Combining Approaches to IR

W. Bruce Croft
Computer Science Department
University of Massachusetts
Amherst, MA 01003-4610
croft@cs.umass.edu

ABSTRACT

The combination of different text representations and search strategies has become a standard technique for improving the effectiveness of information retrieval. Combination, for example, has been studied extensively in the TREC evaluations and is the basis of the "meta-search" engines used on the Web. We examine the development of this technique, including both experimental results and the retrieval models that have been proposed as formal frameworks for combination. We show that combining approaches for information retrieval can be modeled as combining the outputs of multiple classifiers based on one or more representations, and that this simple model can provide explanations for many of the experimental results. We also discuss recent results with mixture models and language modeling approaches.