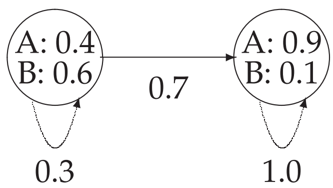
3.2 Skryté Markovove modely

Doposiaľ najflexibilnejší a najúspešnejší prístup v oblasi rozpoznávania rečových signálov sú skryté Markovove modely (HMM). V tejto sekcii je popísaný základný princíp HMM, ako aj algoritmy pre ich trénovanie a používanie. Skryté Markovove modely používajú štatistické modelovanie spracovávaného signálu. Variabilita rečového signálu, ktorá sa dá pomerne dobre vystihnút' pomocou teórie pravdepodobnosti, predurčuje HMM na použitie v tejto oblasti.

Skrytý Markovov model je určený množinou stavov, pravdepodobnosťami prechodou medzi nimi a pravdepodobnosťami generovania výstupných symbolov v jednotlivých stavoch, ako ukazuje obrázok (3.2). Začína sa v počiatočnom stave a v každom časovom kroku nastáva prechod do nového stavu, pričom je vygenerovaný jeden výstupný symbol. Prechody ako aj generovanie výstupných symbolov je náhodné, určené svojou funkciou hustoty pravdepodobnosti. HMM si možno predstavit' ako „čiernu skrinku", ktorá generuje postupnost' výstupných symbolov. Postupnost' stavov, cez ktoré systém v čase prechádza, zostáva pre pozorovateľa skrýtá - z toho je odvodený názov skryté Markovove modely.



Obr. 3.3: Príklad jednoduchého HMM

HMM pozostáva z nasledovných elementov:

- $\{s\}$ - množina stavov
- $\{a_{ij}\}$ - množina pravdepodobností prechodov, kde a_{ij} je pravdepodobnost' prechodu zo stavu i do stavu j.

- $\{\pi_i\}$ - množina pravdepodobností počiatočného stavu. π_i vyjadruje pravdepodobnosť s ktorou daný model začne v stave i.

- $\{b_i(u)\}$ - množina pravdepodobností generovania výstupných symbolov, kde $b_i(u)$ je pravdepodobnosť vygenerovania symbolu u v stave i.

Pre úplnnosť zaveďme ešte označenie λ pre súbor parametrov Markovovho modelu.

$$\lambda = (a, b, \pi) \tag{3.2}$$

Keďže a, b a π sú pravdepodobnosti, musia spĺňať tieto vlastnosti:

$$\sum_i \pi_i = 1 \quad \forall i \tag{3.3}$$

$$a_{ij} \geq 0 \qquad b_i(u) \geq 0 \quad \forall i, j, n \tag{3.4}$$

$$\sum_j a_{ij} = 1 \quad \forall i \qquad \sum_u b_i(u) = 1 \quad \forall i \tag{3.5}$$

Ďalej budeme predpokladať, že uvedený HMM je prvého rádu, inými slovami povedané, pravdepodobnosť nasledujúceho stavu závisí len na aktuálnom stave a nezávisí od sekvencie predchádzajúcich stavov.

Modelovanie pomocou HMM sa nezaobíde bez zvládnutia troch základných problémov:

- **Stanovenie pravdepodobnasti pozorovania**, používa sa pri rozpoznávaní izolovaných slov

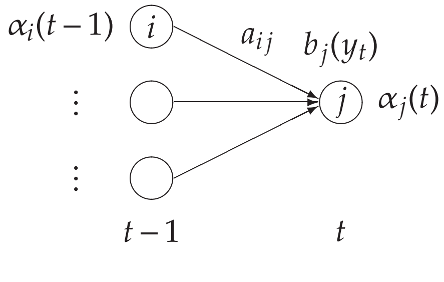
- **Určenie postupnosti stavov**, cez ktorý model prechádzal pri generovaní výstupnej postupnosti

- **Trénovanie modelu**, používané na nastavenie parametrov HMM

3.2.1 Stanovenie pravdepodobnasti pozorovania

Pri rozpoznávaní izolovaných slov musíme byt' schopný určit' pravdepodobnost' generovania pozorovanej postupnosti na daných HMM, ktoré popisujú referenčné slová. Porovnaním týchto pravdepodobností je možné určit' model, na ktorom bola pozorovaná postupnost' vygenerovaná (model s najväčšou pravdepodobnosťou).

Predpokladajme, že HMM je určený hodnotami $\{a_{ij}\}, \{b_i(u)\}, \{\pi_i\}$,. Vyjadríme pravdepodobnost', že bola vygenerovaná postupnost' $y_1^T = (y_1, y_2, \ldots, y_T)$. Pretože v každom stave i môže byt' vygenerovaný výstupný symbol u s pravdepodobnosťou $b_i(u)$, každá postupnost' stavov dĺžky T prispieva k celkovej pravdepodobnmosti. Pomocou „hrubej sily“ by bolo možné nájst' všetky možné postupnosti stavov dĺžky T a spočítat' ich pravdepodobnosti vygenerovania y_1^T, čo by však nebolo vďaka svojej výpočtovej náročnosti praktické. Omnoho efektívnejšie riešenie poskytuje *dopredná procedúra*. Definujme $\alpha_j(t)$ ako pravdepodobnost' generovania časti postupnosti y_1^t, ktorej posledný symbol bol vygenerovaný v stave j a čase t. Na začiatku je $\alpha_i(t = 0)$ inicializované na hodnoty $\pi_i \cdot b_i(y_1)\ \forall i$. Ak poznáme $\alpha_i(t-1)$ pre všetky stavy i v predchádzajúcom časovom intervale $t-1$, potom $\alpha_j(t)$ môže byt' určené rekurzívne, ako suma pravdepodobností prechodu do stavu j zo všetkých možných stavov i, pričom bol vygenerovaný výstupný symbol y_t (obr. 3.4).



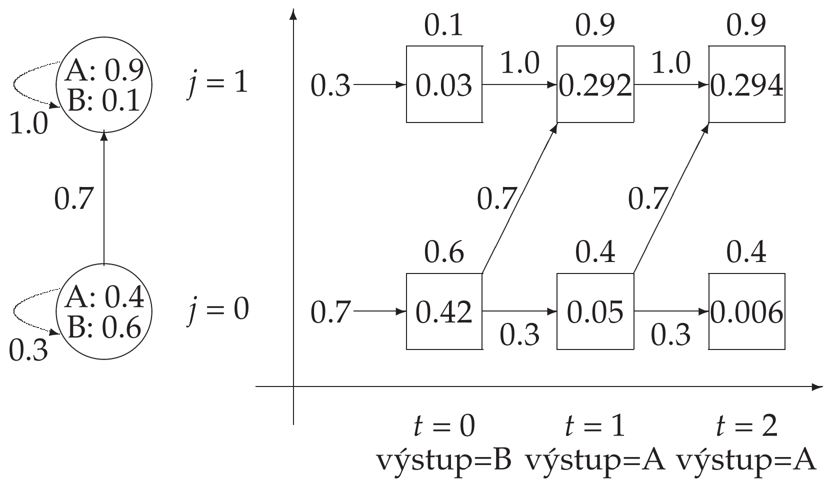
Obr. 3.4: Ilustrácia doprednej procedúry

$$\alpha_j(t) = \sum_i \alpha_i(t-1)a_{ij}b_j(y_t) \tag{3.6}$$

Ak k je množina konečných stav, potom pravdepodobnosť vygenerovania celej postupnosti y_1^T daným HMM je:

$$\mathrm{P}(y_1^T|\lambda) = \sum_k \alpha_k(T) \tag{3.7}$$

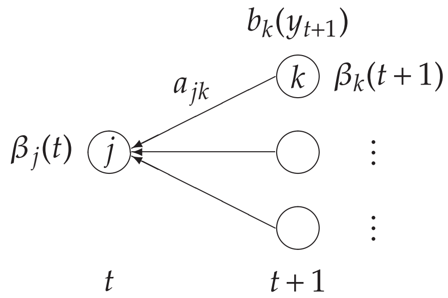
Obrázok (3.5) ilustruje príklad, v ktorom HMM z obrázku (3.2) vygeneroval vystupnú postupnosť $y_1^3 = (B, A, A)$, pričom počiatočná pravdepodobnosť $\pi = \{0.7, 0.3\}$. Bunka na pozícii (t, j) obsahuje hodnotu $\alpha_j(t)$. Hodnota pravdepodobnosti vygenerovania danej postupnosti je v našom prípade $\mathrm{P}(y_1^3 = (B, A, A) \mid \lambda) = 0.30024$.



Obr. 3.5: Ilustrácia doprednej („forward") procedúry na príklade

V predchádzjúcom texte sme definovali $\alpha_j(t)$ ako pravdepodobnosť vygenerovania časti postupnostî y_1^t, ktorej posledný symbol bol vygenerovaný v stave j a čase t. Teraz definujeme jej zrkadlový obraz $\beta_j(t)$ ako pravdepodobnosť vygenerovania zvyšku postupnosti $y_{t+1}^T = (y_{t+1}, y_{t+2}, \ldots, y_T)$ a toho, že v čase t bol model v stave j, $\alpha_j(t)$ sa nazýva dopredný („forward") člen a $\beta_j(t)$ spätný („backward") člen. Podobne ako v prípade $\alpha_j(t)$ aj $\beta_j(t)$ sa počíta rekurzívne (obr. 3.6).

$$\beta_j(t) = \sum_k a_{jk} b_k(y_{t+1}) \beta_k(t+1) \tag{3.8}$$



Obr. 3.6: Ilustrácia spätnej procedúry

Rekurzia začína v čase T nastavením $\beta_k(T)$ na 1.0 pre konečné stavy a 0.0 pre všetky ostatné stavy. Ak i reprezentuje množinu počiatočných stavov, potom výsledná pravdepodobnosť pozorovania je potom daná vzťahom:

$$\mathrm{P}(y_1^T|\lambda) = \sum_i \pi_i b_i(y_1) \beta_i(1) \tag{3.9}$$

3.2.2 Určenie postupnosti stavov

Ako už bolo spomenuté, dopredná procedúra je vhodná na rozpoznávanie izolovaných slov. Rozpoznávanie súvislej reči vyžaduje použitie odlišného prítupu, pretože by bolo dosť nepraktické mať rozdielne HMM pre každú možnú vetu. V tomto prípade je potrebné zistiť postupnosť stavov, ktorá vygenerovala výstupnú postupnosť. Z postupností stavov je následne možné určiť postupnosť slov. Nanešťastie, aktuálna postupnosť stavov je z definície HMM skrytá a nemôže byť jednoznačne určená, pretože každá možná cesta množinou stavov mohla vygenerovať pozorovanú postupnosť s určitou pravdepodobnosťou. Najlepšie čo môžeme v danej situácii urobiť, je určiť najpravdepodobnejšiu sekvenciu stavov, ktorá vygenerovala pozorovanú postupnosť. Tak ako pri riešení prvého problému je aj tu možné použiť „hrubú silu“, vypočítať pravdepodobnosti všetkých možných postupností stavov a určiť tú najpravdepodobnejšiu. Rozumné riešenie poskytuje *Viterbiho algoritmus*, ktorý je veľmi podobný *doprednej procedúre*.

$$V_j(t) = MAX_i[v_i(t-1)a_{ij}b_j(y_t)] \tag{3.10}$$

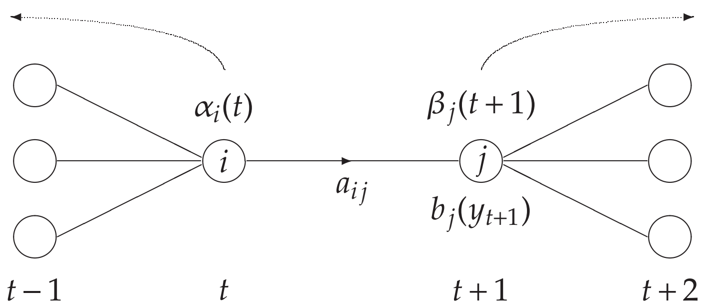
3.2.3 Trénovanie modelu

Pri ténovaní HMM je najväčším problémom nastavenie parametrov modelu tak, aby sa maximalizovala pravdepodobnosť generovania danej postupnosti na požadovanom modeli. Zatiaľ neexistuje jeho analytické riešenie a musí sa preto použiť iteratívny spôsob známy ako *Baum-Welch algoritmus*, ktorý stručne popíšeme.

Definujme $\gamma_{ij}(t)$ ako pravdepodobnosť prechodu zo stavu i do stavu j v čase t a vygenerovania celej výstupnej postupnosti y_1^T.

$$\gamma_{ij}(t) = \mathrm{P}(i_t \to j \,|\, y_1^T) = \frac{\alpha_i(t)a_{ij}b_j(y_{t+1})\beta_j(t+1)}{\sum_k \alpha_k(T)} \tag{3.11}$$

Vyznam čitateľa ilustruje obrázok (3.7). Menovateľ odzrkadľuje fakt, že posledný symbol y_1^T môže byť vygenerovaný v k konečných stavoch modelu.



Obr. 3.7: Ilustrácia výpočtu $\gamma_{ij}(t)$ pomocou „forward-backward“ algoritmu

Definujme teraz $N(i \to j)$ ako očakávaný počet prechodov zo stavu i do stavu j v čase 1 až T:

$$N(i \rightarrow j) = \sum_t \gamma_{ij}(t) \tag{3.12}$$

Sčítaním uvedených členov cez všetky cieľové stavy j, dostaneme $N(i \rightarrow *)$ čo predstavuje očakávaný počet výskytov modelu v stave i v čase 1 až T:

$$N(i) = N(i \rightarrow *) = \sum_t \sum_j \gamma_{ij}(t) \tag{3.13}$$

Vybraním len tých stavov i, v ktorých došlo k vygenerovaniu výstupného symbolu u definujeme $N(i,u)$:

$$N(i,u) = \sum_{t:(y_t=u)} \sum_j \gamma_{ij}(t) \tag{3.14}$$

Parameter a_{ij} (pravdepodobnosť prechodu zo stavu i do stavu j) je možné určiť ako podiel priemerného počtu prechodov zo stavu i do stavu j a priemerného počtu výskytov modelu v stave i v čase 1 až T. Pre odhad nového parametra $\overline{a_{ij}}$ platí vzťah:

$$\overline{a_{ij}} = \mathrm{P}(i \rightarrow j) = \frac{N(i \rightarrow j)}{N(i)} = \frac{\sum_t \gamma_{ij}(t)}{\sum_t \sum_j \gamma_{ij}(t)} \tag{3.15}$$

Podobne možno určiť parameter $b_i(u)$, ako podiel počtu výskytov symbolu u v stave i a počtu výskytov modelu v stave i v čase 1 až T. Pre odhad nového parametra $\overline{b_i(u)}$ platí vzťah:

$$\overline{b_i(u)} = \mathrm{P}(u|i) = \frac{N(i,u)}{N(i)} = \frac{\sum_{t:(y_t=u)} \sum_j \gamma_{ij}(t)}{\sum_t \sum_j \gamma_{ij}(t)} \tag{3.16}$$

Dá sa dokázať, že substitúciou $\{\overline{a}, \overline{b}\}$ za $\{a, b\}$ sa $\mathrm{P}(y_1^T)$ zvyšuje až do dosiahnutia lokálneho maxima, čím opakovanie tejto procedúry zaručí optimalizáciu HMM pre trénovacie dáta.