# Acoustic Model Format

The structure of acoustic model is simple:

- only triphone context is supported
- all phones have same number of states
- senones are shared in triphones with the same phones
- filler phones are context-independent
- transition matrices are usually same for same base phone

The acoustic model is usually a folder with the following contents. Some of the files could be missing:

- `feat.params` – feature extraction parameters, a list of options used to configure feature extractoion.
- `mdef` – the definition of mapping between the triphone contexts to GMM ids (senones)
- `means` – gaussian codebook means
- `variances` – gaussian codebook variances
- `mixture_weights` – mixtures for gaussians (could be missing if sendump is present)
- `sendump` – compressed and quantized mixtures (could replace mixture_weights)
- `feature_transform` – feature transformation matrix
- `noisedict` – the dictionary for filler words
- `transition_matrices`– HMM transition matrices

Binary files usually consist of the header which points to the number of streams and data dimensions and then the raw float data. Last value is usually a checksum. They arrays are stored sequentially, float by float.

The arrays like means, mixture weights or variances usually have multiple dimensions. For means it is feature stream id, then for example gaussian id, then the vector of means. For mixture weights first index is stream id, then senone id, then mixture weights for each gaussian for the senone. Sendump file contains quantized and processed mixture weights. One can use printp tool from sphinxtrain to convert this binary representation to text readable format.

Mdef file is a text file listing the mapping from the triphone context to state senone ids.