

Axiomatic Analysis and Optimization of Information Retrieval Models

ChengXiang Zhai

University of Illinois at Urbana-Champaign
czhai@cs.uiuc.edu

Abstract. Development of optimal retrieval models is an important, yet challenging research problem in information retrieval. Although many effective retrieval models have been proposed, there is still no clear single winner, making it interesting to ask the question whether there exists a single optimal retrieval model that is better than all others. However, this question is theoretically ill defined unless we can formally characterize what properties must be satisfied by an optimal retrieval model. In this talk, I will present a number of formal constraints that an optimal retrieval model are expected to satisfy, and show that these constraints not only provide a formal framework for analytically assessing the optimality of a retrieval model, but also are necessary for diagnosing deficiencies of existing models and improving them. I will use several examples to show that such an axiomatic analysis is required in order to better understand and bridge the gap between theoretically motivated models and empirically effective retrieval functions. Finally, I will discuss some interesting challenges in developing a complete axiomatic analysis framework for seeking an ultimately optimal retrieval model.