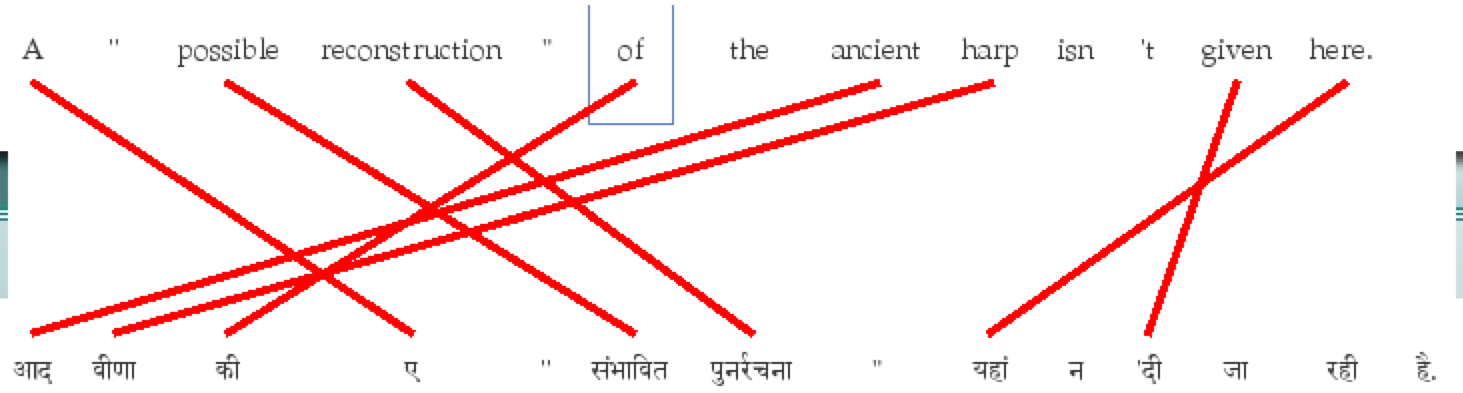
A " possible reconstruction " of the ancient harp isn 't given here.
आद् वीणा की ए " संभावित पुनर्रचना " यहां न 'दी जा रही है.

A Set of Annotation Interfaces for Alignment of Parallel Corpora

Anil Kumar Singh

IIT (BHU), Varanasi, India

Motivation and Goal

- Aligned parallel annotated data a precious resource
- From sentence alignment to treebank alignment
  - Interfaces needed: For manual alignment from scratch or for correction of auto-alignments
- Building a single system for all the steps
- Unified by a common representation scheme
  - Shakti Standard Format (SSF)
- Easy to use, flexible and extensible

Implemented System

- Sentence alignment interface
- Word/phrase/group alignment interface
  - Aligned shallow parsed corpus
- Parallel syntactic annotation alignment interface
  - Based on a facility-rich syntactic annotation interface
  - Aligned treebank: Work in progress
- Parallel corpus markup interface
  - Different from the rest of the interfaces in the set
  - Does not use SSF: Uses sentence level offsets