Combined Phone set

What is my task?

My task is to obtain combined phone set.

- Suppose I have three languages Hindi, Tamil and English. Each language have its own phone-set.
- I want to train a single neural network with these three data sources with a common phone-set.

Related Work?

- In the paper by **Kanishka Rao -** they trained a **neural network with multiple data sources** and improved WER on each data source.
- They trained the neural network with a **common character-set** in the loss function.

What is their Innovation?

They had an **accent specific phone-set** which acted as **soft-guide** and they optimized loss function on the **common character set**. **They jointly learned**:

- feature vector to accent specific phoneme mapping
- accent specific phoneme to grapheme mapping

What is the difference in their setup and our setup:

- They got it successfully working with multiple data sources because the character set is common (all accents are for english language).

What are our Challenges and Innovations:

we cannot use their approach directly!!

- They had one language and multiple accents. We have multiple languages. Hence we cannot train with multiple data sources directly with a common phone-set.
- We need a **many to one mapping** from **language specific phone-set to a language independent phone-set**.

How to get this many to one phone-mapping?

Answer: We should use some clustering approach to cluster phones from different languages together. Clustering approaches known to me are:

- K-means clustering
- AHC or Tree based clustering

Either use sequence of vectors of phone or perform embedding:

- **Embedding approaches**: RNN based classification, I-vector, X-vector.

To-do list

- [ ] Choose clustering approach
- [ ] Choose approach to get embedding

Sub-pages

Clustering approach

Embedding approach