

Senior Design Project Description

We would like to apply a log-linear model based on those found in machine translation systems to the problem of creating a harmony for a given melody. The log-linear model will consist of two main components: the language model and the translation model. The language model will model the likelihood of a given harmony line $H = \{h_0, h_1, \dots, h_{l-1}\}$ where $l = |H|$. We propose it use 3-grams for context and apply a penalty function d to discourage large differences in notes' positions on the staff. More formally, the language model can be defined as:

$$p_{LM}(H) = \prod_{i=1}^l P[h_i|h_{i-1}, h_{i-2}, h_{i-3}] * d(h_i - h_{i-1})$$

The second component of our model will be the translation model, which will model the probability that a harmony H is a good translation for a melody M . It can be defined as:

$$p_{TM}(M|H) = \prod_{i=1}^l P[m_i|h_i]$$

Putting these two components together, given a melody M we would like to find a harmony H s.t.

$$H = \operatorname{argmax}_H p(H|M) = p_{TM}(M|H)p_{LM}(H).$$

To simplify computation, we will actually minimize the log of the function above so that

$$\begin{aligned} H &= \operatorname{argmax}_H \log(p_{TM}(M|H)) + \log(p_{LM}(H)) \\ &= \operatorname{argmax}_H \sum_{i=1}^l \log(P[m_i|h_i]) + \log(P[h_i|h_{i-1}, h_{i-2}, h_{i-3}]) + \log(d(h_i - h_{i-1})) \end{aligned}$$

The problem of melody to harmony translation is somewhat simpler than the standard problem of translating between two natural languages because there is no concept of reordering in the former problem. Based on how we have chosen to define the problem, a harmony note corresponding to some melody note will always occur in the same position as the melody note. Additionally, our harmonies and melodies will be of the same length, so no reordering or length penalties need be applied.

In order to minimize the function above, we plan to build up the potential translations $H^j = \{h_0, \dots, h_j\}$ for $M^j = \{m_0, \dots, m_j\}$ where $0 \leq j \leq l$ and prune the potential translations by some threshold value for each increment of j .