Ontology Assisted Crowd Mining

Yael Amsterdamer
Susan B. Davidson*
Tova Milo
Slava Novgorodov
Amit Somech

Goals
- Crowd data sourcing is a powerful data procurement paradigm
- We would like harness it in order to allow users to:
  - specify their information needs in a declarative manner
  - efficiently mine the crowd for relevant data
  - obtain a concise list of answers that represent frequent, significant data patterns

A Motivating Example
Ann is planning a vacation in NYC with her family:

“T’m looking for activities to do at a child-friendly attraction in New York, and a good restaurant near by”

Answers should be relevant, represent popular recommendations, and may include additional advice from the crowd

“You can go bike riding in Central Park and eat at Maoz Vegetarian.
Tips: Rent bikes at the boathouse”

Web search can only extract information from recorded data
- It may be hard to extract combinations of restaurant and activity

Forums can yield targeted answers,
- which require reading, aggregating, identifying consensus, etc.

We propose a new, alternative approach using crowd mining

Efficient Query Evaluation Algorithm
- Lazily construct a semantic subsumption partial order
- Traverse it in a top-down manner
- Automatically generate crowd questions
- Prune insignificant parts

The goal is minimizing the # crowd questions

Additional Aspects of the Algorithm
- Open questions – letting crowd members specify patterns
  “What else do you do when you play basketball in Central Park?”

- Asking a sequence of questions “in context”
- Quick pruning of irrelevant items by crowd members
- Multiple crowd workers in parallel
- Output quality assurance

Experimental Results

OASSIS Screenshots