Using order constraints in crowd data sourcing

We consider the problem of extracting knowledge from crowd users on crowdsourcing platforms. One example application is to match a product to categories in a taxonomy; another application is to find out which combinations of activities are frequent. In such contexts, the system must retrieve many different data values, e.g., the compatibility of the product to each category, or the popularity of each activity combination. Further, these unknown data values are related by order constraints, e.g., the inclusion relationships in the taxonomy, or the monotonicity of the support function in frequent itemset mining. The crowd mining system must therefore leverage these constraints to infer missing information, and to choose the right questions to ask to the crowd. In this talk, we will present our general model for this task, and review our results on frequent itemset mining (ICDT'14) and top-k querying of unknown values under order constraints (ICDT'17).