

## Optimality-Driven Ranking of Joined Search Results

Yaron Kanza, Technion

In complex search tasks, it is often required to pose several basic search queries, join the answers to these queries, where each answer is given as a ranked list of items, and return a ranked list of combinations. However, the join result may include too many repetitions of items, and hence, frequently the entire join is too large to be useful. This can be solved by choosing a small subset of the join result. The focus of this talk is on how to choose this subset. We propose two measures for estimating the quality of result sets, namely, coverage and optimality ratio. Intuitively, maximizing the coverage is including in the result as many as possible appearances of items in their optimal combination. Maximizing the optimality ratio means striving to have each item appearing only in its optimal combination, i.e., only in the most highly ranked combination that contains it. One of the difficulties, when choosing the subset of the join in a complex search, is that there is a conflict between maximizing the coverage and maximizing the optimality ratio.

In the talk, we will present new semantics for complex search queries that intended to provide a high combination of coverage and high optimality ratio. We will compare the quality of the results under existing semantics such as top-k and skyline to the novel semantics, according to the two proposed measures, and we will describe algorithms for answering complex search queries under the new semantics. Finally, we will present an experimental study, over searches in Yahoo! Local Search Web Services, to illustrate the effectiveness of the proposed algorithms.

This is a joint work with Mirit Shalem.