Abstract:

This paper presents two solutions to the question of how to build software for finding the morphological parses of words in the Qashgar dialect of Modern Uyghur. Qashgar Uyghur poses a problem for existing morphological parsing software, since the way in which phonological processes interact in the language can make it difficult to identify stems and affixes. In particular, the phonology of Qashgar Uyghur appears to demand a cyclic analysis, in which a phonological process applying at one morphologically-defined cycle can feed phonological processes applying during later cycles.

I show that it is possible to model some of these cyclic interactions in Qashgar Uyghur with a finite state transducer (see Bessley & Karttunen, 2003), which can then be used for morphological parsing. I also show, though, that there are certain cyclic interactions in natural languages that cannot be modeled with finite state transducers. In these cases, I suggest that it is possible to treat morphological parsing as a search problem, as Anderson (1988) proposes, but that these searches can be made tractable by choosing which derivations to build and test in an informed way. I present my implementation of an informed, downstream search for Qashgar Uyghur, and I describe how informed searches can be set up for other languages.

References:
