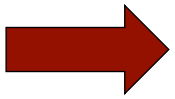
Approval Voting and Incentives in Crowdsourcing

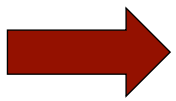
Nihar B. Shah, UC Berkeley

Dengyong Zhou, Microsoft Research

Yuval Peres, Microsoft Research

Crowdsourcing for labeled data

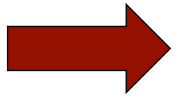


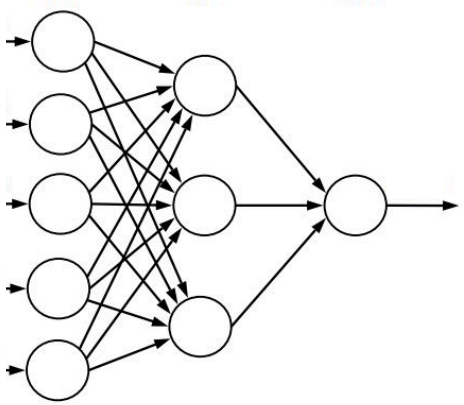


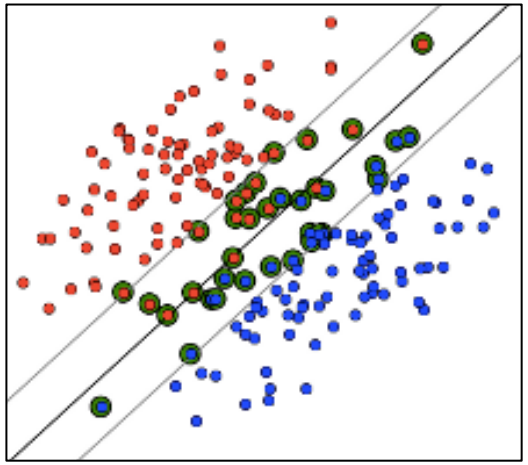
CAT

CAT

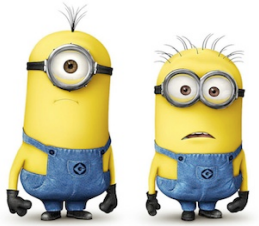
DOG







Data is often of low-quality

1. Workers not experts  , freeloaders 
2. Absence of proper incentives
3. Interface does not allow workers to convey their knowledge accurately

Want to get higher quality data for same or lower costs

Standard interface: “single selection”

Identify the language in the picture

- Hindi
- Chinese
- Hebrew
- Japanese
- French
- German

Worker must select **one** option

Psychology suggests “Approval voting”

Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

- Worker can select any number of options
- Many advantages [Horst 1932; Collet 1971; Brams & Fishburn 1978; Gibbons et al. 1979]
  - More flexibility of expression
  - Utilizes partial knowledge more effectively

Example

Identify the language in the picture

- Hindi
- Chinese
- Hebrew
- Japanese
- French
- German

Workers 1 & 2: Know its either Chinese or Japanese but clueless between the two

Worker 3: Knows its Chinese

Example

Identify the language in the picture

- ○ Hindi
- ○ Chinese
- ○ Hebrew
- ○ Japanese
- ○ French
- ○ German

Workers 1 & 2: Know its either Chinese or Japanese but clueless between the two

Worker 3: Knows its Chinese

Worker 1

- ○ Hindi
- ○ Chinese
- ○ Hebrew
- ● Japanese
- ○ French
- ○ German

Worker 2

- ○ Hindi
- ○ Chinese
- ○ Hebrew
- ● Japanese
- ○ French
- ○ German

Worker 3

- ○ Hindi
- ● Chinese
- ○ Hebrew
- ○ Japanese
- ○ French
- ○ German

Final conclusion: Japanese (wrong)

Example

Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

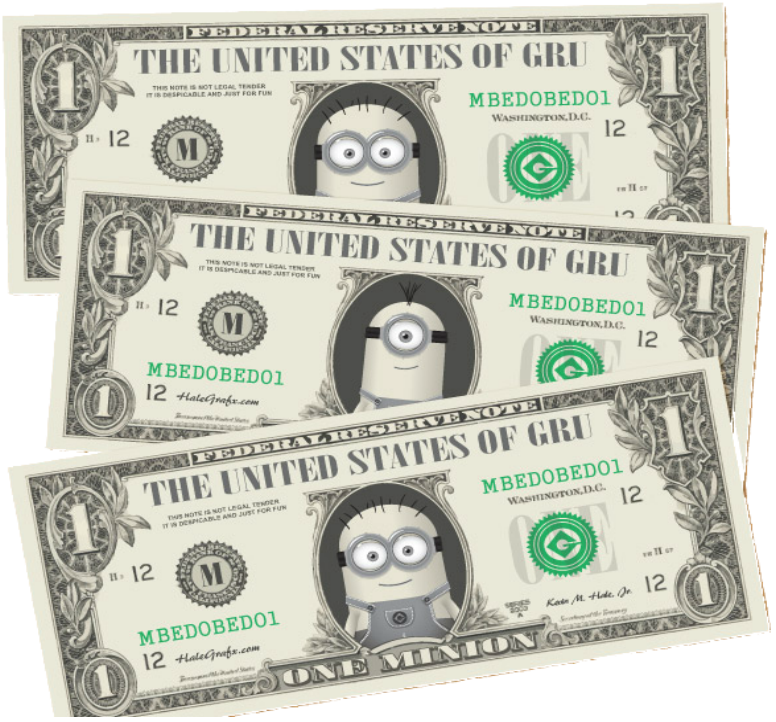
Workers 1 & 2: Know its either Chinese or Japanese but clueless between the two

Worker 3: Knows its Chinese

Worker 1	Worker 2	Worker 3
☐ Hindi	☐ Hindi	☐ Hindi
☑ Chinese	☑ Chinese	☑ Chinese
☐ Hebrew	☐ Hebrew	☐ Hebrew
☑ Japanese	☑ Japanese	☐ Japanese
☐ French	☐ French	☐ French
☐ German	☐ German	☐ German

Final conclusion: Chinese (correct)

Goal: Design payment mechanisms to incentivize workers to respond “appropriately”



Outline

- Problem setting
- Roadblock, and detour
- Mechanism
- Why this (and only this) mechanism
- Preliminary experiments

Problem setting

- Multiple-choice questions

1. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

2. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

3. Identify the language in the picture

▪
▪
▪

Problem setting

- Multiple-choice questions
- One or more “gold standard” questions: answers known apriori
- Evaluate worker on gold standard

1. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

2. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

3. Identify the language in the picture

▪
▪
▪

Problem setting

- Multiple-choice questions
- One or more “gold standard” questions: answers known apriori
- Evaluate worker on gold standard
- Worker aims to maximize expected payment

1. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

2. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

3. Identify the language in the picture

▪
▪
▪

Problem setting

- Multiple-choice questions
- One or more “gold standard” questions: answers known apriori
- Evaluate worker on gold standard
- Worker aims to maximize expected payment
- Payment non-negative

1. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

2. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

3. Identify the language in the picture

▪
▪
▪

Problem setting

- Multiple-choice questions
- One or more “gold standard” questions: answers known apriori
- Evaluate worker on gold standard
- Worker aims to maximize expected payment
- Payment non-negative
- Budget “B” = maximum payment to a worker

1. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

2. Identify the language in the picture

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

3. Identify the language in the picture

▪
▪
▪

Which options do we want the worker to select?

Select ALL options that could be the language in this image

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

- Coombs 1953: "should be instructed to cross out all the alternatives which they consider wrong"
- Coombs, Milholland, Womer 1956: experimental verification of usefulness

Belief of worker for a question

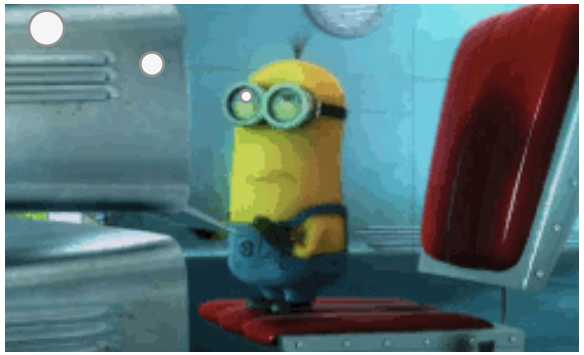
= probability distribution over all options

- ☐ Hindi
- ☑ Chinese
- ☐ Hebrew
- ☑ Japanese
- ☐ French
- ☐ German

Chance of being correct:

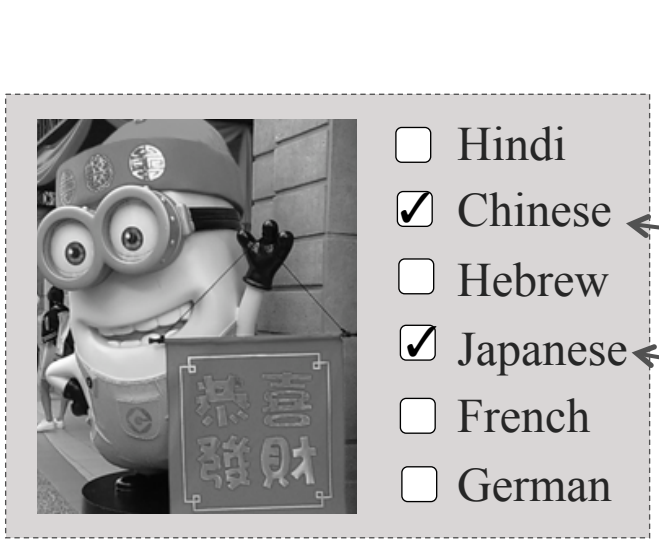
50%

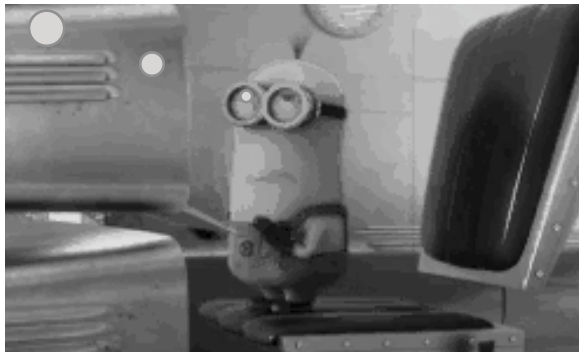
50%



Belief of worker for a question

= probability distribution over all options





Goal: Design payment mechanism such that for each question, worker incentivized to select all options for which belief > 0

- In other words, elicit **support** of belief
- "**Incentive-compatible** mechanism"

An impossibility result

THEOREM

No payment mechanism can be incentive compatible.

Coarse beliefs

Extensive literature in psychology on coarseness of processing and perception in humans

[Miller 1956, Shiffrin & Nosofsky 1994, Saaty & Ozdemir 2003, Mullainathan et al. 2008, Siddiqi 2011, Jones & Loe 2013]

Coarse beliefs assumption: For some (fixed and known) value $\rho > 0$, all non-zero values in the belief are $>\rho$

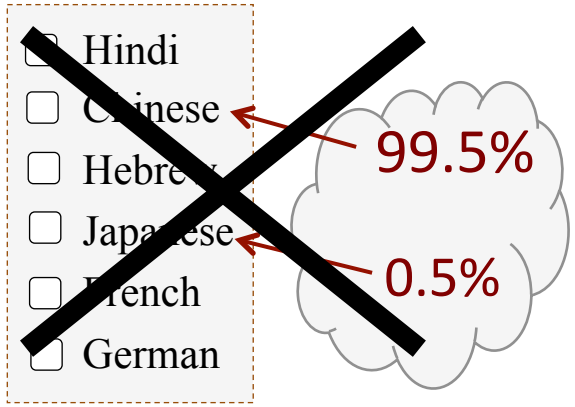
Coarse beliefs

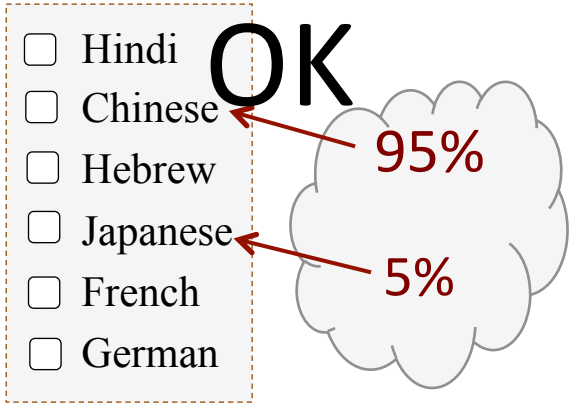
Extensive literature in psychology on coarseness of processing and perception in humans

[Miller 1956, Shiffrin & Nosofsky 1994, Saaty & Ozdemir 2003, Mullainathan et al. 2008, Siddiqi 2011, Jones & Loe 2013]

Coarse beliefs assumption: For some (fixed and known) value $\rho > 0$, all non-zero values in the belief are $>\rho$

Example: $\rho = 1\%$





Our mechanism

Input: coarse belief parameter ρ, budget B,
worker's answers to gold standard questions

Output:

- Payment starts at B
- For every question in gold standard:
  - For every wrong option selected, payment is multiplied by (1-ρ)
  - If the correct option is not selected, payment becomes zero

Example (ρ = 5%, B = $1)

1.

- ☐ Hindi
- ☑ **Chinese**
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

2.

- ☐ Hindi
- ☐ Chinese
- ☐ Hebrew
- ☐ Japanese
- ☑ **French**
- ☐ German

Payment = B = $ 1

Example (ρ = 5%, B = $1)

1.

- ☐ Hindi
- ☑ **Chinese**
- ☐ Hebrew
- ☑ Japanese
- ☐ French
- ☐ German

2.

- ☑ Hindi
- ☑ Chinese
- ☑ Hebrew
- ☑ Japanese
- ☑ **French**
- ☑ German

Payment = $B (1 - \rho)^6 = \$ .74$

Example (ρ = 5%, B = $1)

1.

- ☐ Hindi
- ☑ **Chinese**
- ☐ Hebrew
- ☐ Japanese
- ☐ French
- ☐ German

2.

- ☐ Hindi
- ☐ Chinese
- ☑ Hebrew
- ☐ Japanese
- ☐ **French**
- ☑ German

Payment = $ 0

Our mechanism: analytical guarantees

THEOREM

Our mechanism is incentive-compatible under the coarse-beliefs assumption.

Continues to do something desirable even when beliefs are not coarse
(ask me offline)

There could be many other incentive-compatible mechanisms

Why use this mechanism?

Minimum possible expenditure on freeloaders

Selecting all options ⇒ Free money

- ☑ Hindi
- ☑ Chinese
- ☑ Hebrew
- ☑ Japanese
- ☑ French
- ☑ German

THEOREM

Among all possible incentive-compatible mechanisms, our mechanism makes the *strictly* smallest payment to a freeloader.

Unique under a no-free-lunch requirement

If **all** attempted answers are wrong, then payment must be zero.

[Shah and Zhou, 2014]

Unique under a no-free-lunch requirement

If **all** attempted answers are wrong, then payment must be zero.

doesn’t select all options

Unique under a no-free-lunch requirement

If **all** attempted answers are wrong, then payment must be zero.

doesn't select all options

doesn't select correct option

Unique under a no-free-lunch requirement

If **all** attempted answers are wrong, then payment must be zero.

doesn't select all options

doesn't select correct option

This is a weak requirement: no restriction if even one of the answers is correct

Unique under a no-free-lunch requirement

If **all** attempted answers are wrong, then payment must be zero.

doesn't select all options

doesn't select correct option

This is a weak requirement: no restriction if even one of the answers is correct

THEOREM

Our mechanism is the one and only incentive-compatible mechanism that satisfies no-free-lunch.

Preliminary experiments

Goal: to evaluate the primary hypotheses underlying the theory

- Are workers making judicious use of the approval voting setup?
- Does presence of a mechanism improve quality?
- Is there is any opposition from the workers?

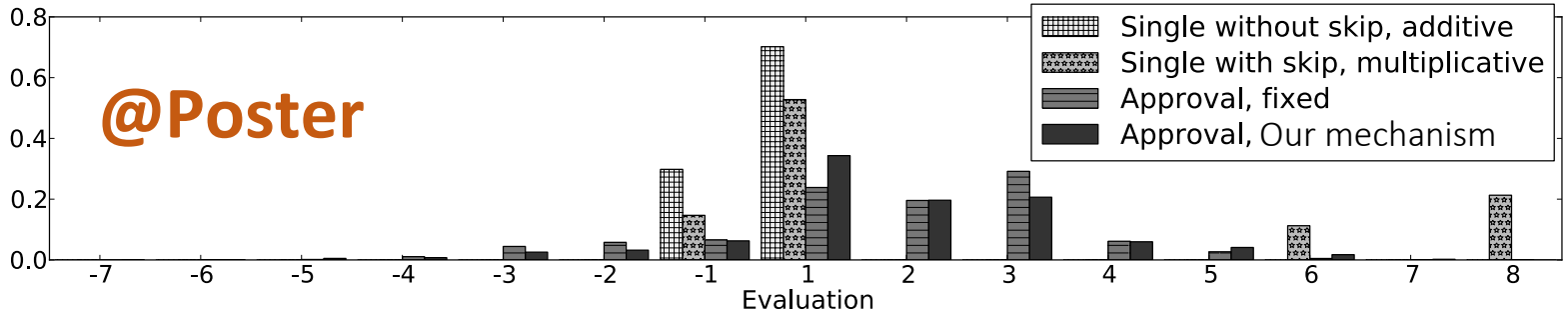
Preliminary experiments

Goal: to evaluate the primary hypotheses underlying the theory

- Are workers making judicious use of the approval voting setup?
- Does presence of a mechanism improve quality?
- Is there is any opposition from the workers?

400 workers

(Collected data available online)



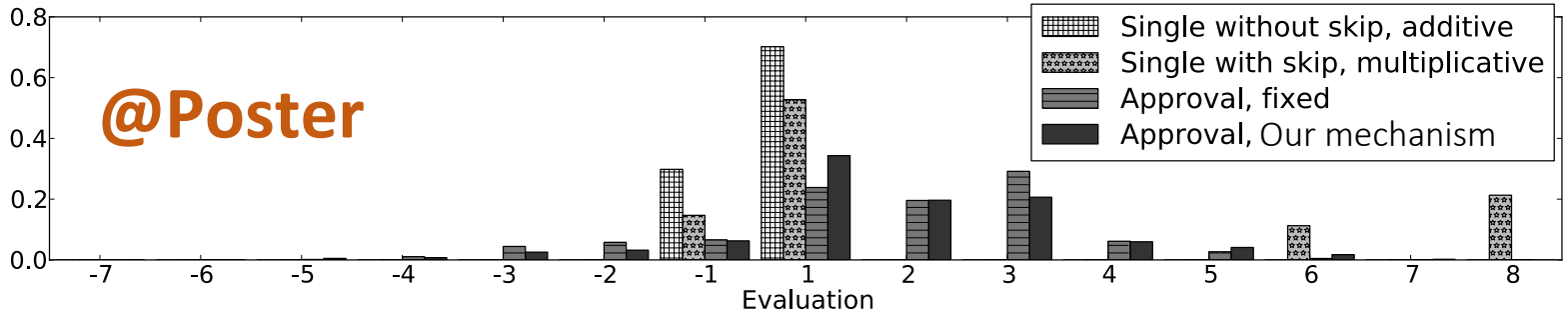
Preliminary experiments

Goal: to evaluate the primary hypotheses underlying the theory

- Are workers making judicious use of the approval voting setup? **Yes**
- Does presence of a mechanism improve quality? **Yes**
- Is there is any opposition from the workers? **No**

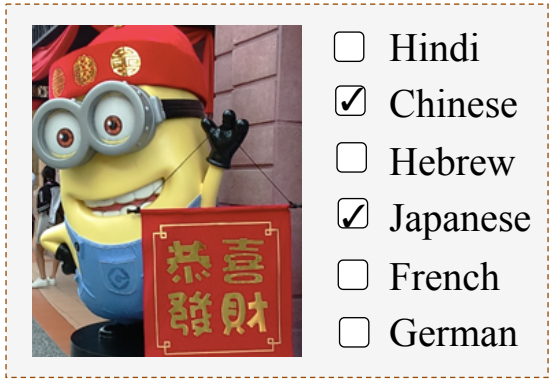
400 workers

(Collected data available online)



Summary and future work

- "Approval voting" interface for crowdsourcing
- Design mechanism:
  - Incentive-compatible
  - Strictly minimum payment to freeloaders
  - Only one to satisfy a natural "no-free-lunch" requirement
  - Simple
- Future work: Design statistical aggregation algorithms
  - Incorporate approval voting setting
  - Exploit structure of data due to mechanism



Thanks! Questions?

(Hope to chat more
with you at the poster)

nihar@berkeley.edu