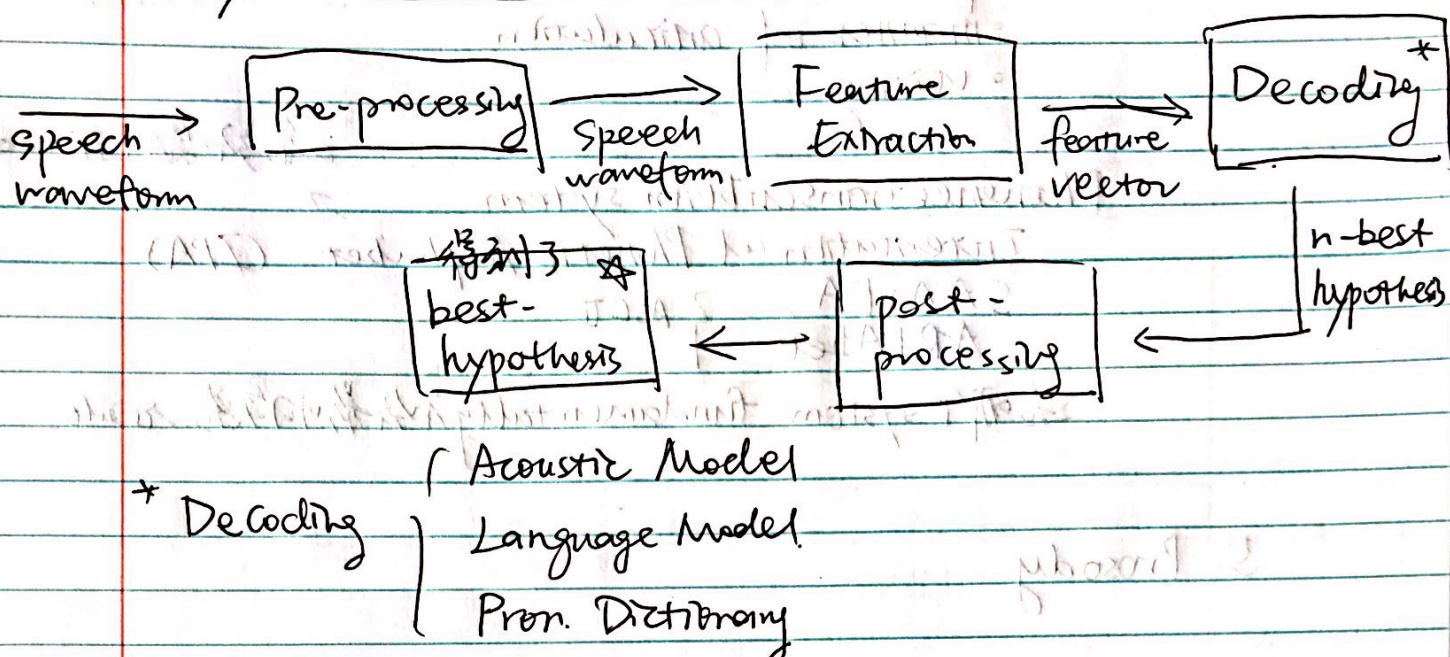


# Automatic Speech Recognition (ASR)

## 1. SR System Architecture!



2. { Discrete Prob. Distribution: random experiments  
Continuous Prob. Distribution:  $\int_{-\infty}^{\infty} p(x) dx = 1$ .

↓

通常用 mixture of Gaussians 来 approximate.

$$\text{density. } p(x) = \sum_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_i} \exp \left[ -\frac{(x-\mu_i)^2}{2\sigma_i^2} \right]$$

$\mu_i, \sigma_i$ : mean, variance of the  $i$ -th Gaussian.

## 3. HMM.

algorithm.  
trained with Baum-Welch (forward-backward)  
→ there is no analytical method to determine HMM parameters  
B-W is an iterative estimation, monotonously improves  
para. set  $\Phi = (A, B)$ .

(MAP).  
→ Speaker adaptation: Maximum a posteriori adaptation  
MAP estimate to substitute HMM's.



#### 4. Phonemes

- ~~Categories~~ Characterized by the way they are produced.
  - place of articulation
  - manner of articulation
  - voicing

- Phonemes transcription system

International Phonetic Alphabet (IPA)

SAMPA, } ASCII  
ARPAbet }

另外是ASCII encoded

选哪个 system fundamentally 不会影响实验 results.

#### 5. Prosody