

# A future for universal grapheme-phoneme transduction modeling with neuralized finite-state transducers

**Chu-Cheng Lin**

Johns Hopkins University  
kitsing@cs.jhu.edu

## Abstract

We propose a universal grapheme-phoneme transduction model using neuralized finite-state transducers.

Many computational models of grapheme-phoneme transduction nowadays are based on the (autoregressive) sequence-to-sequence string transduction paradigm. While such models have achieved state-of-the-art performance, they suffer from theoretical limitations of autoregressive models. On the other hand, neuralized finite-state transducers (NFSTs) have shown promising results on various string transduction tasks. NFSTs can be seen as a generalization of weighted finite-state transducers (WFSTs), and can be seen as pairs of a featurized finite-state machine (‘marked finite-state transducer’ or MFST in NFST terminology), and a string scoring function. Instead of taking a product of local contextual feature weights on FST arcs, NFSTs can employ arbitrary scoring functions to weight global contextual features of a string transduction, and therefore break the Markov property. Furthermore, NFSTs can be formally shown to be more expressive than (autoregressive) seq2seq models. Empirically, joint grapheme-phoneme transduction NFSTs have consistently outperformed vanilla seq2seq models on grapheme-to-phoneme and phoneme-to-grapheme transduction tasks for English. Furthermore, they provide interpretable aligned string transductions, thanks to their finite-state machine component.

In this talk, we propose a multilingual extension of the joint grapheme-phoneme NFST. We achieve this goal by modeling typological and phylogenetic features of languages and scripts as optional latent variables using a finite-state machine. The result is a versatile grapheme-phoneme transduction model: in addition to standard monolingual and multilingual transduction, the proposed multilingual NFST can also be used in various controlled generation scenarios, such as phoneme-to-grapheme transduction of an unseen language-script pair. We also plan to release an NFST software package.