Each phone (or other acoustic unit) is an HMM with a number of states. It follows all words, and sentences are HMMs as well, since they are blocks of words. We also apply...

Common Practices

State Transition

Parameter Update

Parameter estimation for parameter re-estimation. In (8), the posterior probability of state \( s_t \) at time \( t \) is often the indicator function of the event that \( s_t \) be the indicator function of the event that \( s_t \) is the state at time \( t \), and \( s_t \) is a state transition between two successive states. The expectation value of the total number of transitions from state \( i \) to state \( j \) is often

\[
E[D] = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t)
\]

Using (7) in (2), we have

\[
D = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t)
\]

Parameter Update

The parameter \( p_{ij}(t) \) is updated according to

\[
\lambda_{ij} = \frac{1}{|A|} \sum_{t=1}^{N} p_{ij}(t)
\]

The model complexity is often measured in terms of the total number of parameters. Suppose \( |A| = n \), the minimal number of parameters is \( n \). The complexity is therefore a function of the number of states and parameters. From the conditional independence assumption of HMMs, we have

\[
\log p(x) = \sum_{t=1}^{N} \log p(x_t | x_{t-1}) = \sum_{t=1}^{N} \log p(x_t | x_{t-1}) = \sum_{t=1}^{N} \log p(x_t | x_{t-1}) = \sum_{t=1}^{N} \log p(x_t | x_{t-1})
\]

Taking log-priors, we have

\[
I = \sum_{t=1}^{N} \log \frac{p(x_t | x_{t-1})}{p(x_t)} = \sum_{t=1}^{N} \log \frac{p(x_t | x_{t-1})}{p(x_t)} = \sum_{t=1}^{N} \log \frac{p(x_t | x_{t-1})}{p(x_t)} = \sum_{t=1}^{N} \log \frac{p(x_t | x_{t-1})}{p(x_t)}
\]

The parameter \( p_{ij}(t) \) is updated according to

\[
\lambda_{ij} = \frac{1}{|A|} \sum_{t=1}^{N} p_{ij}(t)
\]

For parameter re-estimation. In (8), the posterior probability of state \( s_t \) at time \( t \) is often

\[
\sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t)
\]

Using (7) in (2), we have

\[
D = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t) = \sum_{t=1}^{N} \sum_{i}^{N} \sum_{j}^{N} p_{ij}(t)
\]

Parameter Update

The parameter \( p_{ij}(t) \) is updated according to

\[
\lambda_{ij} = \frac{1}{|A|} \sum_{t=1}^{N} p_{ij}(t)
\]

The model complexity is often measured in terms of the total number of parameters. Suppose \( |A| = n \), the minimal number of parameters is \( n \). The complexity is therefore a function of the number of states and parameters. From the conditional independence assumption of HMMs, we have

\[
\log p(x) = \sum_{t=1}^{N} \log p(x_t | x_{t-1}) = \sum_{t=1}^{N} \log p(x_t | x_{t-1}) = \sum_{t=1}^{N} \log p(x_t | x_{t-1}) = \sum_{t=1}^{N} \log p(x_t | x_{t-1})
\]

Taking log-priors, we have

\[
I = \sum_{t=1}^{N} \log \frac{p(x_t | x_{t-1})}{p(x_t)} = \sum_{t=1}^{N} \log \frac{p(x_t | x_{t-1})}{p(x_t)} = \sum_{t=1}^{N} \log \frac{p(x_t | x_{t-1})}{p(x_t)} = \sum_{t=1}^{N} \log \frac{p(x_t | x_{t-1})}{p(x_t)}
\]