

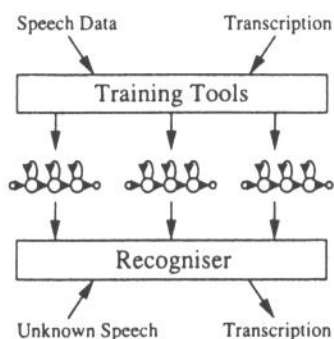
# Beszédinformációs rendszerek

/gyakorlat – vázlat/

## Gépi beszéd felismerés

### Mintaillesztés rejtett Markov-modellekkel

- A beszéd felismerés alapfeladata. (hang  $\rightarrow$  szöveg konverzió)
- Statisztikus alapú beszéd felismerés:
  - Tanítás
  - Felismerés
  - Modellek (akusztikus, kiejtési, nyelvtani)



- A beszéd felismerés 2 fő lépése:
  - Lényegkiemelés vagy parametrizálás
  - Felismerés vagy mintaillesztés

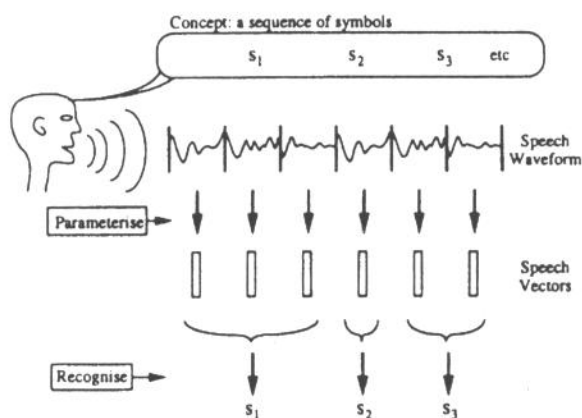


Fig. 1.1 Message Encoding/Decoding

- Parametrizálás: FFT alapú (spektrum számítás), illesztés a halláshoz (lényegkiemelés)
- Mintaillesztés: HMM-mel, másképp fogalmazva: beszédhang-modellek összekapcsolásából készült gráf segítségével.

- Izolált szavas mintaillesztés HMM-mel (elm.)

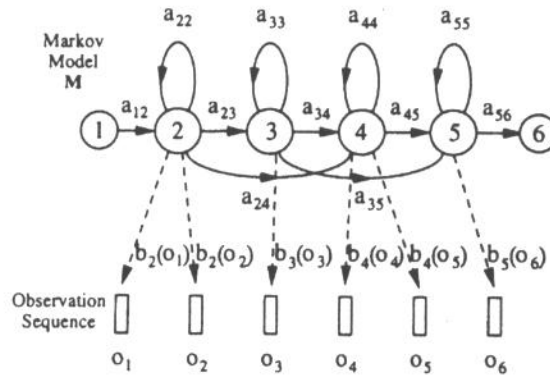


Fig. 1.3 The Markov Generation Model

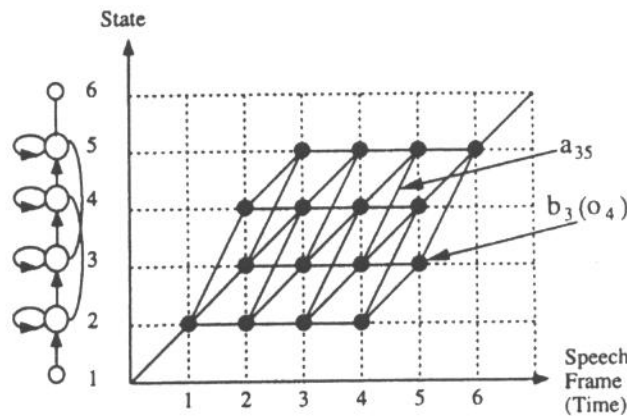


Fig. 1.6 The Viterbi Algorithm for Isolated Word Recognition

Definiáljuk a  $\delta_t(i)$  mennyiséget, amely a kezdőtől az  $i$  állapotig, az első  $t$  db megfigyelés során számított legjobb út valószínűségét jelenti. Ennek segítségével az algoritmus (a modell-állapotok száma  $N$ , a tulajdonságvektoroké  $T$ ):

- Inicializálás:  $\delta_t(i) = 1$ ;  $t = 0$ -nál minden kezdőállapotra
- Indukciós lépés:  $\delta_{t+1}(j) = \max_i [\delta_t(i) \cdot a_{ij}] \cdot b_j(\mathbf{o}_{t+1})$ ;  
minden  $i = 1 \dots N$  és  $j = 1 \dots N$ -re.  $t = 1 \dots T$
- Leállítás: ha  $t = T$ .

A mintaillesztés (felismerés) eredménye a végpontok közötti legnagyobb valószínűségű út. Ez adja meg a beszédvektorokról a szövegre történő leképezést.

- 1 v. 2 szemléltető példa
  - Két szavas felismerés Viterbi-algoritmussal. Használt modell: sima “balról-jobbra” CD-HMM 1 Gaussal. (Jellemzővektorok 1 dimenziósak.)
  - Több szavas felismerés, ragozási gráf-szerű HMM-mel. Csak a lehetséges útvonalak követése a trellis-ben, az algoritmus lényegének szemléltetése.

- DTW-s és HMM-es beszédfelismerő élő bemutatása.
  - Voice-Dial
  - az NTT-BUTE városnév felismerő.

Ajánlott irodalom:

<http://htk.eng.cam.ac.uk/>