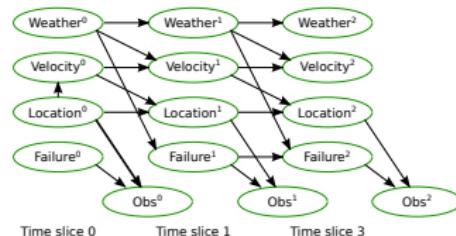
O RBD este o rețea Bayesiană ce reprezintă un model temporal de probabilitate.



(a) Rețeaua Bayesiană în două snapshot-uri



(b) Rețeaua la momentul 0



(c) RBD desfășurată pe 3 momente de timp

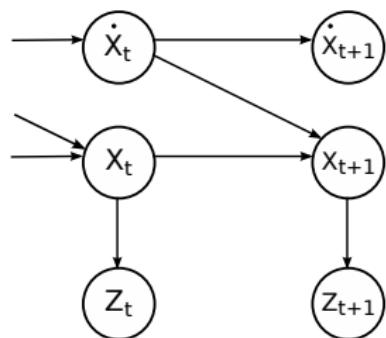
RBD simplificată pentru monitorizarea unui vehicul [KF09]

Aplicată în probleme precum: urmărirea obiectelor, recunoașterea activității umane, sevențierea proteinelor, etc.

Raționament Probabilistic Temporal - Metode Cunoscute

Filtre Kalman (Sistem Dinamice Lineare)

Un model temporal având una sau mai multe variabile care evoluează linear în timp, la care se adaugă **zgomot Gaussian**.



- Poate fi văzut ca o RBD în care toate variabilele sunt continue, iar dependențele sunt linear gaussiane.
- Aplicații multiple în **urmărirea obiectelor**

Structura unei RB pentru un sistem linear dinamic cu variabile de poziție X_t , viteză \dot{X}_t , și măsurare a poziției Z_t

Raționament Probabilistic Temporal - Metode Cunoscute

Modele Markov Ascunse (MMA)

Un MMA (HMM) este un model probabilistic temporal în care *starea* procesului de schimbare este descrisă de ***o singură variabilă aleatoare discretă***. Valorile posibile ale variabilei reprezintă stările posibile ale lumii modelate.

Utilizat cu succes în aplicații precum:

- Recunoașterea Scrisului
- Recunoașterea Gesturilor
- Recunoașterea Vorbirii
- Determinarea Partilor de Vorbire (Part-of-Speech Tagging)
- Secvențiere ADN

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

3 Implementarea MMA

- Problema Evaluării: Algoritmul Forward-Backward
- Problema Interpretării: Algoritmul Viterbi
- Problema Estimării: Algoritmul Baum-Welch

4 Demo: Recunoașterea Simbolurilor

5 Tipuri de MMA

6 Discuții și Concluzii

Cele 3 Probleme ale MMA

3 probleme fundamentale [Rab89]

- Particularizarea inferenței în probleme cu secvențe temporale pe cazul MMA
- Structura restrictionată a MMA permite implementări elegante ale tuturor algoritmilor de bază

Cele 3 Probleme ale MMA

3 probleme fundamentale [Rab89]

- Particularizarea inferenței în probleme cu secvențe temporale pe cazul MMA
- Structura restrictionată a MMA permite implementări elegante ale tuturor algoritmilor de bază

Problema Evaluării

Dându-se un model și o secvență de observații, cum calculăm probabilitatea ca **secvența observată** să fi fost produsă de model?

Cele 3 Probleme ale MMA

3 probleme fundamentale [Rab89]

- Particularizarea inferenței în probleme cu secvențe temporale pe cazul MMA
- Structura restricționată a MMA permite implementări elegante ale tuturor algoritmilor de bază

Problema Evaluării

Dându-se un model și o secvență de observații, cum calculăm probabilitatea ca secvența observată să fi fost produsă de model?

Problema Interpretării (cea mai bună explicație a observațiilor)

Dându-se un model și o secvență de observații, cum alegem o secvență corespunzătoare de stări care dău sens observațiilor?

Cele 3 Probleme ale MMA

3 probleme fundamentale [Rab89]

- Particularizarea inferenței în probleme cu secvențe temporale pe cazul MMA
- Structura restricționată a MMA permite implementări elegante ale tuturor algoritmilor de bază

Problema Evaluării

Dându-se un model și o secvență de observații, cum calculăm probabilitatea ca secvența observată să fi fost produsă de model?

Problema Interpretării (cea mai bună explicație a observațiilor)

Dându-se un model și o secvență de observații, cum alegem o secvență corespunzătoare de stări care dă sens observațiilor?

Problema Estimării (Antrenării) Modelului

Dându-se mai multe secvențe de observații, cum putem ajusta parametrii modelului MMA care explică cel mai bine observațiile făcute?

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

3 Implementarea MMA

- Problema Evaluării: Algoritmul Forward-Backward
- Problema Interpretării: Algoritmul Viterbi
- Problema Estimării: Algoritmul Baum-Welch

4 Demo: Recunoașterea Simbolurilor

5 Tipuri de MMA

6 Discuții și Concluzii

Exemplu: Urmărirea stărilor emotionale

Să considerăm următorul exemplu:

- un robot ce urmărește evoluția stărilor emotionale ale unui om

Senzor:

- cameră video

Să modelăm împreună această problemă definind componentele unui Model Markov Ascuns!

Exemplu: Urmărirea stărilor emoționale

s₁:vesel

s₂:trist

s₃:nervos

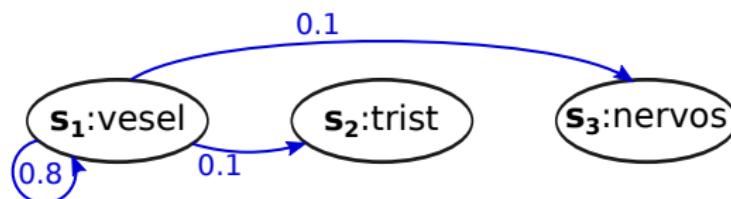
N - numărul de stări ascunse

$$\mathbf{N} = 3$$

Stări:

- **s₁:** vesel
- **s₂:** trist
- **s₃**: nervos

Exemplu: Urmărirea stărilor emoționale



A - matricea distribuțiilor de probabilitate ale tranzițiilor între stări

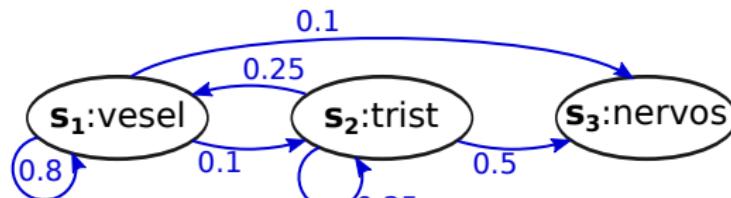
$$\mathbf{A} = \{a_{i,j}\}, 1 \leq i, j \leq N$$

$$a_{i,j} = P(q_{t+1} = s_j | q_t = s_i)$$

$$\sum_{j=1}^N a_{i,j} = 1, \quad 1 \leq i \leq N$$

$$\mathbf{A} = \begin{matrix} & s_1 & s_2 & s_3 \\ s_1 & 0.8 & 0.1 & 0.1 \\ s_2 & & & \\ s_3 & & & \end{matrix}$$

Exemplu: Urmărirea stărilor emoționale



A - matricea distribuțiilor de probabilitate ale tranzițiilor între stări

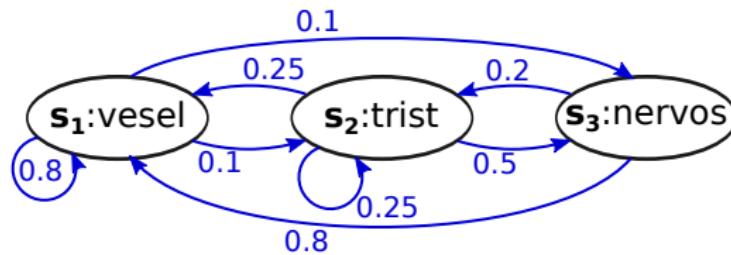
$$\mathbf{A} = \{a_{i,j}\}, 1 \leq i, j \leq N$$

$$a_{i,j} = P(q_{t+1} = s_j | q_t = s_i)$$

$$\sum_{j=1}^N a_{i,j} = 1, \quad 1 \leq i \leq N$$

$$\mathbf{A} = \begin{matrix} & s_1 & s_2 & s_3 \\ s_1 & 0.8 & 0.1 & 0.1 \\ s_2 & 0.25 & 0.25 & 0.5 \\ s_3 & & & \end{matrix}$$

Exemplu: Urmărirea stărilor emoționale



A - matricea distribuțiilor de probabilitate ale tranzițiilor între stări

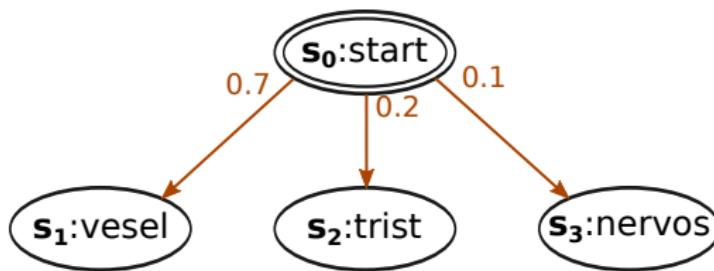
$$\mathbf{A} = \{a_{i,j}\}, 1 \leq i, j \leq N$$

$$a_{i,j} = P(q_{t+1} = s_j | q_t = s_i)$$

$$\sum_{j=1}^N a_{i,j} = 1, \quad 1 \leq i \leq N$$

$$\mathbf{A} = \begin{matrix} & s_1 & s_2 & s_3 \\ s_1 & 0.8 & 0.1 & 0.1 \\ s_2 & 0.25 & 0.25 & 0.5 \\ s_3 & 0.8 & 0.2 & 0 \end{matrix}$$

Exemplu: Urmărirea stărilor emoționale



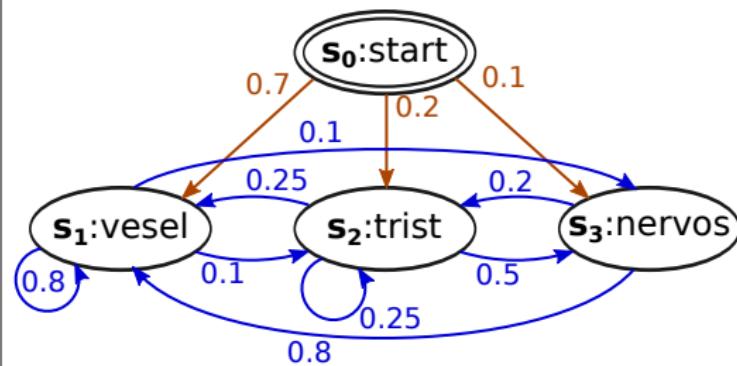
Π - distribuția stării inițiale

$$\Pi = \{\pi_i\}, \quad 1 \leq i \leq N$$

$$\pi_i = P(q_1 = s_i)$$

$$\Pi = \begin{pmatrix} s_1 & s_2 & s_3 \\ 0.7 & 0.2 & 0.1 \end{pmatrix}$$

Exemplu: Urmărirea stărilor emoționale

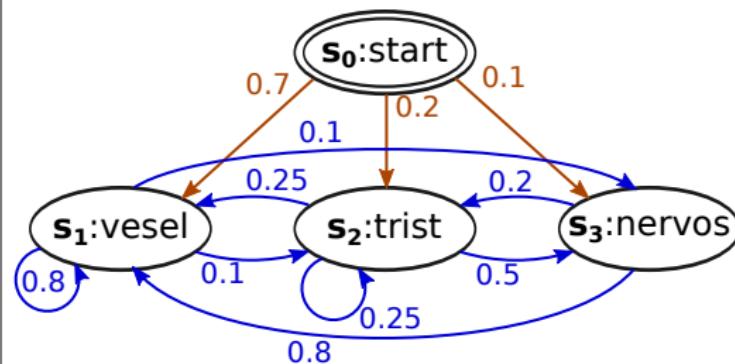


Deocamdată am descris un lanț Markov.

$$A = \begin{pmatrix} s_1 & s_2 & s_3 \\ s_1 & 0.8 & 0.1 & 0.1 \\ s_2 & 0.25 & 0.25 & 0.5 \\ s_3 & 0.8 & 0.2 & 0 \end{pmatrix}$$

$$\Pi = \begin{pmatrix} s_1 & s_2 & s_3 \\ (0.7 & 0.2 & 0.1) \end{pmatrix}$$

Exemplu: Urmărirea stărilor emoționale



vesel → vesel → nervos → trist

$$\mathbf{Q} = [q_1:s_1 \ q_2:s_1 \ q_3:s_3 \ q_4:s_2]$$

Deocamdată am descris un lanț Markov.

$$A = \begin{pmatrix} s_1 & s_2 & s_3 \\ s_1 & 0.8 & 0.1 & 0.1 \\ s_2 & 0.25 & 0.25 & 0.5 \\ s_3 & 0.8 & 0.2 & 0 \end{pmatrix}$$

$$\Pi = \begin{pmatrix} s_1 & s_2 & s_3 \\ 0.7 & 0.2 & 0.1 \end{pmatrix}$$

Notație: $\mathbf{Q} = [q_1 q_2 \cdots q_T]$

$$P(Q|A, \Pi) = \pi_{q_1} a_{q_1, q_2} \cdots a_{q_{T-1}, q_T}$$

$$P(s_1, s_1, s_3, s_2 | A, \Pi) = \pi_{q_1} \cdot a_{1,1} \cdot a_{1,3} \cdot a_{3,2} \\ = 0.8 \cdot 0.8 \cdot 0.1 \cdot 0.2 = 0.0128$$

Exemplu: Urmărirea stărilor emoționale

s₁: vesel

s₂: trist

s₃: nervos

v₁: zâmbet
/ rânjet

v₂: nimic

v₃: încruntare

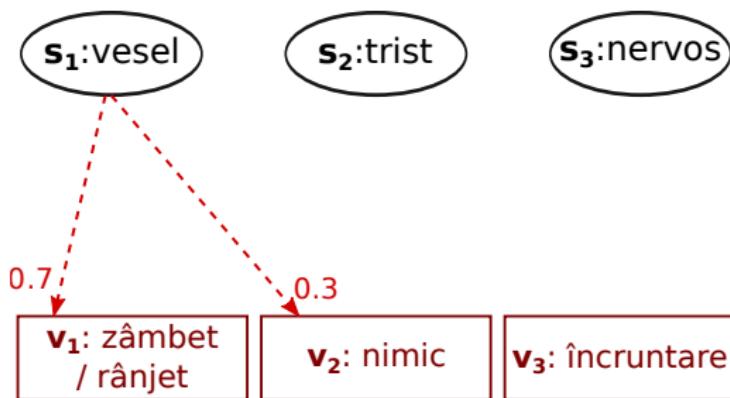
M - numărul de valori observabile distincte

$$M = 3$$

valori observabile:

- **v₁:** zâmbet / rânjet
- **v₂:** nimic
- **v₃:** încruntare

Exemplu: Urmărirea stărilor emoționale



B - matricea distribuțiilor de probabilitate ale valorilor observabile

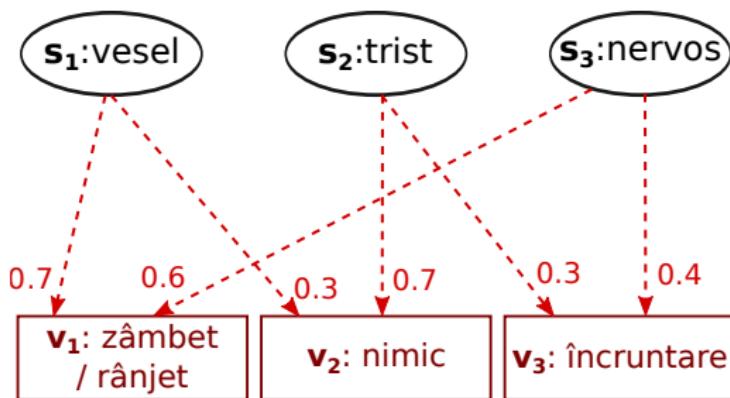
$$\mathbf{B} = \{b_{j,k}\}_{1 \leq j \leq N, 1 \leq k \leq M}$$

$$\begin{aligned} b_{j,k} &= b_j(v_k) \\ &= P(o_t = v_k | q_t = s_j) \end{aligned}$$

$$\sum_{k=1}^M b_{j,k} = 1, \quad 1 \leq j \leq N$$

$$\mathbf{B} = \begin{matrix} s_1 & v_1 & v_2 & v_3 \\ s_2 & 0 & 0.2 & 0.8 \\ s_3 & & & \end{matrix}$$

Exemplu: Urmărirea stărilor emoționale



B - matricea distribuțiilor de probabilitate ale valorilor observabile

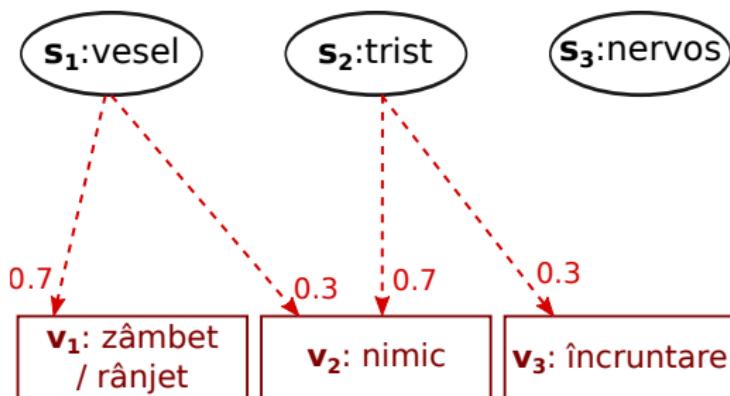
$$\mathbf{B} = \{b_{j,k}\}_{1 \leq j \leq N, 1 \leq k \leq M}$$

$$\begin{aligned} b_{j,k} &= b_j(v_k) \\ &= P(o_t = v_k | q_t = s_j) \end{aligned}$$

$$\sum_{k=1}^M b_{j,k} = 1, \quad 1 \leq j \leq N$$

$$\mathbf{B} = \begin{matrix} & v_1 & v_2 & v_3 \\ s_1 & 0 & 0.2 & 0.8 \\ s_2 & 0.3 & 0.7 & 0 \\ s_3 & 0.4 & 0.3 & 0.3 \end{matrix}$$

Exemplu: Urmărirea stărilor emoționale



B - matricea distribuțiilor de probabilitate ale valorilor observabile

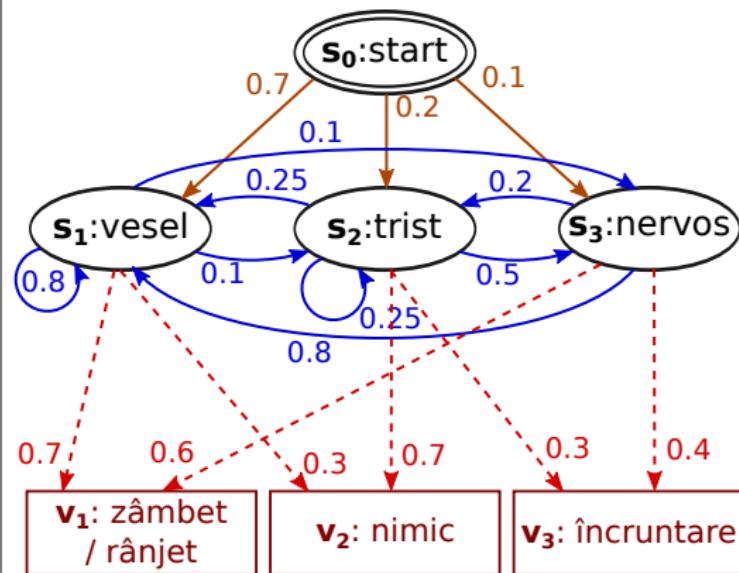
$$\mathbf{B} = \{b_{j,k}\}_{1 \leq j \leq N, 1 \leq k \leq M}$$

$$\begin{aligned} b_{j,k} &= b_j(v_k) \\ &= P(o_t = v_k | q_t = s_j) \end{aligned}$$

$$\sum_{k=1}^M b_{j,k} = 1, \quad 1 \leq j \leq N$$

$$\mathbf{B} = \begin{matrix} & v_1 & v_2 & v_3 \\ s_1 & 0 & 0.2 & 0.8 \\ s_2 & 0.3 & 0.7 & 0 \\ s_3 & 0.4 & 0 & 0.6 \end{matrix}$$

Exemplu: Urmărirea stărilor emoționale



λ - parametrii Modelului Markov Ascuns

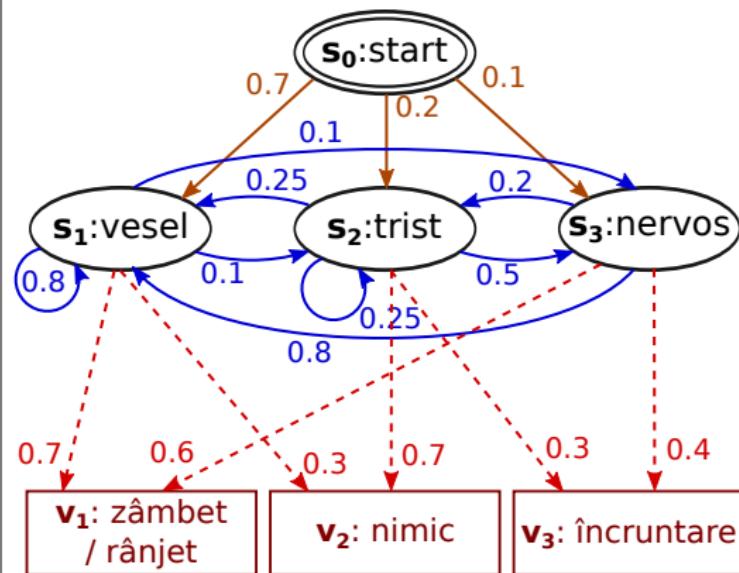
$$\lambda = (A, B, \Pi)$$

A - matricea distribuțiilor de probabilitate ale tranzițiilor între stări

B - matricea distribuțiilor de probabilitate ale valorilor observabile

Π - distribuția stării initiale

Exemplu: Urmărirea stărilor emotionale



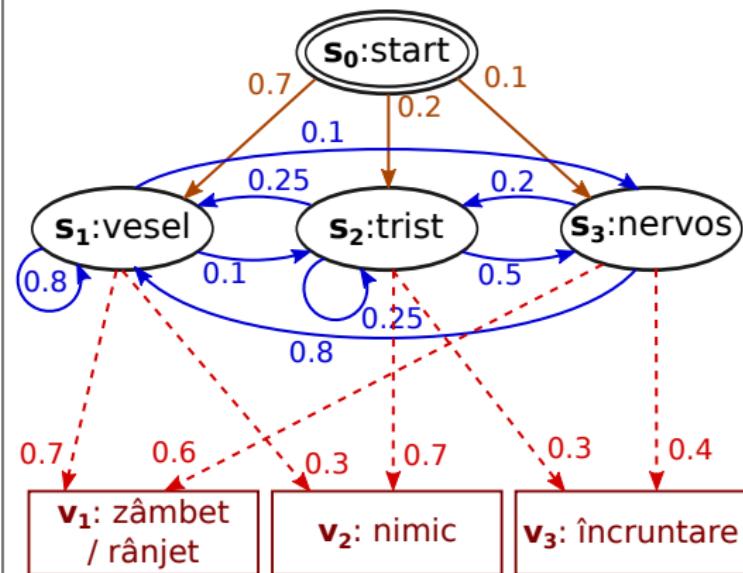
O - secvența de observații

T - lungimea secvenței de observații

$$O = [o_1 o_2 \cdots o_T]$$

$$O = [o_1:zâmbet \quad o_2:nimic \quad o_3:nimic \\ o_4:încruntare \quad o_5:zâmbet]$$

Exemplu: Urmărirea stărilor emotionale



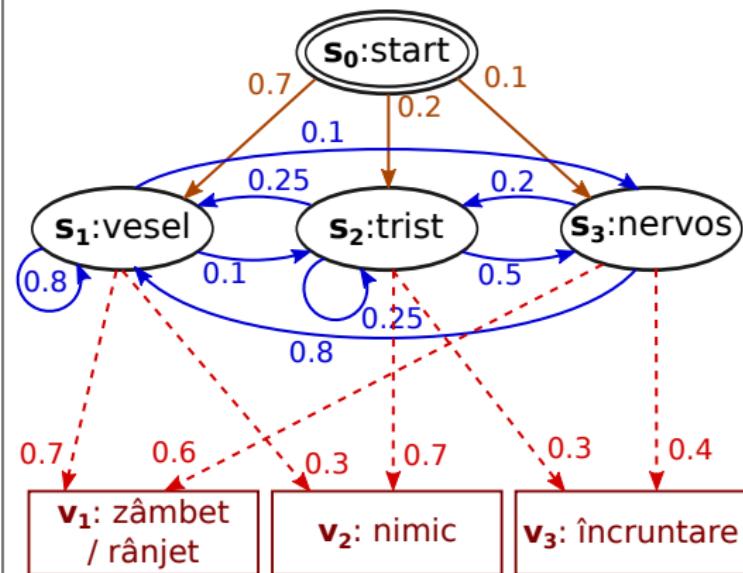
O - secvența de observații

T - lungimea secvenței de observații

$$O = [o_1 o_2 \cdots o_T]$$

zâmbet, nimic, nimic, încruntare, zâmbet

Exemplu: Urmărirea stărilor emotionale



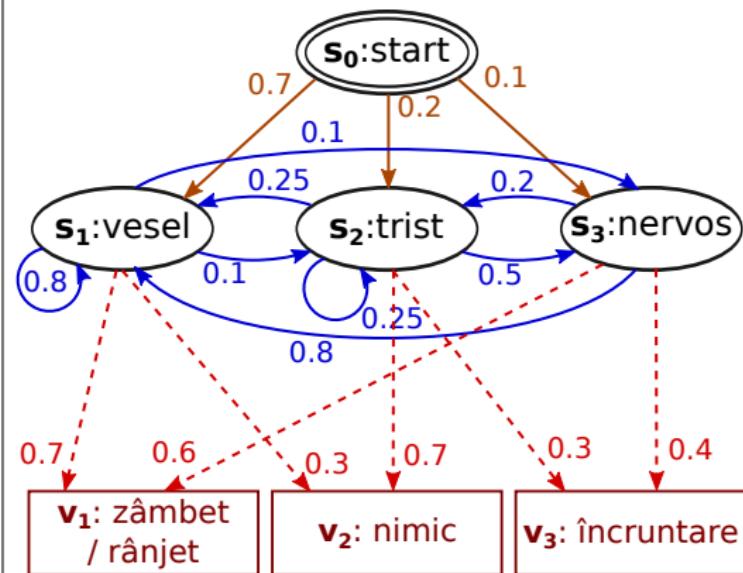
$\text{vesel} \rightarrow \text{vesel} \rightarrow \text{vesel} \rightarrow \text{trist} \rightarrow \text{nervos}$
 $\text{zâmbet}, \text{nimic}, \text{nimic}, \text{încruntare}, \text{zâmbet}$

O - secvența de observații

T - lungimea secvenței de observații

$$O = [o_1 o_2 \cdots o_T]$$

Exemplu: Urmărirea stărilor emotionale



O - secvența de observații

T - lungimea secvenței de observații

$$O = [o_1 o_2 \cdots o_T]$$

Exemplu: Urmărirea stărilor emoționale

- Exemplul a fost adaptat după:

R. Zubek. Introduction to hidden markov models. *AI Game Programming Wisdom*, 3:633–646, 2006

Reformularea celor 3 probleme fundamentale ale MMA

Problema evaluării

Date fiind un model și o secvență de observații , cum calculăm probabilitatea ca secvența de observații să fi fost generată de acel model?

Reformularea celor 3 probleme fundamentale ale MMA

Problema evaluării

Date fiind un model $\lambda = (A, B, \Pi)$ și o secvență de observații , cum calculăm probabilitatea ca secvența de observații să fi fost generată de acel model?

Reformularea celor 3 probleme fundamentale ale MMA

Problema evaluării

Date fiind un model $\lambda = (A, B, \Pi)$ și o secvență de observații $O = [o_1 o_2 \cdots o_T]$, cum calculăm probabilitatea ca secvența de observații să fi fost generată de acel model?

Reformularea celor 3 probleme fundamentale ale MMA

Problema evaluării

Date fiind un model $\lambda = (A, B, \Pi)$ și o secvență de observații $O = [o_1 o_2 \cdots o_T]$, cum calculăm probabilitatea $P(O|\lambda)$ ca secvența de observații să fi fost generată de acel model?

Reformularea celor 3 probleme fundamentale ale MMA

Problema evaluării

Date fiind un model $\lambda = (A, B, \Pi)$ și o secvență de observații $O = [o_1 o_2 \cdots o_T]$, cum calculăm probabilitatea $P(O|\lambda)$ ca secvența de observații să fi fost generată de acel model?

- Prin enumerarea tuturor secvențelor posibile de stări:

$$P(O|\lambda) = \sum_{\text{all } Q} P(O|Q, \lambda) \cdot P(Q|\lambda) \quad (1)$$

Reformularea celor 3 probleme fundamentale ale MMA

$$P(O|\lambda) = \sum_{\text{all } Q} P(O|Q, \lambda) \cdot P(Q|\lambda) \quad (1)$$

Reformularea celor 3 probleme fundamentale ale MMA

$$P(O|\lambda) = \sum_{\text{all } Q} P(O|Q, \lambda) \cdot P(Q|\lambda) \quad (1)$$

$$P(O|Q, \lambda) = \prod_{t=1}^T P(o_t|q_t, \lambda) = \prod_{t=1}^T b_{q_t}(o_t) = b_{q_1}(o_1) \cdot \dots \cdot b_{q_T}(o_T) \quad (2)$$

Reformularea celor 3 probleme fundamentale ale MMA

$$P(O|\lambda) = \sum_{\text{all } Q} P(O|Q, \lambda) \cdot P(Q|\lambda) \quad (1)$$

$$P(O|Q, \lambda) = \prod_{t=1}^T P(o_t|q_t, \lambda) = \prod_{t=1}^T b_{q_t}(o_t) = b_{q_1}(o_1) \cdot \dots \cdot b_{q_T}(o_T) \quad (2)$$

$$P(Q|\lambda) = \pi_{q_1} \prod_{t=2}^T a_{q_{t-1}, q_t} = \pi_{q_1} \cdot a_{q_1, q_2} \cdot a_{q_2, q_3} \cdot \dots \cdot a_{q_{T-1}, q_T} \quad (3)$$

Reformularea celor 3 probleme fundamentale ale MMA

$$P(O|\lambda) = \sum_{\text{all } Q} P(O|Q, \lambda) \cdot P(Q|\lambda) \quad (1)$$

$$P(O|Q, \lambda) = \prod_{t=1}^T P(o_t|q_t, \lambda) = \prod_{t=1}^T b_{q_t}(o_t) = b_{q_1}(o_1) \cdot \dots \cdot b_{q_T}(o_T) \quad (2)$$

$$P(Q|\lambda) = \pi_{q_1} \prod_{t=2}^T a_{q_{t-1}, q_t} = \pi_{q_1} \cdot a_{q_1, q_2} \cdot a_{q_2, q_3} \cdot \dots \cdot a_{q_{T-1}, q_T} \quad (3)$$

$$\begin{aligned} P(O|\lambda) &= \sum_{\text{all } Q} P(O, Q|\lambda) = \sum_{\text{all } Q} P(O, |Q, \lambda) \cdot P(Q, \lambda) \\ &= \sum_{\text{all } Q} \left(\pi_{q_1} \cdot b_{q_1}(o_1) \cdot \prod_{t=2}^T b_{q_t}(o_t) a_{q_{t-1}, q_t} \right) \end{aligned} \quad (1)$$

Reformularea celor 3 probleme fundamentale ale MMA

Problema explicării unei secvențe de observații

Date fiind un model și o secvență de observații , cum alegem o secvență corespunzătoare de stări care să dea un înțeles observațiilor? Cum descoperim partea ascunsă a modelului?

Reformularea celor 3 probleme fundamentale ale MMA

Problema explicării unei secvențe de observații

Date fiind un model $\lambda = (A, B, \Pi)$ și o secvență de observații
, cum alegem o secvență corespunzătoare de stări
care să dea un înțeles observațiilor? Cum descoperim
partea ascunsă a modelului?

Reformularea celor 3 probleme fundamentale ale MMA

Problema explicării unei secvențe de observații

Date fiind un model $\lambda = (A, B, \Pi)$ și o secvență de observații

$O = [o_1 o_2 \cdots o_T]$, cum alegem o secvență corespunzătoare de stări
care să dea un înțeles observațiilor? Cum descoperim
partea ascunsă a modelului?

Reformularea celor 3 probleme fundamentale ale MMA

Problema explicării unei secvențe de observații

Date fiind un model $\lambda = (A, B, \Pi)$ și o secvență de observații

$O = [o_1 o_2 \cdots o_T]$, cum alegem o secvență corespunzătoare de stări

$Q = [q_1 q_2 \cdots q_T]$ care să dea un înțeles observațiilor? Cum descoperim partea ascunsă a modelului?

Reformularea celor 3 probleme fundamentale ale MMA

Problema explicării unei secvențe de observații

Date fiind un model $\lambda = (A, B, \Pi)$ și o secvență de observații

$O = [o_1 o_2 \cdots o_T]$, cum alegem o secvență corespunzătoare de stări

$Q = [q_1 q_2 \cdots q_T]$ care să dea un înțeles observațiilor? Cum descoperim partea ascunsă a modelului?

- Există mai multe criterii pentru *cea mai bună* sevență
 - Secvența celor mai probabile stări (luate individual):

$$Q_{\text{best}} = [\hat{q}_1 \hat{q}_2 \dots \hat{q}_T], \quad \hat{q}_t = \underset{s_i}{\operatorname{argmax}} P(q_t = s_i | O, \lambda) \quad (4)$$

Reformularea celor 3 probleme fundamentale ale MMA

Problema explicării unei secvențe de observații

Date fiind un model $\lambda = (A, B, \Pi)$ și o secvență de observații

$O = [o_1 o_2 \cdots o_T]$, cum alegem o secvență corespunzătoare de stări

$Q = [q_1 q_2 \cdots q_T]$ care să dea un înțeles observațiilor? Cum descoperim partea ascunsă a modelului?

- Există mai multe criterii pentru *cea mai bună* sevență
 - Secvența celor mai probabile stări (luate individual):

$$Q_{\text{best}} = [\hat{q}_1 \hat{q}_2 \dots \hat{q}_T], \quad \hat{q}_t = \underset{s_i}{\operatorname{argmax}} P(q_t = s_i | O, \lambda) \quad (4)$$

- Cea mai bună *cale* (de dimensiune T)

$$Q_{\text{best}} = \underset{Q}{\operatorname{argmax}} P(Q | O, \lambda) = \underset{Q}{\operatorname{argmax}} P(Q, O | \lambda) \quad (5)$$

Reformularea celor 3 probleme fundamentale ale MMA

Problema Estimării Modelului (Învățării)

Date fiind niște secvențe de observații , cum ajustăm parametrii ai unui MMA astfel încât să explice cel mai bine observațiile?

Reformularea celor 3 probleme fundamentale ale MMA

Problema Estimării Modelului (Învățării)

Date fiind niște secvențe de observații $\mathcal{O} = [O_1 O_2 \cdots O_L]$, cum *ajustăm parametrii* ai unui MMA astfel încât să explice cel mai bine observațiile?

Reformularea celor 3 probleme fundamentale ale MMA

Problema Estimării Modelului (Învățării)

Date fiind niște secvențe de observații $\mathcal{O} = [O_1 O_2 \cdots O_L]$, cum *ajustăm parametrii $\lambda = (A, B, \Pi)$* ai unui MMA astfel încât să explice cel mai bine observațiile?

Reformularea celor 3 probleme fundamentale ale MMA

Problema Estimării Modelului (Învățării)

Date fiind niște secvențe de observații $\mathcal{O} = [O_1 O_2 \cdots O_L]$, cum *ajustăm parametrii $\lambda = (A, B, \Pi)$* ai unui MMA astfel încât să explice cel mai bine observațiile?

- Întrebarea se poate reformula matematic:

$$\lambda_{\text{best}} = \underset{\lambda}{\operatorname{argmax}} P(\mathcal{O}|\lambda) \quad (6)$$

:)

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

3 Implementarea MMA

- **Problema Evaluării: Algoritmul Forward-Backward**
- Problema Interpretării: Algoritmul Viterbi
- Problema Estimării: Algoritmul Baum-Welch

4 Demo: Recunoașterea Simbolurilor

5 Tipuri de MMA

6 Discuții și Concluzii

:)

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

3 Implementarea MMA

- Problema Evaluării: Algoritmul Forward-Backward
- **Problema Interpretării: Algoritmul Viterbi**
- Problema Estimării: Algoritmul Baum-Welch

4 Demo: Recunoașterea Simbolurilor

5 Tipuri de MMA

6 Discuții și Concluzii

:)

Outline

1 Aplicații în Învățarea Automată pentru MMA

- Învățarea Automată
- MMA în Învățarea Automată

2 Teoria MMA

- Cele Trei Probleme ale MMA
- Fundamente Matematice

3 Implementarea MMA

- Problema Evaluării: Algoritmul Forward-Backward
- Problema Interpretării: Algoritmul Viterbi
- **Problema Estimării: Algoritmul Baum-Welch**

4 Demo: Recunoașterea Simbolurilor

5 Tipuri de MMA

6 Discuții și Concluzii

Învățarea din observații - Amintire

Problema Estimării (Antrenării) Modelului

Dându-se un set de secvențe observate , cum ajustam parameterii ai unui MMA care încearcă să explice cel mai bine acele observații?

Învățarea din observații - Amintire

Problema Estimării (Antrenării) Modelului

Dându-se un set de secvențe observate $\mathcal{O} = [O_1 O_2 \dots O_L]$, cum ajustam parameterii ai unui MMA care încearcă să explice cel mai bine acele observații?

Secvențele de observații folosite pentru ajustarea parametrilor modelului se numesc secvențe **de antrenare**.

Problema antrenării este esențială - ea permite crearea celor mai bune modele pentru fenomene reale.

Învățarea din observații - Amintire

Problema Estimării (Antrenării) Modelului

Dându-se un set de secvențe observate $\mathcal{O} = [O_1 O_2 \dots O_L]$, cum ajustam parameterii $\lambda = (A, B, \Pi)$ ai unui MMA care încearcă să explice cel mai bine acele observații?

Secvențele de observații folosite pentru ajustarea parametrilor modelului se numesc secvențe **de antrenare**.

Problema antrenării este esențială - ea permite crearea celor mai bune modele pentru fenomene reale.

Învățarea din observații - Abordare

Învățarea din observații - Abordare

Problemă

Nu se cunoaște o metodă analitică de căutare a parametrilor modelului care *maximizează* probabilitatea secvențelor observate.

Învățarea din observații - Abordare

Problemă

Nu se cunoaște o metodă analitică de căutare a parametrilor modelului care *maximizează* probabilitatea secvențelor observate.

Soluție

Putem totuși găsi $\lambda = (A, B, \Pi)$, astfel încât $\max_{\lambda} P(O|\lambda)$ corespunde unui **maxim local**, utilizând o **procedură iterativă** precum *algoritmul Baum-Welch*.

Această metodă este o instanță a *algoritmului EM (Expectation Maximization) [DLR77]* pentru cazul MMA.

Algoritmul Baum-Welch (I)

Procedura în descriere conceptuală:

- ① Avem MMA $\lambda = (A, B, \Pi)$ și o secvență observată O

Algoritmul Baum-Welch (I)

Procedura în descriere conceptuală:

- ① Avem MMA $\lambda = (A, B, \Pi)$ și o secvență observată O
- ② Calculăm folosindu-ne de parametrii $\alpha_t(i)$ și $\beta_t(i)$
 - nr. estimat de tranziții din S_i , pentru fiecare $1 \leq i \leq N$
 - nr. estimat de tranziții din S_i la S_j , pentru fiecare $1 \leq i \leq N, 1 \leq j \leq N$
 - nr. estimat de vizite în S_j observând simbolul v_k , pentru fiecare $1 \leq j \leq N, 1 \leq k \leq M$

Algoritmul Baum-Welch (I)

Procedura în descriere conceptuală:

- ① Avem MMA $\lambda = (A, B, \Pi)$ și o secvență observată O
- ② Calculăm folosindu-ne de parametrii $\alpha_t(i)$ și $\beta_t(i)$
 - nr. estimat de tranziții din S_i , pentru fiecare $1 \leq i \leq N$
 - nr. estimat de tranziții din S_i la S_j , pentru fiecare $1 \leq i \leq N, 1 \leq j \leq N$
 - nr. estimat de vizite în S_j observând simbolul v_k , pentru fiecare $1 \leq j \leq N, 1 \leq k \leq M$
- ③ Dacă modelul este corect ne așteptăm ca
 - (a) $\Pi_i = \text{nr. estimat de vizite în starea } S_i \text{ la momentul } (t=1) = \bar{\Pi}_i$
 - (b) $a_{i,j} = \frac{\text{nr. estimat de tranziții din } s_i \text{ la } s_j}{\text{nr. estimat de tranziții din } s_i} = \bar{a}_{i,j}$
 - (c) $b_{j,k} = \frac{\text{nr. estimat de vizite în } s_j \text{ observând simbolul } v_k}{\text{nr. estimat de vizite în } s_j} = \bar{b}_{j,k}$

Algoritmul Baum-Welch (I)

Procedura în descriere conceptuală:

- ① Avem MMA $\lambda = (A, B, \Pi)$ și o secvență observată O
- ② Calculăm folosindu-ne de parametrii $\alpha_t(i)$ și $\beta_t(i)$
 - nr. estimat de tranziții din S_i , pentru fiecare $1 \leq i \leq N$
 - nr. estimat de tranziții din S_i la S_j , pentru fiecare $1 \leq i \leq N, 1 \leq j \leq N$
 - nr. estimat de vizite în S_j observând simbolul v_k , pentru fiecare $1 \leq j \leq N, 1 \leq k \leq M$
- ③ Dacă modelul este corect ne așteptăm ca
 - (a) $\Pi_i = \text{nr. estimat de vizite în starea } S_i \text{ la momentul } (t=1) = \bar{\Pi}_i$
 - (b) $a_{i,j} = \frac{\text{nr. estimat de tranziții din } s_i \text{ la } s_j}{\text{nr. estimat de tranziții din } s_i} = \bar{a}_{i,j}$
 - (c) $b_{j,k} = \frac{\text{nr. estimat de vizite în } s_j \text{ observând simbolul } v_k}{\text{nr. estimat de vizite în } s_j} = \bar{b}_{j,k}$
- ④ Vedem că raporturile calculate din presupunerea noastră explică mai bine observația decât parametrii anteriori, i.e. $P(O|\bar{\lambda}) > P(O|\lambda)$

Algoritmul Baum-Welch (I)

Procedura în descriere conceptuală:

- ① Avem MMA $\lambda = (A, B, \Pi)$ și o secvență observată O
- ② Calculăm folosindu-ne de parametrii $\alpha_t(i)$ și $\beta_t(i)$
 - nr. estimat de tranziții din S_i , pentru fiecare $1 \leq i \leq N$
 - nr. estimat de tranziții din S_i la S_j , pentru fiecare $1 \leq i \leq N, 1 \leq j \leq N$
 - nr. estimat de vizite în S_j observând simbolul v_k , pentru fiecare $1 \leq j \leq N, 1 \leq k \leq M$
- ③ Dacă modelul este corect ne așteptăm ca
 - (a) $\Pi_i = \text{nr. estimat de vizite în starea } S_i \text{ la momentul } (t=1) = \bar{\Pi}_i$
 - (b) $a_{i,j} = \frac{\text{nr. estimat de tranziții din } s_i \text{ la } s_j}{\text{nr. estimat de tranziții din } s_i} = \bar{a}_{i,j}$
 - (c) $b_{j,k} = \frac{\text{nr. estimat de vizite în } s_j \text{ observând simbolul } v_k}{\text{nr. estimat de vizite în } s_j} = \bar{b}_{j,k}$
- ④ Vedem că raporturile calculate din presupunerea noastră explică mai bine observația decât parametrii anteriori, i.e. $P(O|\bar{\lambda}) > P(O|\lambda)$
- ⑤ Atunci repetăm procesul până ce suntem mulțumiți (convergență: $P(O|\bar{\lambda}) - P(O|\lambda) \leq \epsilon$)

Algoritmul Baum-Welch (II)

Algoritmul Baum-Welch (II)

Definim întâi niște variabile auxiliare:

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

Probabilitatea de a fi în starea s_i la momentul t și în starea s_j la momentul $t + 1$, condiționat de parametrii modelului curent și secvența observată.

Algoritmul Baum-Welch (II)

Definim întâi niște variabile auxiliare:

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

Probabilitatea de a fi în starea s_i la momentul t și în starea s_j la momentul $t + 1$, condiționat de parametrii modelului curent și secvența observată.

$$\gamma_{t,i} = \gamma_t(i) = P(q_t = s_i | O, \lambda)$$

Probabilitatea de a fi în starea s_i la momentul t , condiționat de parametrii modelului curent și secvența observată.

Algoritmul Baum-Welch (II)

Definim întâi niște variabile auxiliare:

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

Probabilitatea de a fi în starea s_i la momentul t și în starea s_j la momentul $t + 1$, condiționat de parametrii modelului curent și secvența observată.

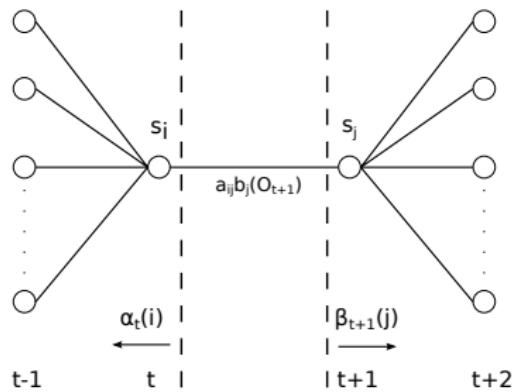
$$\gamma_{t,i} = \gamma_t(i) = P(q_t = s_i | O, \lambda)$$

Probabilitatea de a fi în starea s_i la momentul t , condiționat de parametrii modelului curent și secvența observată.

Din definiții rezultă că:

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i,j)$$

Algoritmul Baum-Welch (III)



Secvența de operații necesară pentru calculul evenimentului mixt ca sistemul se află în starea S_i la momentul t și în starea S_j la momentul $t + 1$

$$\alpha_{t,i} = P(o_1, o_2, \dots, o_t, q_t = S_i | \lambda) \quad (7)$$

$$\beta_{t,i} = P(o_{t+1} o_{t+2} \cdots o_T | q_t = S_i, \lambda) \quad (8)$$

$$\begin{aligned} \xi_t(i,j) &= \frac{\alpha_{t,i} \cdot a_{i,j} \cdot b_j(o_{t+1}) \cdot \beta_{t+1,j}}{P(O|\lambda)} \\ &= \frac{\alpha_{t,i} \cdot a_{i,j} \cdot b_j(o_{t+1}) \cdot \beta_{t+1,j}}{\sum_{k=1}^N \sum_{l=1}^N \alpha_{t,k} \cdot a_{k,l} \cdot b_l(o_{t+1}) \cdot \beta_{t+1,l}} \end{aligned} \quad (9)$$

Algoritmul Baum-Welch (IV)

Cum ne ajută aceste variabile auxiliare?

$$\sum_{t=1}^{T-1} \gamma_t(i) = \text{numărul estimat de tranzitii din } S_i$$

$$\sum_{t=1}^{T-1} \xi_t(i,j) = \text{numărul estimat de tranzitii din } S_i \text{ la } S_j$$

Algoritmul Baum-Welch (V)

$\bar{\pi}_i = \text{nr. estimat de vizite în starea } S_i \text{ la momentul } (t = 1) = \gamma_1(i)$ (10)

Algoritmul Baum-Welch (V)

$$\bar{\pi}_i = \text{nr. estimat de vizite în starea } S_i \text{ la momentul } (t = 1) = \gamma_1(i) \quad (10)$$

$$\begin{aligned} \bar{a}_{i,j} &= \frac{\text{nr. estimat de tranziții din } S_i \text{ la } S_j}{\text{nr. estimat de tranziții din } S_i} \\ &= \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \end{aligned} \quad (11)$$

Algoritmul Baum-Welch (V)

$$\bar{\pi}_i = \text{nr. estimat de vizite în starea } S_i \text{ la momentul } (t=1) = \gamma_1(i) \quad (10)$$

$$\begin{aligned} \bar{a}_{i,j} &= \frac{\text{nr. estimat de tranziții din } S_i \text{ la } S_j}{\text{nr. estimat de tranziții din } S_i} \\ &= \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \end{aligned} \quad (11)$$

$$\begin{aligned} \bar{b}_{j,k} &= \frac{\text{nr. estimat de vizite în } S_j \text{ observând simbolul } v_k}{\text{nr. estimat de vizite în } S_j} \\ &= \frac{\sum_{t=1, O_t=v_k}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)} \end{aligned} \quad (12)$$

Algoritmul Baum-Welch (VI)

Algorithm 1 Algoritm Baum-Welch

```

1: intrări:  $O \leftarrow$  secvența de observații,  $\epsilon \leftarrow$  prag de convergență
2:
   {Initializare}
3: init. uniformă  $\Pi$  ( $\Pi_i = 1/N, 1 \leq i \leq N$ )
4: init. aleatoare  $a_{i,j}$ , a. î.  $\sum_{j=1}^N a_{i,j} = 1, \forall i = 1, N$ 
5: init. uniformă  $b_{j,k}$  ( $b_{j,k} = 1/M, 1 \leq j \leq N, 1 \leq k \leq M$ )
6: init  $\log(P(O|\bar{\lambda})) = 0$ 

7: repeat
8:    $\log(P(O|\lambda)) = \log(P(O|\bar{\lambda}))$ 
9:
   {E STEP - calculeaza variantele scalate pentru  $\alpha$  și  $\beta$  și probabilitatea curentă (log likelihood -  $\log(P(O|\bar{\lambda}))$ ) a secvenței observate}
10:   $[\log(P(O|\bar{\lambda})), \hat{\alpha}, \hat{\beta}, Scale] = forward\_backward(O, \Pi, A, B)$ 
11:
   {M STEP - recalculeaza update-uri pentru  $\Pi$ ,  $A$  și  $B$ }
12:   $\Pi = update\_pi\_procedure(\hat{\alpha}, \hat{\beta}, Scale)$ 
13:   $A = update\_A\_procedure(O, \hat{\alpha}, \hat{\beta}, Scale)$ 
14:   $B = update\_B\_procedure(O, \hat{\alpha}, \hat{\beta}, Scale)$ 
15: until  $\log(P(O|\bar{\lambda})) - \log(P(O|\lambda)) < \epsilon$ 

```

Algoritmul Baum-Welch (VI)

Algorithm 2 Algoritm Baum-Welch

```
1: Function update_pi_procedure( $\hat{\alpha}$ ,  $\hat{\beta}$ , Scale)
2:   for  $i = 1$  to  $N$  do
3:      $\Pi_i = \frac{\alpha_1(i) \cdot \beta_1(i) / Scale(1)}{\sum_{j=1}^N \alpha_1(j) \cdot \beta_1(j) / Scale(1)}$ 
4:   end for
5:   return  $\Pi$ 
6: EndFunction

7: Function update_A_procedure( $O$ ,  $\hat{\alpha}$ ,  $\hat{\beta}$ , Scale)
8:   for  $i = 1$  to  $N$  do
9:     for  $j = 1$  to  $N$  do
10:       $a_{i,j} = \frac{\sum_{t=1}^{T-1} \hat{\alpha}_{t,i} \cdot a_{i,j} \cdot b_j(o_{t+1}) \cdot \hat{\beta}_{t+1,j}}{\sum_{t=1}^{T-1} \sum_{j=1}^N \hat{\alpha}_{t,i} \cdot a_{i,j} \cdot b_j(o_{t+1}) \cdot \hat{\beta}_{t+1,j}}$ 
11:   end for
12: end for
13: return  $a$ 
14: EndFunction
```

Algoritmul Baum-Welch (VI)

Algorithm 3 Algoritm Baum-Welch

```
1: Function update_B_procedure( $O$ ,  $\hat{\alpha}$ ,  $\hat{\beta}$ , Scale)
2:   for  $j = 1$  to  $N$  do
3:     for  $k = 1$  to  $M$  do
4:        $b_{j,k} = \frac{\sum_{t=1, O(t)=v_k}^T \hat{\alpha}_t(j) \cdot \hat{\beta}_t(j) / Scale(t)}{\sum_{t=1}^T \hat{\alpha}_t(j) \cdot \hat{\beta}_t(j) / Scale(t)}$ 
5:     end for
6:   end for
7:   return  $b$ 
8: EndFunction
```

Baum-Welch - Să scriem niște cod

LET'S WRITE SOME CODE :-)

O aplicație simplă de recunoaștere a simbolurilor

Features:

O aplicație simplă de recunoaștere a simbolurilor

Features:

Definire

Definire, organizare și vizualizare a unui set de date de simboluri captate prin mișcarea mouse-ului.

O aplicație simplă de recunoaștere a simbolurilor

Features:

Definire

Definire, organizare și vizualizare a unui set de date de simboluri captate prin mișcarea mouse-ului.

Antrenare

Antrenarea unui motor de recunoaștere a simbolurilor prin metoda MMA.

O aplicație simplă de recunoaștere a simbolurilor

Features:

Definire

Definire, organizare și vizualizare a unui set de date de simboluri captate prin mișcarea mouse-ului.

Antrenare

Antrenarea unui motor de recunoaștere a simbolurilor prin metoda MMA.

Recunoaștere

Recunoașterea unui nou simbol și vizualizarea unor metriki de clasificare.

O aplicație simplă de recunoaștere a simbolurilor

Features:

Definire

Definire, organizare și vizualizare a unui set de date de simboluri captate prin mișcarea mouse-ului.

Antrenare

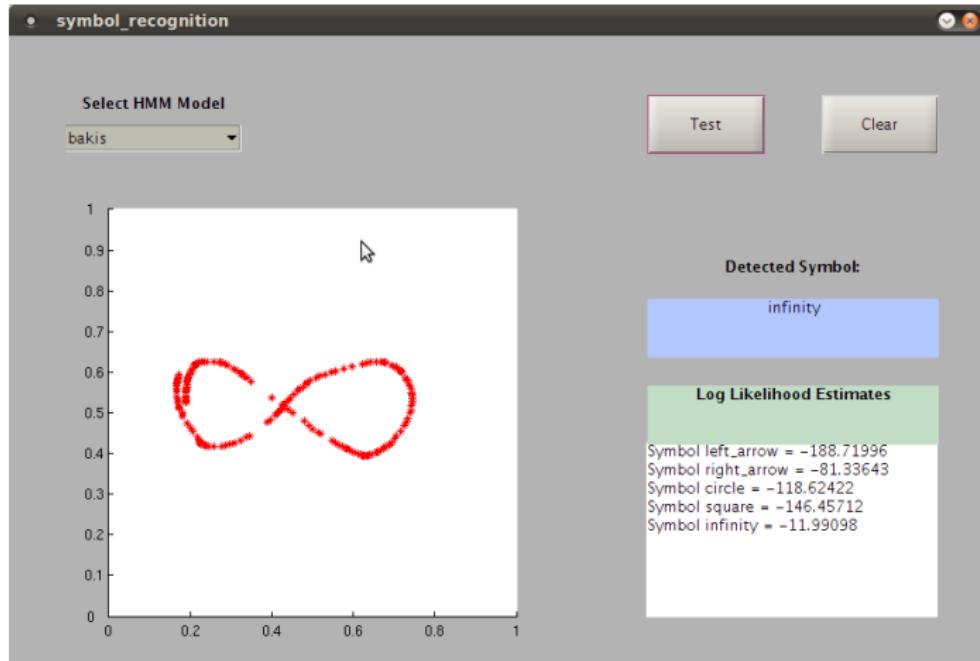
Antrenarea unui motor de recunoaștere a simbolurilor prin metoda MMA.

Recunoaștere

Recunoașterea unui nou simbol și vizualizarea unor metrii de clasificare.

Simboluri default incluse: **săgeată stânga, săgeată dreapta, cerc, pătrat, infinit**

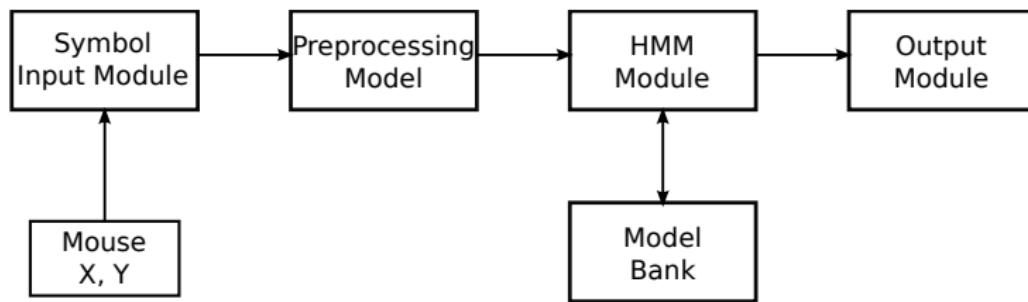
O aplicație simplă de recunoaștere a simbolurilor - Vizualizare



O vizualizare cu GUI-ul aplicației de recunoaștere a simbolurilor

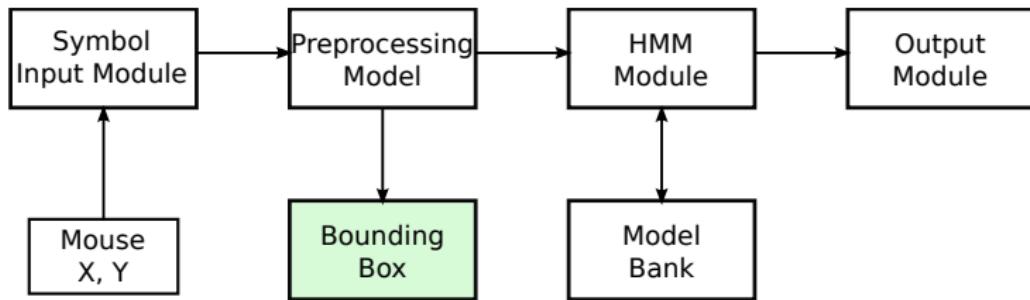
O aplicație simplă de recunoaștere a simbolurilor - Abordare (I)

Adapted from [YX94].



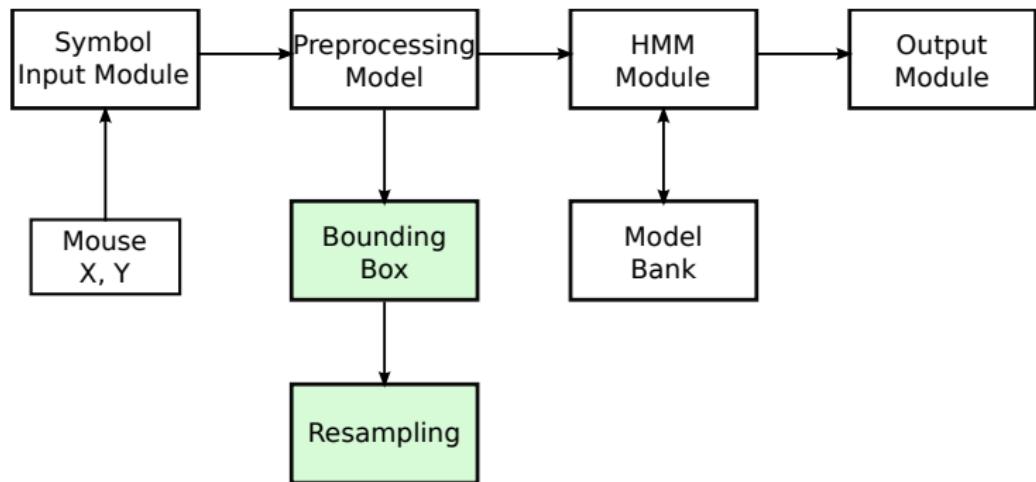
O aplicație simplă de recunoaștere a simbolurilor - Abordare (I)

Adapted from [YX94].



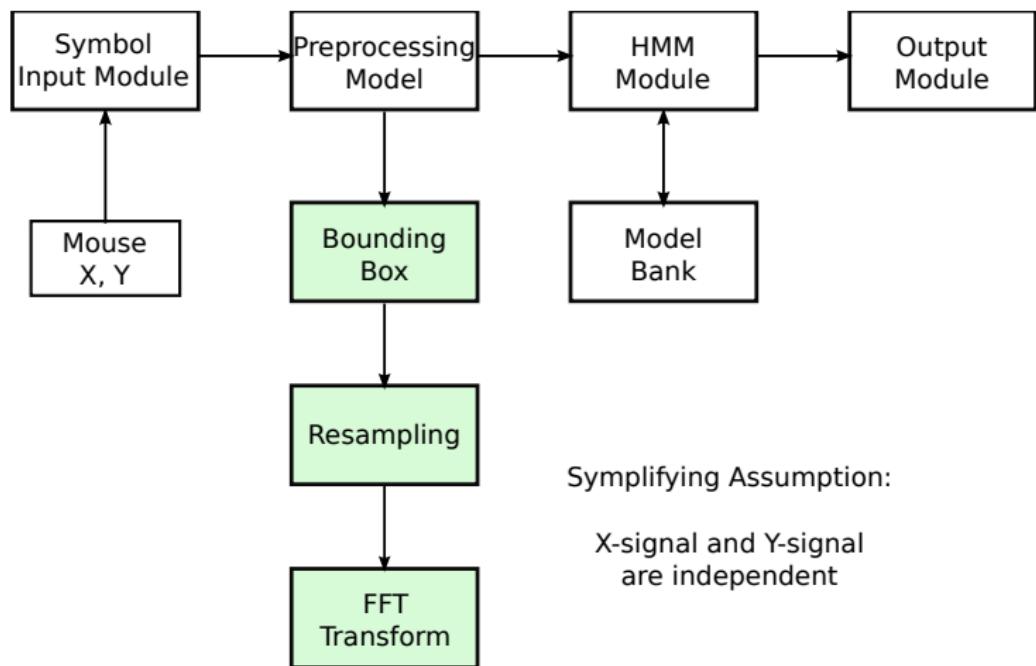
O aplicație simplă de recunoaștere a simbolurilor - Abordare (I)

Adapted from [YX94].



O aplicație simplă de recunoaștere a simbolurilor - Abordare (I)

Adapted from [YX94].



Simplifying Assumption:

X-signal and Y-signal
are independent

O aplicație simplă de recunoaștere a simbolurilor - Abordare (II)

Adaptare după [YX94].

Structura MMA

N (număr de stări) = 8

2 variabile observabile de tip discrete per stare - $\text{coef}_{FFT}(x)$, $\text{coef}_{FFT}(y)$

M (număr de valori pentru fiecare variabilă observabilă) = 256

Modelul de tranziție:

- Bakis
- Ergodic

O aplicație simplă de recunoaștere a simbolurilor - Abordare (III)

Procedură de recunoaștere:

- ① Construim și antrenăm câte un MMA pentru fiecare simbol în parte

O aplicație simplă de recunoaștere a simbolurilor - Abordare (III)

Procedură de recunoaștere:

- ① Construim și antrenăm câte un MMA pentru fiecare simbol în parte
- ② Folosim un set de date (simboluri) de validare pentru a stabili *praguri* de recunoaștere, i.e. dacă probabilitatea secvenței observate este prea mică pentru fiecare MMA, marcăm simbolul drept *necunoscut*

O aplicație simplă de recunoaștere a simbolurilor - Abordare (III)

Procedură de recunoaștere:

- ① Construim și antrenăm câte un MMA pentru fiecare simbol în parte
- ② Folosim un set de date (simboluri) de validare pentru a stabili *praguri* de recunoaștere, i.e. dacă probabilitatea secvenței observate este prea mică pentru fiecare MMA, marcăm simbolul drept *necunoscut*
- ③ Recunoaștere:
 - Calculăm $P(O|\lambda_i)$ pentru fiecare MMA construit pentru simbolurile $i = 1, \dots, nr_simboluri$
 - Alegem $\max(P(O|\lambda_i))$ ca și simbol candidat. Dacă $P(O|\lambda_i) > prag_i$, atunci am recunoscut *simbolul i*, altfel marcăm *necunoscut*

O aplicație simplă de recunoaștere a simbolurilor - Rezultate

Mărime set de date

**5 simboluri: săgeată stânga, săgeată dreapta, cerc, pătrat, infinit
100 exemple per simbol: 50 antrenare, 10 validare, 40 testare**

```
>> symbol_performance_test('ergodic')
----- Testing trained HMM models -----
## Results for the model of symbol "left_arrow":
Accuracy: 0.97500
Precision: 1.00000
Recall: 0.97500
Confusion matrix line: 39 0 1 0 0 0

## Results for the model of symbol "right_arrow":
Accuracy: 1.00000
Precision: 1.00000
Recall: 1.00000
Confusion matrix line: 0 40 0 0 0 0

## Results for the model of symbol "circle":
Accuracy: 0.90244
Precision: 0.97368
Recall: 0.92500
Confusion matrix line: 0 0 37 2 1 0

## Results for the model of symbol "square":
Accuracy: 0.95238
Precision: 0.95238
Recall: 1.00000
Confusion matrix line: 0 0 0 40 0 0

## Results for the model of symbol "infinity":
Accuracy: 0.97561
Precision: 0.97561
Recall: 1.00000
Confusion matrix line: 0 0 0 0 40 0
```

```
>> symbol_performance_test('bakis')
----- Testing trained HMM models -----
## Results for the model of symbol "left_arrow":
Accuracy: 0.90000
Precision: 1.00000
Recall: 0.90000
Confusion matrix line: 36 0 1 0 0 3

## Results for the model of symbol "right_arrow":
Accuracy: 1.00000
Precision: 1.00000
Recall: 1.00000
Confusion matrix line: 0 40 0 0 0 0

## Results for the model of symbol "circle":
Accuracy: 0.97561
Precision: 0.97561
Recall: 1.00000
Confusion matrix line: 0 0 40 0 0 0

## Results for the model of symbol "square":
Accuracy: 0.97500
Precision: 1.00000
Recall: 0.97500
Confusion matrix line: 0 0 0 39 0 1

## Results for the model of symbol "infinity":
Accuracy: 1.00000
Precision: 1.00000
Recall: 1.00000
Confusion matrix line: 0 0 0 0 40 0
```

:)

- [DLR77] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38, 1977.
- [KF09] D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press, 2009.
- [Rab89] L.R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
- [YX94] J. Yang and Y. Xu. Hidden markov model for gesture recognition. Technical report, DTIC Document, 1994.
- [Zub06] R. Zubek. Introduction to hidden markov models. *AI Game Programming Wisdom*, 3:633–646, 2006.

Thank you!

Baftă, șoalor!