מודלים מרקוביים חבויים

Hidden Markov Models (HMM)

מודלים סטטיסטיים לרצפים

רצף (sequence): סידרה בזמן של תצפיות.

מודל סטוכסטי של רצפים: מודל סטוכסטי הינו תהליך אקראי אשר בכל זמן t $t = 1,2,\ldots,T$ נמצא במצב כלשהו מתוך קבוצת מצבים $\{s_i\}$ $i=1,\ldots,n$.

תהליך חבוי: התצפיות שאנו רואים אינם המצב שבו התהליך נמצא, אלא מ"מ אשר התפלגותו הינה פונקציה של המצב.

נתאר את המצב שבו נמצא התהליך בזמן t ע"י הסימון q_t.

ננתח את ההסתברות לרצף מצבים $q_1,q_2,\ldots,q_T$:

$$\Pr(q_1,q_2,\ldots,q_T) = \Pr(q_1) * \Pr(q_2|q_1) * \ldots * \Pr(q_T|q_1,\ldots,q_{T-1})$$

נניח הנחה מרקובית מסדר 1: ההסתברות שתהליך יעבור בזמן $t+1$ למצב s_i תלויה אך ורק במצב בו נמצאים בזמן t.

כלומר: $\Pr(q_{t+1}|q_1,\ldots,q_t) = \Pr(q_{t+1}|q_t)$

מודלים סטטיסטיים לרצפים (2)

ולכן ההסתברות לרצף מצבים $q_1,q_2,\ldots,q_T$:

$$Pr(q_1,q_2,\ldots,q_T) = Pr(q_1) * Pr(q_2|q_1) * Pr(q_3|q_2) * \ldots * Pr(q_T|q_{T-1})$$

כלומר תהליך מרקובי מסדר ראשון מוגדר באופן מלא ע"י הסתברויות התחלה בכל אחד מן המצבים, והסתברויות מעבר בין המצבים.

<u>מודל מרקובי סטציונרי</u>: הסתברות המעבר ממצב i למצב j אינה תלויה בזמן:

$$Pr(q_t = s_j \mid q_{t-1} = s_i) = Pr(q_r = s_j \mid q_{r-1} = s_i) = a_{ij}$$

נתאר את התהליך ע"י וקטור הסתברויות התחלה Π, ומטריצת הסתברויות מעבר A.

<u>נגדיר רצף תצפיות $o_1,o_2,\ldots,o_T$:</u>

בזמן t התהליך נמצא במצב q_t ופולט תצפית o_t.

מודלים מרקובים חבויים

הנחה: התצפית הנפלטת בזמן t תלויה (סטוכסטית) אך ורק במצב q_t
לא ב:

זמן t, תצפיות בזמנים שונים מ-t, מצב בזמנים שונים מ-t.

הגדרה פורמלית:

$$\Pr(x = o_t \mid q_1,\ldots,q_T, o_1,o_2,\ldots,o_{t-1}) = \Pr(x = o_t \mid q_t = k) = b_k(o_t)$$

נתאר את אופן יצירת רצף תצפיות ע"י HMM:

1. בחר מצב התחלתי עפ"י התפלגות $\Pi \rightarrow q_1$, t=1.
2. פלוט תצפית O_t עפ"י ההתפלגות $\Pr(O_t \mid q_t)$.
3. בחר את המצב הבא ע"י ההתפלגות $\Pr(s_i \mid s_j) \rightarrow q_t$, t=t+1.
4. אם $t < T$ חזור לשלב 2.
5. סיים.

מודלים מרקובים חבויים

דוגמא:

states = 2 (0,1)

$\Pi_1 = 1.0$ $\Pi_2 = 0.0$

$a_{11} = 0.9$ $a_{12} = 0.1$

$a_{21} = 0.5$ $a_{22} = 0.5$

Observation space = {+,-}

$b_0(+) = 0.9$ $b_0(-) = 0.1$

$b_1(+) = 0.1$ $b_1(-) = 0.9$

for example:

O = ++-++++-----+-+----+++-+-+++++

שלושת הבעיות הבסיסיות של HMM

1. איך נחשב את הסתברות הפליטה של רצף תצפיות בהינתן מודל HMM:

$Pr(X = o_1,o_2,\ldots,o_T \mid \Pi , A , b)$

2. איך נמצא את רצף המצבים המסתבר ביותר בהינתן מודל ורצף תצפיות:

$\mathrm{argmax}_Q\ Pr(Q = q_1,q_2,\ldots,q_T \mid \Pi , A , b, X= o_1,o_2,\ldots,o_T)$

3. איך נאמן את המודל – איך נמצא את ערכי Π , A , b עפ"י אוסף של רצפי תצפיות :

$\mathrm{argmax}_{\Pi , A , b}\ Pr(\Pi , A , b \mid O)$

פתרון ישיר לבעיה מס' 1: חישוב $\Pr(X \mid M = \Pi, A, b)$

$$\Pr(X \mid M) = \sum_Q \Pr(X,Q \mid M) = \sum_Q \Pr(X \mid Q, M) \Pr(Q \mid M)$$

ניעזר במשוואות:

$$\Pr(Q \mid M) = \Pi_{q_1} a_{q_1,q_2} a_{q_2,q_3} * \ldots * a_{q_{T-1},q_T}$$

$$\Pr(X \mid Q,M) = b_{q_1}(o_1) * \ldots * b_{q_T}(o_T)$$

לפיכך:

$$\Pr(X \mid M) = \sum_Q \Pi_{q_1} a_{q_1,q_2} a_{q_2,q_3} * \ldots * a_{q_{T-1},q_T} b_{q_1}(o_1) * \ldots * b_{q_T}(o_T)$$

הבעיה:

זמן החישוב לינארי במספר רצפי המצבים האפשרי - $2Tn^T$.

פתרון יעיל לבעיה מס' 1: חישוב $\Pr(X \mid M = \Pi, A, b)$

נגדיר פונקצית עזר $\alpha(t,s_i)$:

$\alpha(t,s_i) = \Pr(X = o_1,o_2,\ldots,o_t,\ q_t = s_i \mid M)$

<u>המשמעות:</u>

ההסתברות לכך שהתהליך יפלוט את תת-רצף התצפיות $o_1,o_2,\ldots,o_t$ ויגיע בזמן t למצב s_i.

<u>מוטיבציה:</u>

מערכי $\alpha(t,s_i)$ ניתן לחשב את $\Pr(X|M)$:

$\Pr(X|M) = \sum_{i=1,\ldots,n} \alpha(T,s_i)$

<u>נחשב את ערכי $\alpha(t,s_i)$ בתכנות דינאמי:</u>

<u>Initialize:</u> $\alpha(1,s_i) = \Pi_i\, b_i(o_1) \quad 1 \le i \le n$

<u>Recurrence:</u> $\alpha(t+1,s_j) = [\sum_{i=1,\ldots,n} \alpha(t,s_i) * a_{ij}] * b_j(o_{t+1}) \quad 1 \le t \le T-1,\ 1 \le j \le n$

פתרון יעיל לבעיה מס' 1: חישוב $Pr(X|M = \pi, A, b)$

<u>סיבוכיות:</u>

בטבלה $\alpha(t,s_j)$ יש Tn ערכים, כל ערך מחושב ב- $O(n)$ פעולות, סה"כ $O(Tn^2)$.

<u>עבור מטריצת מעברים A דלילה (טופולוגית מצבים שמאל-לימין ללא דילוגים):</u>

בטבלה $\alpha(t,s_j)$ יש Tn ערכים, כל ערך מחושב ב- $O(1)$ פעולות, סה"כ $O(Tn)$.

<u>זכרון:</u>

אנו זקוקים לשורה האחרונה של α בלבד, וכל שורה ניתן לחשב עפ"י השורה שלפניה, לפיכך יש צורך ב- $O(n)$ זכרון בלבד.

פתרון לבעיה מס' 2: מציאת רצף המצבים המסתבר ביותר

$\text{argmax}_Q \Pr(Q = q_1,q_2,\ldots,q_T \mid M, X= o_1,o_2,\ldots,o_T) = ???$

עפ"י חוק Bayes:

$$\Pr(Q = q_1,q_2,\ldots,q_T \mid M, X= o_1,o_2,\ldots,o_T) = \frac{\Pr(X= o_1,o_2,\ldots,o_T \mid M, Q = q_1,q_2,\ldots,q_T) * \Pr(Q = q_1,q_2,\ldots,q_T \mid M)}{\Pr(X= o_1,o_2,\ldots,o_T \mid M)}$$

הגודל $\Pr(X= o_1,o_2,\ldots,o_T \mid M)$ אינו תלוי ב-Q ולפיכך אפשר להוציאו מן המיקסום ונקבל:

$$\text{argmax}_Q \{\Pr(X= o_1,o_2,\ldots,o_T \mid M, Q = q_1,q_2,\ldots,q_T) * \Pr(Q = q_1,q_2,\ldots,q_T \mid M)\}$$

עפ"י הנחות HMM (מודל מרקובי מסדר ראשון וסטציונריות):

(I) $\Pr(X= o_1,o_2,\ldots,o_T \mid M, Q = q_1,q_2,\ldots,q_T) =$

$$\Pr(X= o_1,o_2,\ldots,o_{T-1} \mid M, Q = q_1,q_2,\ldots,q_{T-1}) * \Pr(X=o_T \mid M, q_T)$$

(II) $\Pr(Q = q_1,q_2,\ldots,q_T \mid M) = \Pr(Q = q_1,q_2,\ldots,q_{T-1} \mid M) * \Pr(q_T \mid M, q_{T-1})$

פתרון לבעיה מס' 2: מציאת רצף המצבים המסתבר ביותר

נציב את I,II בביטוי שאותו אנו ממקסמים:

$$\text{argmax}_Q \{Pr(X= o_1,o_2,\ldots,o_T \mid \Pi , A , b, Q = q_1,q_2,\ldots,q_T) * Pr(Q = q_1,q_2,\ldots,q_T \mid M)\} =$$

$$\text{argmax}_{Q,q_T} \left\{ \begin{array}{l} Pr(X= o_1,o_2,\ldots,o_{T-1} \mid M, Q = q_1,q_2,\ldots,q_{T-1}) \; * \\ Pr(Q = q_1,q_2,\ldots,q_{T-1} \mid M) * \\ Pr(X=o_T \mid M, q_T) \; * Pr(q_T \mid M, q_{T-1}) \end{array} \right\}$$

$$= \text{argmax}_{Q,q_T} \left\{ \begin{array}{l} Pr(X= o_1,o_2,\ldots,o_{T-1}, Q = q_1,q_2,\ldots,q_{T-1} \mid M) \; * \\ Pr(X=o_T \mid M, q_T) \; * Pr(q_T \mid M, q_{T-1}) \end{array} \right\}$$

נסתכל כעת על כל רצפי המצבים באורך t המסתיימים במצב s_i. נגדיר את ההסתברות המשותפת

של רצף המצבים המסתבר ביותר ורצף התצפיות $o_1,o_2,\ldots,o_t$ מתוך כלל רצפי המצבים

המסתיימים במצב s_i כ- $\delta_t(i)$:

$$\delta_t(i) = \max_{q_1,q_2,\ldots,q_{t-1}} Pr(X=o_1,o_2,\ldots,o_t , Q=q_1,q_2,\ldots,q_{t-1} ,q_t = s_i \mid M)$$

פתרון לבעיה מס' 2: מציאת רצף המצבים המסתבר ביותר

<u>טענה</u>:

$\mathrm{argmax}_Q \{Pr(X= o_1,o_2,\ldots,o_T \mid \Pi , A , b, Q = q_1,q_2,\ldots,q_T) * Pr(Q = q_1,q_2,\ldots,q_T \mid M)\} =$

$\mathrm{argmax}_{\ i}\ \delta_T(i)$

<u>הוכחה</u>: נובע מהגדרת $\delta_T(i)$.

כל שנותר הוא למצוא אלגוריתם יעיל לחישוב ערכי $\delta_t(i)$.

נחשב את ערכי $\delta_t(i)$ ע"י תכנות דינאמי:

Initialize: $\delta_1(i) = \Pi_i\, b_i(o_1)$

Recurrence: $\delta_t(j) = \max_{\ i} \{\delta_{t-1}(i)*a_{ij}\}*\, b_j(o_t)$, $1 \le j \le n$

בעת בניית טבלת $\delta_t(j)$ נשמור בטבלה נפרדת מעקב על "מאיזה מצב בזמן t-1 הגענו למצב j בזמן t".

באמצעות טבלת מעקב זו, נוכל לשחזר את רצף המצבים שעבורו מיקסמנו את $\delta_T(i)$.

פתרון לבעיה מס' 2: מציאת רצף המצבים המסתבר ביותר

סיבוכיות:

גודל טבלה $\delta_t(i)$ הינו nT. כל ערך מחושב ב- O(n), סה"כ $O(n^2T)$

עבור מטריצת מעברים A דלילה (טופולוגית מצבים שמאל-לימין):

בטבלה $\delta_t(i)$ nT ערכים, כל ערך מחושב ב- O(1) פעולות, סה"כ O(nT).

זכרון:

עבור $\delta_t(i)$ אנו זקוקים לשורה האחרונה בלבד, וכל שורה ניתן לחשב עפ"י השורה שלפניה,

לפיכך יש צורך ב- O(n) זכרון בלבד.

עבור טבלת המעקב שממנה מחלצים את רצף המצבים הטוב ביותר, יש צורך לשמור טבלה מלאה

בגודל O(nT).

סה"כ: O(nT).

אלגוריתם זה מכונה אלגוריתם Viterbi.

חזרה לבעיה מס' 1: חישוב $Pr(X| M = \Pi , A , b)$

$$Pr(X \mid M) = \sum_Q Pr(X,Q \mid M) = \sum_Q Pr(X \mid Q, M)\, Pr(Q \mid M)$$

ניתן לקרב את הסכום על כל רצפי המצבים ע"י רצף המצבים המסתבר ביותר.

את הרצף הזה נחשב ע“י אלגוריתם Viterbi.

היתרון: הסיבוכיות התיאורטית נשארת זהה, אבל מעשית האלגוריתם מהיר בהרבה – נרצה לחשב

את $\log(Pr(X|M))$ כדי למנוע overflow, אבל האלגוריתם המדויק מכיל חיבורים.

אלגוריתם ה-Viterbi אינו מכיל חיבורים אלא פונקצית מקסימום.

האלגוריתם:

1. מצא את רצף המצבים המסתבר ביותר עפ"י אלגוריתם Viterbi.
2. הסתברות רצף התצפיות מתוך רצף המצבים המסתבר ביותר מתקבל כתוצאת לוואי של אלגוריתם ה-Viterbi.

פתרון לבעיה מס' 3: אימון מודל HMM

$$\arg\max_{\Pi , A , b} \Pr(\Pi , A , b \mid O) = ???$$

<u>הגדרת הבעיה:</u> בהינתן אוסף של רצפי תצפיות, היינו רוצים לשערך את פרמטרי מודל ה-HMM

באופן אופטימלי (Maximum Likelihood).

<u>פתרון:</u> EM.

<u>מוטיבציה:</u> אם היינו יודעים מהם רצפי המצבים ה"נכונים" המתאימים לרצפי התצפיות, הפתרון

הוא פשוט:

Π_i = number of times the HMM was in state S_i at time 1

number of observation sequences

$a_{i,j}$ = number of times the transition S_i to S_j was used

number of transitions starting from S_i

$b_i(x) \leftarrow$ parametric/non-parametric estimation according to observations emitted in state i.

פתרון לבעיה מס' 3: אימון מודל HMM

אפשרות א': אלגוריתם אימון Baum-Welch

1. אתחול: אתחל בצורה כלשהי את פרמטרי מודל ה-HMM.

2. Expectation:

שערך עבור כל רצף תצפיות r, כל זמן t, וכל מצב i,j את ההסתברות להיות בזמן t במצב i עבור רצף התצפיות ה-r:

$\gamma_{r,t}(i) := \Pr\{q_t(r)=s_i \mid M , o_{r,1},\ldots,o_{r,Tr}\}$

ואת ההסתברות להיות בזמן t במצב i ובזמן t+1 במצב j עבור רצף התצפיות ה-r:

$\xi_{r,t}(i,j) := \Pr\{q_t(r)=s_i, q_{t+1}(r)=s_j \mid M , o_{r,1},\ldots,o_{r,Tr}\}$

השערוך יתבצע ע"י נוסחאות Forward-Backward (תכנות דינאמי).

3. Maximization:

שערך את פרמטרי ה-HMM עפ"י ההסתברויות $\gamma_{r,t}(i)$, $\xi_{r,t}(i,j)$.

4. חזור ל-2 עד להתכנסות.

פתרון לבעיה מס' 3: אימון מודל HMM

איך נשערך את פרמטרי המודל מתוך ההסתברויות $\gamma_{r,t}(i)$, $\xi_{r,t}(i,j)$?

$$\Pi_i = \frac{\sum_{r=1}^{\#\,seq.} \gamma_{r,1}(i)}{\#\,seq.}$$

$$a_{i,j} = \frac{\sum_{r=1}^{\#\,seq.} \sum_{t=1}^{len(r)-1} \xi_{r,t}(i,j)}{\sum_{r=1}^{\#\,seq.} \sum_{t=1}^{len(r)-1} \gamma_{r,t}(i)}$$

$b_i(x) \leftarrow$ parametric/non-parametric estimation according to all observations, weighted by their probability of being emitted by state i - $\gamma_{r,t}(i)$.

פתרון לבעיה מס' 3: אימון מודל HMM

נוסחאות Forward-Backward

כזכור הגדרנו פונקצית עזר $\alpha(t,s_i)$:

$\alpha(t,s_i) = \Pr(X = o_1,o_2,\ldots,o_t,q_t= s_i \mid M)$

המשמעות: ההסתברות לכך שהתהליך יפלוט את תת-רצף התצפיות $o_1,o_2,\ldots,o_t$ ויגיע בזמן t למצב s_i.

נרחיב את ההגדרה עבור קבוצת דוגמאות אימון:

$\alpha_r(t,s_i) = \Pr(X = o_{r,1},o_{r,2},\ldots,o_{r,t}, q_{r,t}= s_i \mid M)$

נגדיר בצורה מקבילה פונקצית עזר $\beta_r(t,s_i)$:

$\beta_r(t,s_i) = \Pr(X = o_{r,t+1},o_{r,t+2},\ldots,o_{r,T_r}, q_{r,t}= s_i \mid M)$

המשמעות: ההסתברות לכך שהתהליך יפלוט את רצף התצפיות $o_{r,t+1},o_{r,t+2},\ldots,o_{r,T_r}$ בזמנים $t+1,\ldots,T_r$

בהינתן שידוע שבזמן t התהליך היה במצב s_i.

את $\beta_r(t,s_i)$ נחשב באמצעות תכנות דינאמי בצורה דומה וביעילות זהה ל- $\alpha_r(t,s_i)$.

פתרון לבעיה מס' 3: אימון מודל HMM

$$\xi_{r,t}(i,j) = \Pr\{q_t(r)=s_i,\ q_{t+1}(r)=s_j \mid M,\ o_{r,1},\ldots,o_{r,Tr}\} =$$

$$= \frac{\Pr\{q_t(r)=s_i,\ q_{t+1}(r)=s_j,\ o_{r,1},\ldots,o_{r,Tr} \mid M\}}{\Pr(o_{r,1},\ldots,o_{r,Tr} \mid M)} = \frac{\alpha_r(t,s_i)\ a_{i,j}\ b_j(o_{r,t+1})\ \beta_r(t+1,s_j)}{\Pr(o_{r,1},\ldots,o_{r,Tr} \mid M)}$$

את α, β ראינו כיצד מחשבים.

חישוב $\Pr(o_{r,1},\ldots,o_{r,Tr} \mid M)$ הוא בעיה מס' 1 ואותה כבר פתרנו.

לפיכך, פתרנו את בעיית חישוב ξ.

נחשב את γ באופן הבא:

$$\gamma_{r,t}(i) = \Pr\{q_t(r)=s_i \mid M,\ o_{r,1},\ldots,o_{r,Tr}\} =$$

$$= \sum_{j=1}^{n} \xi_{r,t}(i,j)$$

פתרון לבעיה מס' 3: אימון מודל HMM

<u>אפשרות ב': אלגוריתם אימון Viterbi</u>

1. אתחול: אתחל בצורה כלשהי את פרמטרי מודל ה-HMM.
2. <u>Expectation</u>:

   מצא עבור כל רצף תצפיות את רצף המצבים המסתבר ביותר ע"י אלגוריתם Viterbi (בהינתן מודל ה-HMM הנוכחי).
3. <u>Maximization</u>:

   שערך את פרמטרי ה-HMM עפ"י רצפי המצבים המסתברים ביותר.
4. חזור ל-2 עד להתכנסות.

<u>אתחול פרמטרי מודל HMM (עבור Baum-Welch / Viterbi)</u>

1. E: התאם לכל רצף תצפיות, רצף מצבים הגיוני ככל האפשר.
2. M: בצע איטרצית שיערוך של פרמטרי ה-HMM.

זיהוי דיבור באמצעות
מודלים מרקוביים חבויים

הגדרת הבעיה

נתון

סידרת דגימות של אות אקוסטי

פלט רצוי

רצף המילים שנאמר.

חילוץ מאפיינים

חלוקת האות למסגרות (10 מילי-שניות).

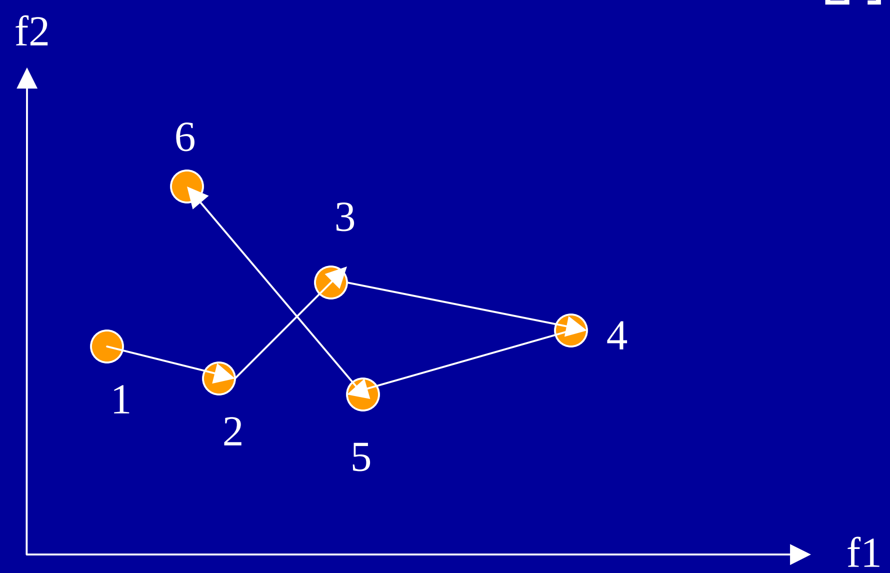
חישוב וקטור Mel-Cepstrum לכל מסגרת.

הרחבת ווקטור המאפיינים של כל מסגרת ע"י נגזרת ראשונה (ושנייה).

התוצאה: סידרת ווקטורי מאפיינים.

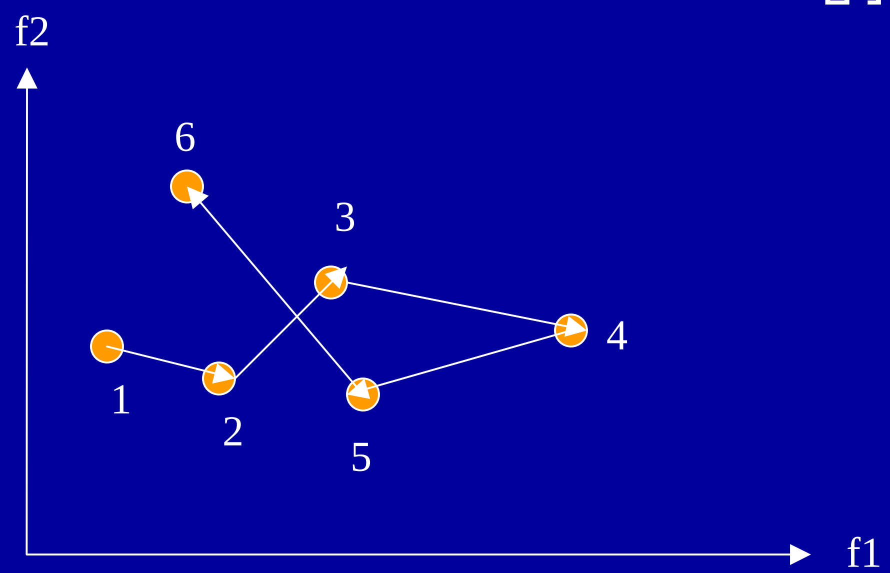
מסלול במרחב המאפיינים

כל אמירה היא מסלול במרחב המאפיינים



מסלול במרחב המאפיינים

כל אמירה היא מסלול במרחב המאפיינים



אקראיות המסלול במרחב המאפיינים

אותה מילה המבוטאת פעמיים לא תתאים בדיוק לאותו מסלול במרחב המאפיינים:

- דובר שונה (אופן היגוי, Pitch, פורמנטים, קצב).
- שונות של אותו דובר
- רעש, ערוץ

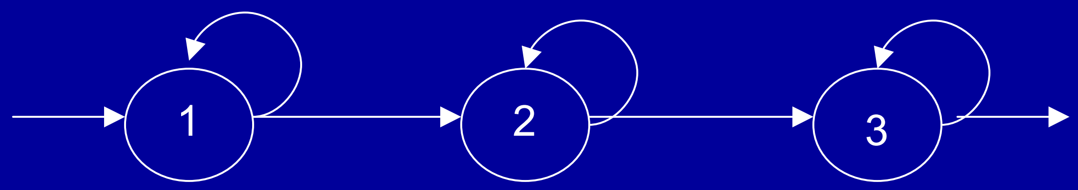
האקראיות היא גם בציר הזמן וגם במרחב המאפיינים.

נמדל את תהליך הדיבור ע"י HMM.

מידול דיבור ע"י HMM

כל פונמה אפשר לתאר כתהליך העונה לרוב ההגדרות של HMM.

נתאר כל פונמה ע"י תהליך המורכב משלושה מצבים: תחילה-אמצע-סוף.



הסתברויות הפליטה ימודלו ע"י תערובות גאוסיינים

(8-64 גאוסיינים לתערובת, תלוי בגודל קבוצת האימון).

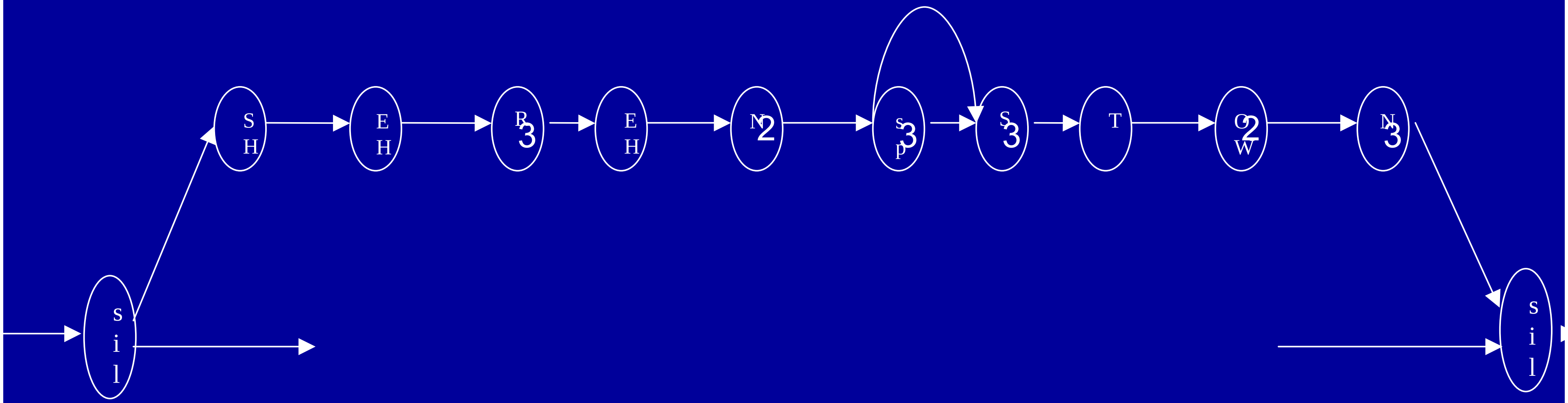
מילים ימודלו ע"י שירשור המצבים של הפונמות.

משפטים ימודלו ע"י שירשור מילים.

זיהוי שמות

נמדל כל שם ע"י שירשור של פונמות:

Sharon Stone : SH EH R_ EH N_ [sp] S_ T_ OW N



זיהוי משפטים – דקדוק BNF

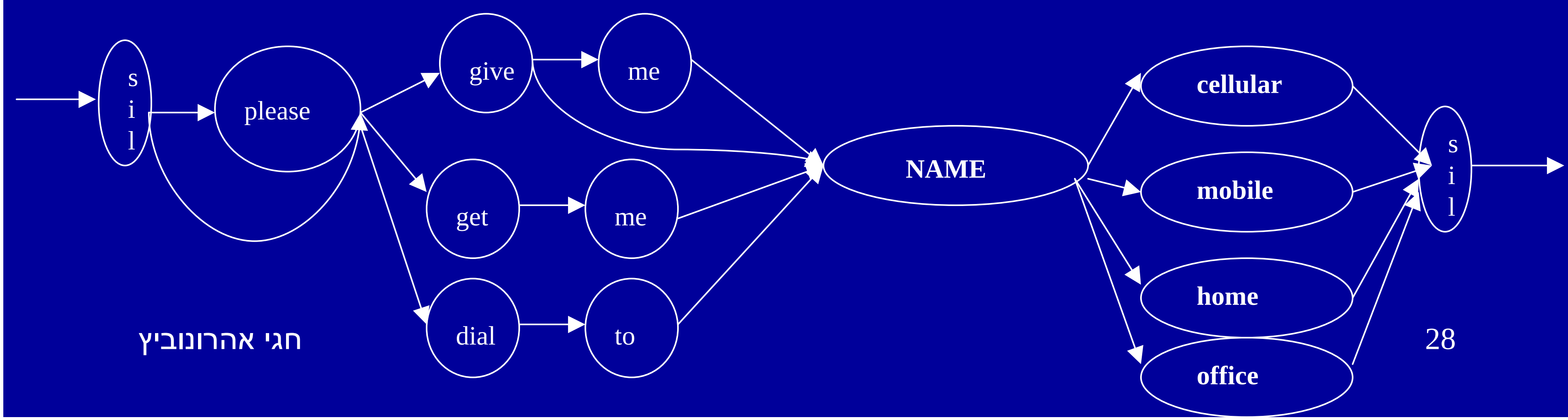
נתאר את הדקדוק בדקדוק חסר הקשר:

SENTENCE → sil [please] VERB NAME [cellular | mobile | home | office] sil

VERB → give me | get [me] | dial to

NAME → Sharon Stone | Harrison Ford | …

בניית הדקדוק תיעשה ע"י מומחה או בצורה אוטומטית מתוך דוגמאות של משפטים תקינים.



זיהוי טקסט חופשי

אוצר מילים: עשרות אלפים.

אפשרות א: ללא מודל שפה:

SENTENCE = WORD*

WORD = word-1 | word-2 | … | word-30,000

אפשרות ב: מודל שפה uni-gram (1-gram).

אפשרות ג: מודל שפה bi-gram (2-gram).

אפשרות ד: מודל שפה tri-gram (3-gram) – לא מודל מרקובי מסדר ראשון.

אפשרות ה: מודל שפה מורכב יותר – לא משתלב עם אלגוריתמי ה-HMM.

בכל מקרה, נארגן את המילון באמצעות עץ.

פונמות תלויות הקשר

הבחנה: אותה פונמה נשמעת אחרת עם היא נמצאת בהקשר פונטי אחר.

מסקנה: במקום להשתמש בפונמות, נרצה להשתמש בפונמות תלויות הקשר – בדרך-כלל tri-phones.

בעיות: אין מספיק data לאימון כל ה-triphones.

פתרון: אישכול של ה-triphones לקבוצות.

אחוז זיהוי כפונקציה של גודל מילון

ככל שגודל המילון גדול יותר, יש יותר אפשרויות להתבלבל בין מילים, ולכן אחוז הזיהוי יורד.

נניח עבור גודל מילון 1,000 אחוז הזיהוי הוא 90%.

מה יהיה אחוז הזיהוי עבור מילון בגודל 100 מילים? 10 מילים?

תשובה:

לא ניתן לדעת ללא בדיקה אמפירית.

מודל 1: ב-10% מן האמירות יש רעש חזק / הגייה לא ברורה / וכו', ולכן אחוז הזיהוי יהיה 90% ללא תלות בגודל המילון.

מודל 2: כל זוג מילים מן המילון הוא או "מבלבל" או "לא-מבלבל". הסיכוי לזוג להיות מבלבל הוא p. לכן עבור 100 מילים אחוז הזיהוי יהיה כ-99%, ועבור 10 מילים אחוז הזיהוי יהיה כ-99.9%.

המציאות: שילוב בין שני המודלים.

תוצאות אמפיריות (עבור שמות ערים באנגלית):

100 – 99%   250 – 95%   500 – 90%

אחוז זיהוי כפונקציה של אורך המילה

ככל שהמילה ארוכה יותר, יש יותר אינפורמציה דיסקרמינטיבית, ולכן אחוז הזיהוי עולה.

לדוגמה:

"על" לעומת "אנציקלופדיה".

שם-פרטי לעומת שם-מלא.

אחוז זיהוי כפונקציה של גמישות הדקדוק

מה מספר המשפטים השונים החוקיים?

האם הסתברות המשפטים השונים זהה?

באמצעות תורת האינפורמציה אפשר לחשב את תוחלת מספר המילים אשר ניתן לומר לאחר סיום

מילה שרירותית (preplexity), וגודל זה משפיע על אחוז הזיהוי.

לדוגמא: מילון בגודל 2 מילים, ההסתברות לאמירת המילה הראשונה היא 99% -> אחוז הזיהוי

יהיה גבוה יותר מאשר אם ההסתברות הייתה 50%.

לדוגמא:

מערכת הכתבה (30,000 מילים) תיכשל לחלוטין ללא אדפטציה לדובר, ואחרי אדפטציה תגיע

ל-90%.

מערכת לזיהוי מילים מבודדות עם 30,000 תשיג ביצועים נחותים בהרבה, כי אין מודל שפה

שמוריד את ה-preplexity.

אחוז זיהוי כפונקציה של עוצמת הרעש

ככל שעוצמת הרעש גבוהה יותר, אחוז הזיהוי יורד.

ניתן להקטין את הרגישות ע"י שימוש בטכניקות שונות, למשל אימון בתנאי רעש.

תוצאות: (100 שמות אנשים)

חדר שקט: 97%

רכב עומד: 90%

רכב נוסע ב-90 קמ"ש: 60%

אחוז זיהוי כפונקציה של אמצעי ההקלטה

מיקרופון קרוב ב-90 קמ"ש: 80%

דיבורית ב-90 קמ"ש: 60%.

שילוב זיהוי דיבור במוצרים מסחריים

שלב א': בחינת כדאיות

1. מערכות זיהוי דיבור אינן מזהות באופן מושלם – האם יש הסכמה למערכת לא מושלמת?
2. האם יש fall-back ?
3. מהו פרופיל השימוש: גודל מילון, דובר קבוע? רעש?
4. מהן החלופות?
5. כדאיות כלכלית? (זיכרון, CPU, עלות פיתוח/תגמולים)
6. **מהו הסיכוי לעמידה במעטפת הביצועים?**

שלב ב': בדיקת היתכנות

1. פיתוח עצמי / התבססות על shareware (כמו HTK) / התבססות על צד ג'.
2. בניית אב-טיפוס ובדיקת מעטפת הביצועים.

שלב ג': פיתוח המוצר

דוגמאות לפרוייקטים (מוצלחים)

1. זיהוי דובר של Persay ("הישיר הראשון").
2. אימות זיהות של חברת Nuance (בנקאות).
3. חיוג אוטומטי מרכב.
4. מערכות הכתבה (Microsoft, IBM, L&H).
5. הפניית שיחות טלפון (AT&T).
6. 144 (Phonetic Systems).
7. סינון שיחות טלפון למטרות ביון.
8. גלישה קולית באינטרנט.
9. גישה למאגרי מידע קוליים Voice-Portals.