NEW DELHI MATHEMATICAL ASSOCIATION (SMA)

A fuzzy matching algorithm for open offline machine translation

Shayon Sengupta

January 15, 2012

1 ABSTRACT

We implement a translation memory database and match the following neural network with it: a 50 node first layer, 25 node second layer and 30 node output layer with 10 neurons for linguistic optimization. This algorithm accounts for all grammatical nuances, and classifies them into the second layer, and further, upon reaching a nuance that has not been previously identified, it adds it to the markov chain to handle for future translations in $O(n^2)$ time.
The algorithm is given by:

$$P(w_i^2) = \prod_{i=1}^{L} P(w_i | w_{i-n+1}^{i-1})$$

This is a variant of the SOUL MT algorithm, however, when coupled with the fuzzy variant :

$$k' = argmax\left\{\sum_{m}^{M} i^2 \lambda_m h_m(k, f)\right\}$$

I haven't had access to a data set to test this yet. But I'd like to.