Lecture 17:
CS 5306 / INFO 5306:
Crowdsourcing and
Human Computation

Human Computation, Chapter 6
The Art of Asking Questions

“the design of tasks … can influence the way human computers compute – motivating them to tell the truth, enhancing (or degrading) the quality of their outputs or making them reach an answer faster”

Points of Intervention for Quality Control

- Before computation:
  routing tasks to the right worker
- After computation:
  appropriate aggregation and filtering
- During computation:
  “safeguards placed *at* the time of computation”

Elements of a Task

Task = “an actual piece of work ... that is performed by a human worker”

Elements of a Task

- Basic information:
  - Inputs
  - What is being computed / the question being asked
  - Allowable outputs
- Conditions for success:
  - ESP Game: words match
- Incentives:
  - ESP Game: 10 points per match

Task Design Decisions

1. Information given to the worker:
   - Task performance is influenced by the information presented to workers
   - Quality of instructions (precise? unambiguous? …)
   - Do workers get value from seeing other workers' solutions?

Task Design Decisions

2. Granularity of task:
   - Is the task well-defined or too large – should it be decomposed into simpler subtasks?
   - Will workers understand the task in the “right” way to compute the intended function?
3. Independence:
   - Will workers do works independently?
   - Can workers communicate? … collaborate?
     - If yes, how?

Task Design Decisions

4. Incentives:
   - Are you paying enough?
   - Are workers being given suitable motivation to work "to the best of their abilities"?
5. Quality control:
   - Is the work correct?

Task Information

“There are plenty of psychology experiments that show that human subjects can be systematically bised by how a question is presented and what information is included”

Task Information

“The list of cognitive biases is long .. But there has been little research on the effects they have on the way workers perform computational tasks”

Relevant Cognitive Biases

- Achoring:
  - Estimating 1 x 2 x 3 x 4 x 5 x 6 x 7 x 8
    versus 8 x 7 x 6 x 5 x 4 x 3 x 2 x 1
- Language: Ask a farmer a question in terms of produce
- Sequential context biases: Preceding answer impacts next answer
  - Intensities of answers are connected
  - Example: If previous images was pretty or ugly, affects answer to attractiveness of next image

Information as Assistance / Bias

- Partial solutions
  - Example: Initial configuration in Foldit
  - Example: Iterative improvement algorithms
  - Example: Earlier information biases subsequent answers, such as in answer length or vocabulary used

Task Granularity

- Example: Clustering using human assessment of similarity

- “Adaptively learning the crowd kernel”, Tamuz, O., Liu, C., Belongie, S., Shamir, O., & Kalai, A. T., Proceedings of the 28th International Conference on Machine Learning, 2011

Task Granularity

- Example: Cluster these ties in subsets that are similar

Task Granularity

- Clustering using human assessment of similarity

- “Adaptively learning the crowd kernel”, Tamuz, O., Liu, C., Belongie, S., Shamir, O., & Kalai, A. T., Proceedings of the 28th International Conference on Machine Learning, 2011

Task Granularity

- Clustering using human assessment of similarity

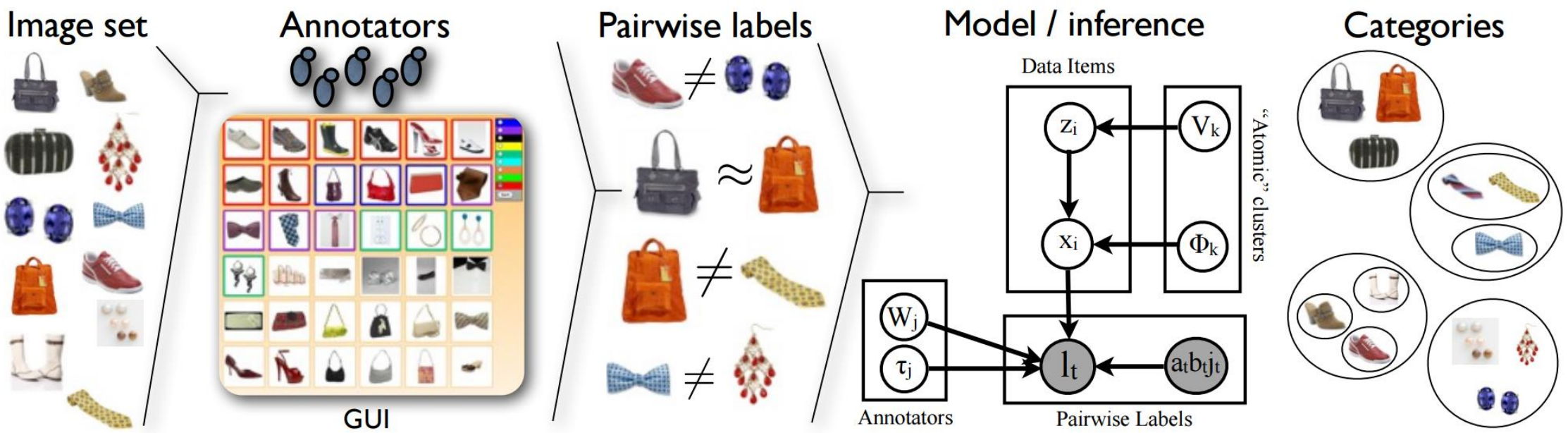
- “Adaptively learning the crowd kernel”, Tamuz, O., Liu, C., Belongie, S., Shamir, O., & Kalai, A. T., Proceedings of the 28th International Conference on Machine Learning, 2011
  - Use: Is A more like B or C?

Task Granularity

- Clustering using human assessment of similarity

- "Crowdclustering", Gomes, R. G., Welinder, P., Krause, A., & Perona, P., *Advances in neural information processing systems* (pp. 558-566), 2011
  - Give M << N items, ask worker to cluster them

Task Granularity



Task Granularity

- Soylent: Find, Fix, Verify
- Iterative improvement algorithm

Task Independence

“While it is common practice that workers perform tasks alone, there are new platforms emerging that will allow workers to perform tasks by interacting with each other”

Task Independence

"While it is common practice that workers perform tasks alone, there are new platforms emerging that will allow workers to perform tasks by interacting with each other"

- No mention of *Infotopia*-like results
- What does task routing look like for a group?

Incentives

- Incentives can impact:
  - Whether workers do any task
  - Which tasks workers do
  - How well they perform the tasks that they do
- Questions:
  1. Form of incentive
  2. How much incentive
  3. Resistance to manipulation

Incentive Taxonomy

- Extrinsic motivation: money, virtual rewards
- Intrinsic motivation:
  - Power: “desire to influence”
  - Curiosity: “to know”
  - Status: “social standing”
  - Social contact: “companionship and play”
  - Competition: “get even”
  - Idealism: “improve society”
  - Ownership: “to collect”

Incentive Taxonomy, Simplified

- Tom Malone, MIT:

Glory

Love

Money

Incentives

- “The interaction between extrinsic and intrinsic motivation is complex”
- Example: Paying to do a task can decrease subsequent internal motivation for the task
- Example: “target earner”
  - Compensated after reaching some milestone
  - Make progress easy to visualize (example: multiple of 5)

Quality Control

- Task design: So the correctness of the worker's "output" can be checked for correctness
  - Verification: Compare to known answer
  - Voting: Select best
  - Filtering: Vote out worst
  - Merging: Combine work to smooth out differences

Use multiple workers

Quality Control

- Social forces and protocols
  - Social norms
  - Sanctions
  - Legal contracts
  - Promise of future work based on performance
  - Sense of community

Quality Control

- Add another question that doesn’t change the amount of work by much but that you knowthe answer to
- Provide signals to workers that they are being monitored
- Put verifiable question before subjective one
  - Dropped incorrect responses by 43%

Quality Control:
Eliciting Truthful Responses

- Game Theory:
  - Example: Vickrey-Clarke-Groves (VCG) auctions
  - “Incentive compatible”: bidders cannot do better by misreporting their true valuation of an item

Quality Control:
Eliciting Truthful Responses

- Games with a purpose
  - Output agreement: ESP Game:
  - Input agreement: TagATune
  - Inversion problems: Peekaboom
  - Complememtary agreement: Polarity