

Truth Inference in Crowdsourcing: Is the Problem Solved?



Yudian Zheng*, Guoliang Li#, Yuanbing Li#, Caihua Shan*, Reynold Cheng*

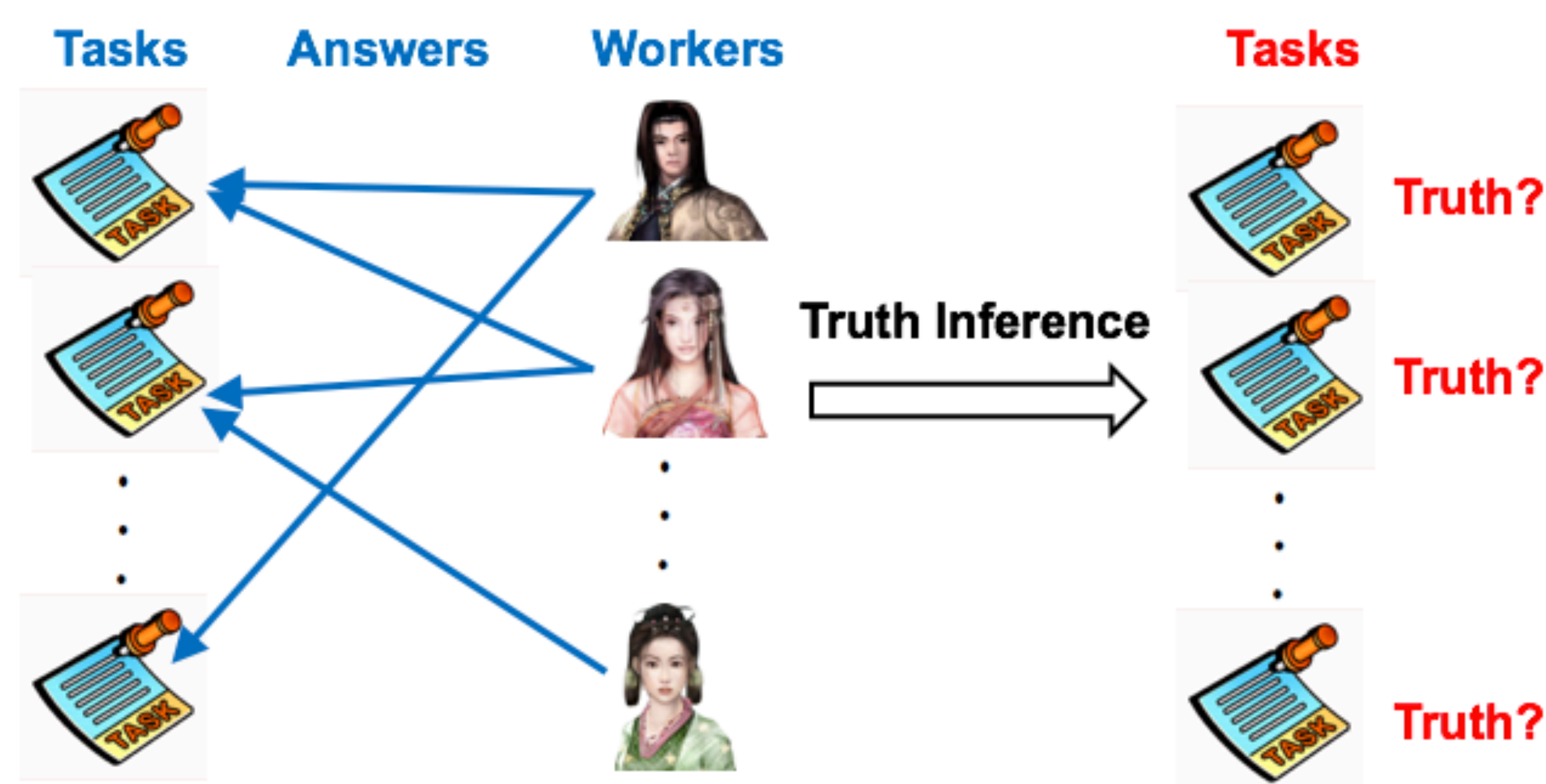
*University of Hong Kong, #Tsinghua University

Motivating Example



Truth Inference Definition

Given **different tasks' answers collected from workers**, the target is to **infer the truth of each task**.



Three Goals

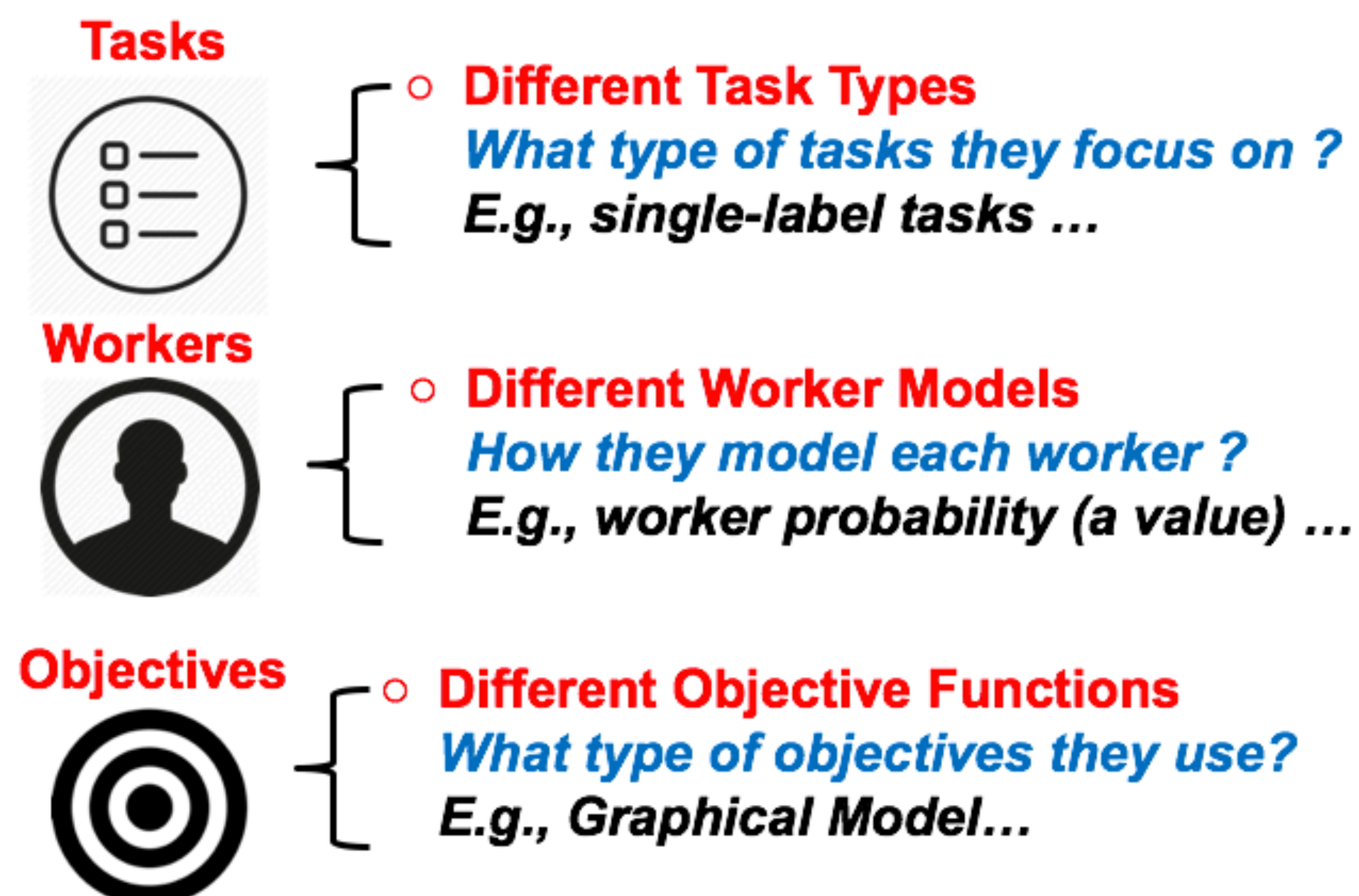
- What are the similarities in existing works?
- What are the differences in existing works?
- Any suggestions to use in practice?

Part I. Unified Framework

- Input: Workers' Answers for all tasks
- Algorithm Framework

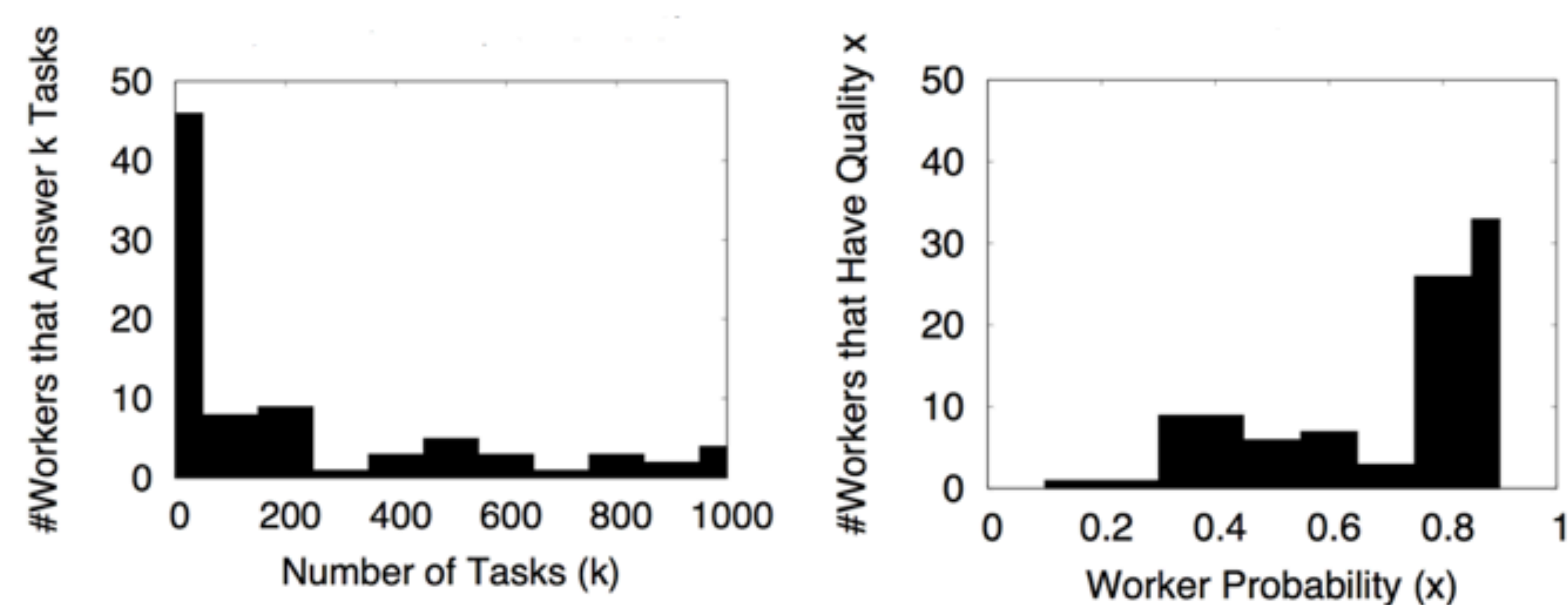
```
Initialize Quality for each worker
While (not converged) {
    Quality for each worker → Truth for each task ;
    Truth for each task → Quality for each worker ;
}
```

Part II. Comparisons



Part III. Experiments and Analysis

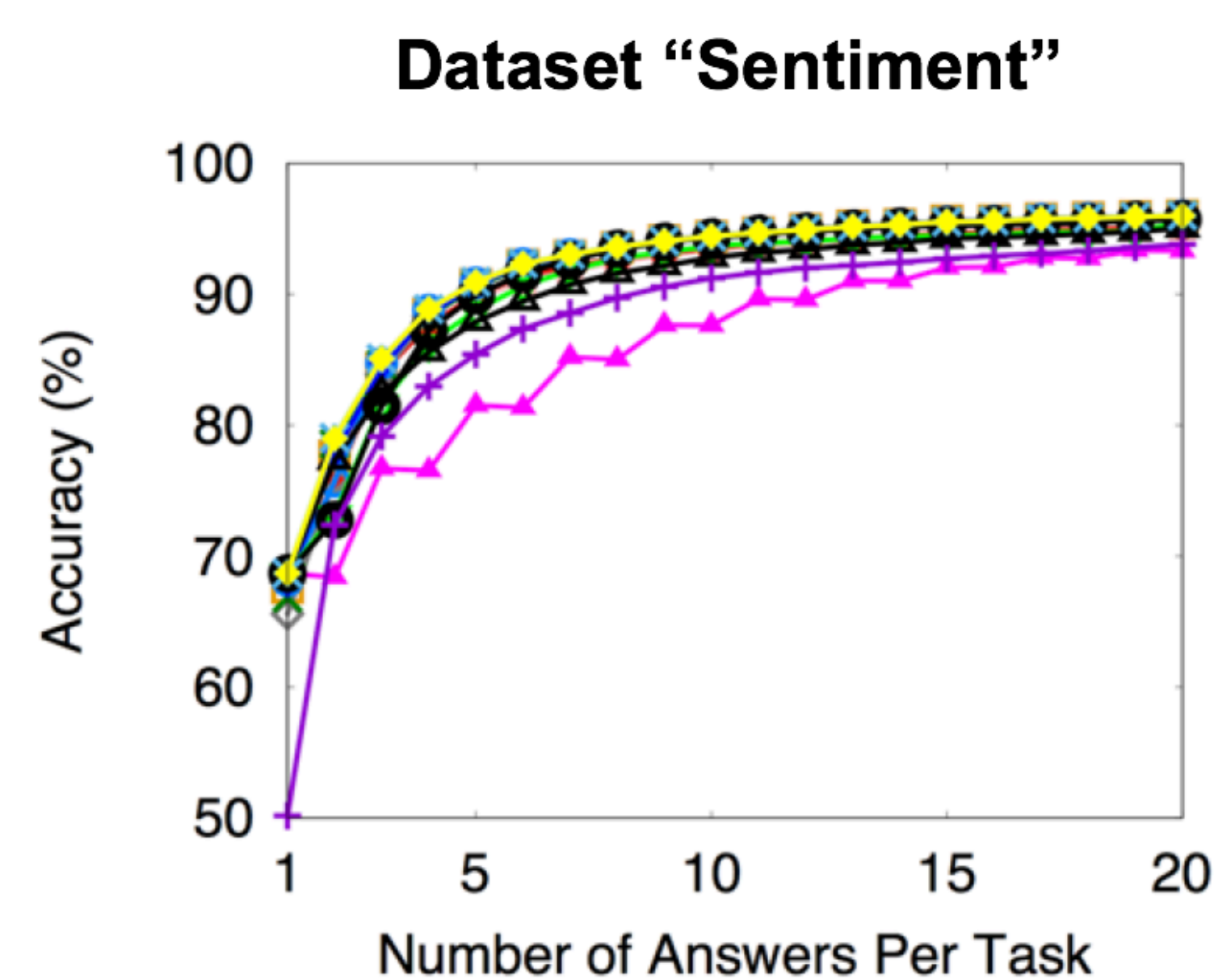
Observations



#workers' answers conform to **long-tail phenomenon**

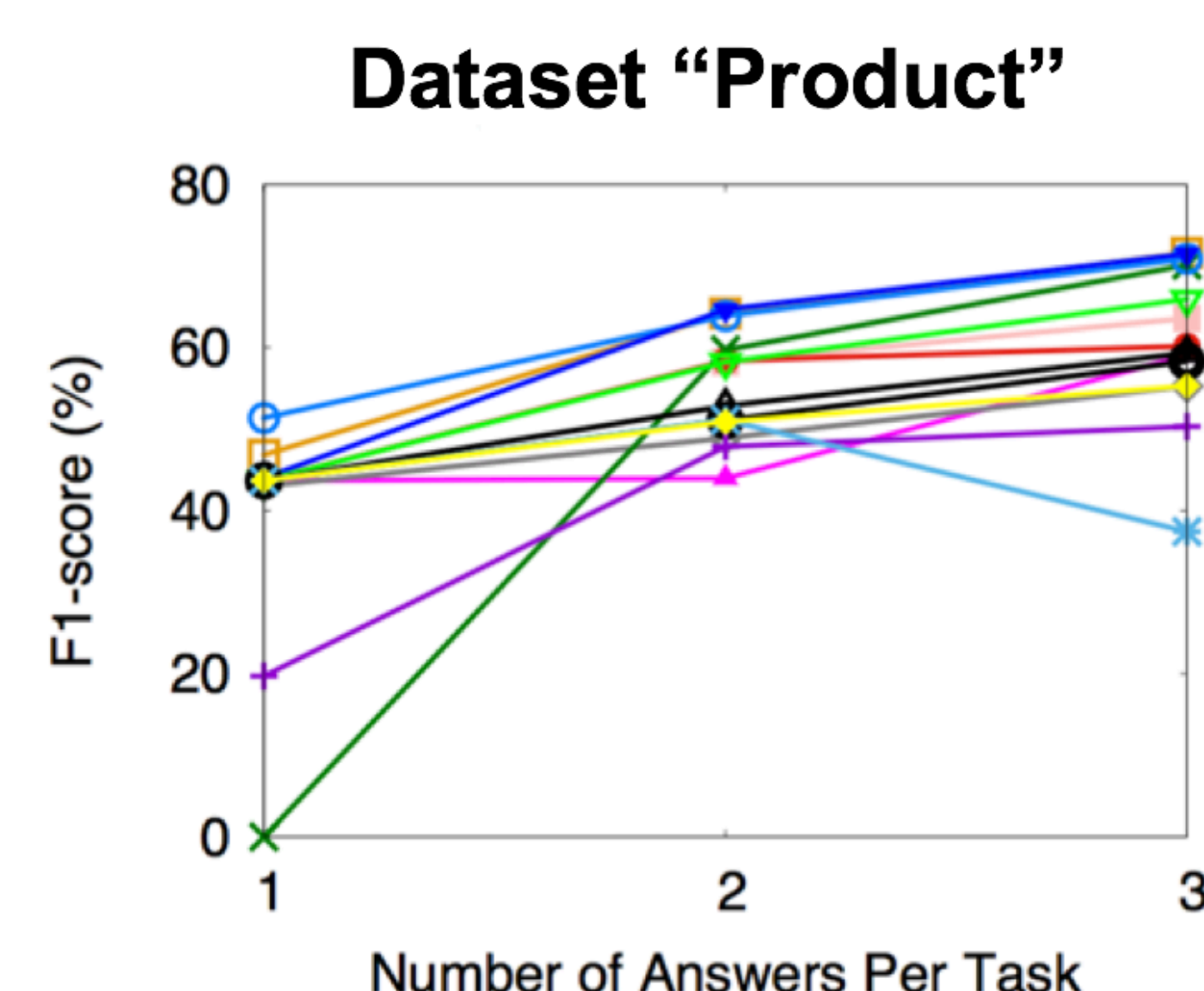
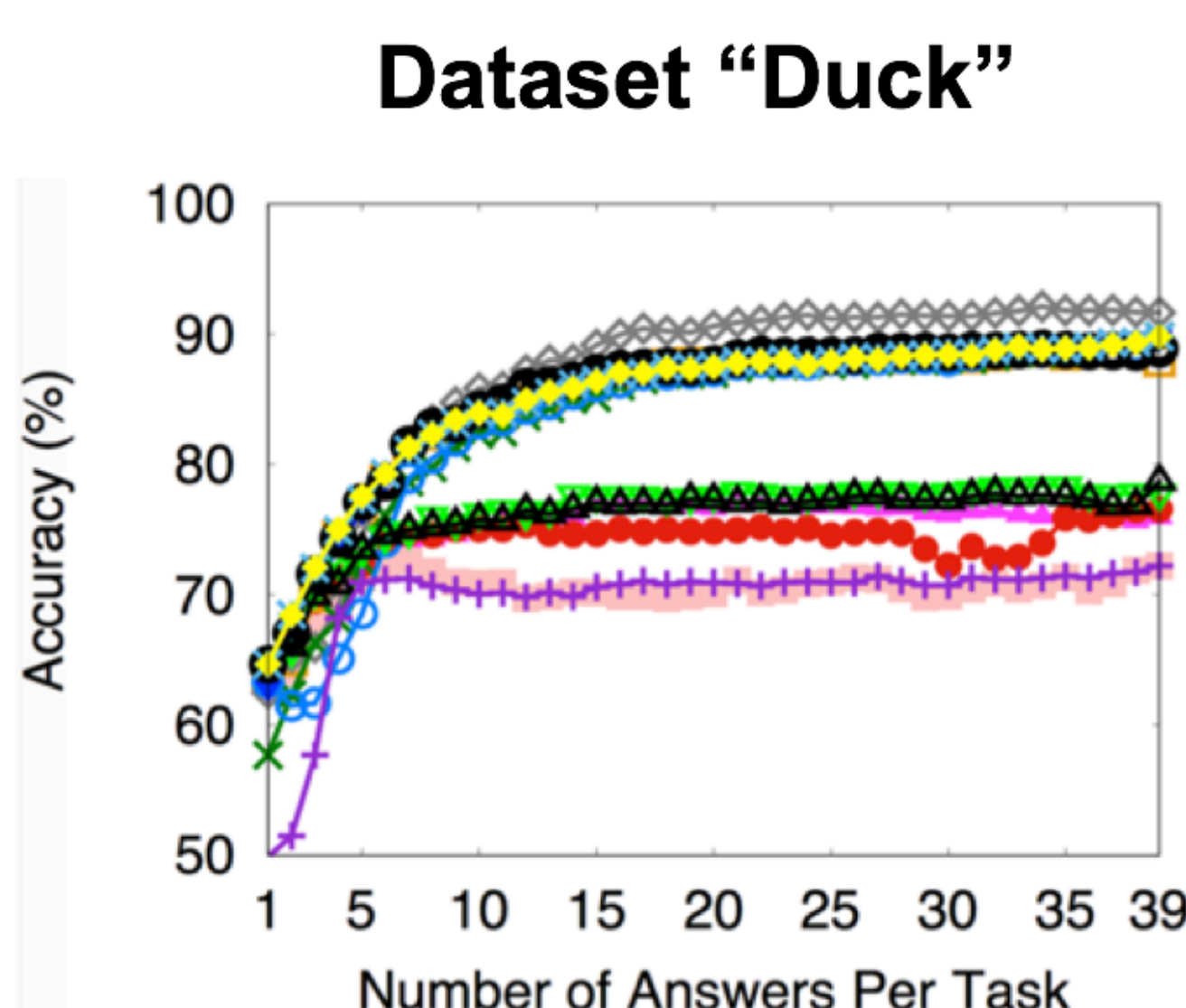
Not all workers are of very high quality

Change of Quality with the #Answers



Observations:

- The quality **increases with #answers**;
- The quality improvement is **significant with few answers**, and is **marginal with more answers**;
- Most methods are similar, except for **Majority Voting** (in pink color).



DOCS: A Domain-Aware Crowdsourcing System Using Knowledge Bases

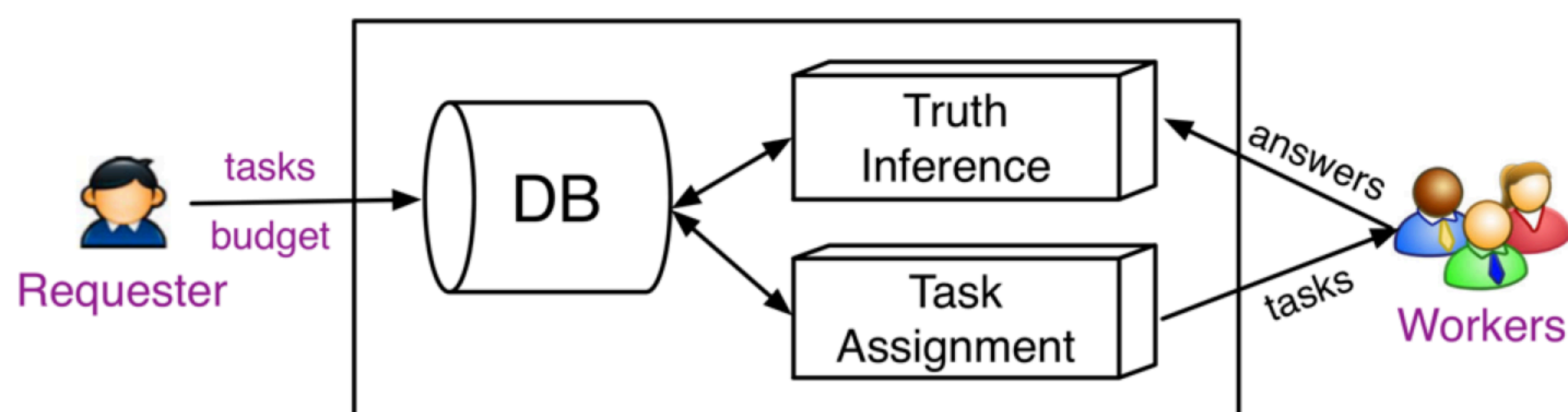


Yudian Zheng*, Guoliang Li#, Reynold Cheng*

*University of Hong Kong, #Tsinghua University

Crowdsourcing Workflow

- Task Assignment:** When a worker comes to the platform, the worker will be assigned to a set of tasks;
- Truth Inference:** When a worker accomplishes tasks, the platform will collect answers from the worker.



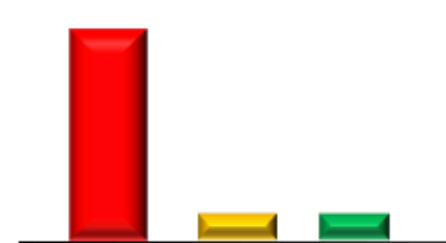
Existing Works Fail in QA Tasks

- Each task is related to different domains

■ Sports ■ Politics ■ Entertainment

Did Michael Jordan win more NBA championships than Kobe Bryant?

→ Sports

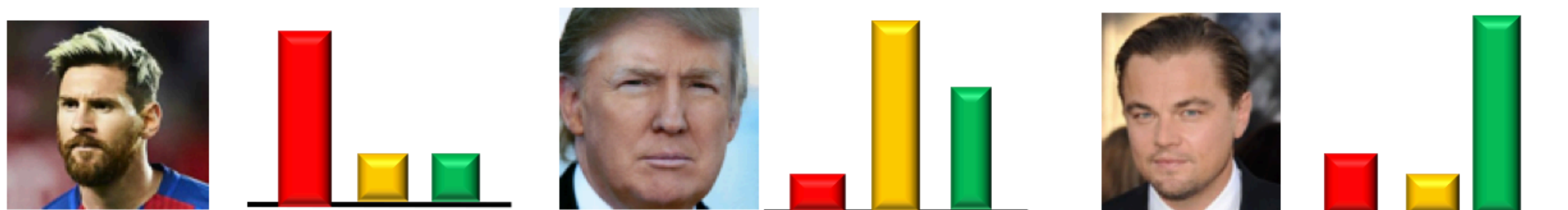


Is there a name for the song that FC Barcelona is known for?

→ Sports & Entertainment

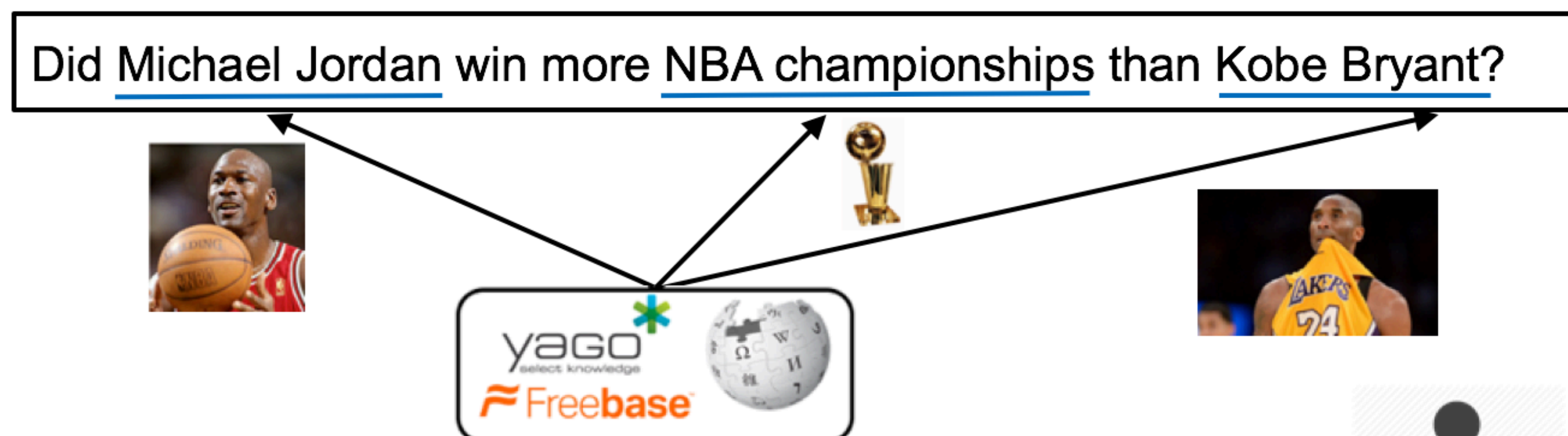


- Each worker has diverse qualities over domains



Build Domain Aware Task Model

- (1) **Entity linking** (map entity to **knowledge bases**)



- (2) Hierarchical domains in knowledge bases
- (3) Obtain the task model (a vector of distribution)

■ Sports ■ Politics ■ Entertainment

Did Michael Jordan win more NBA championships than Kobe Bryant?

→ Sports



Build Domain Aware Worker Model

- Use **qualification test** (like an "exam") amazon mechanical turk Artificial Intelligence
- Two rules for selecting **qualification test**
 - (1) Each selected task should **capture a certain domain**
 - (2) The domain distribution of selected tasks should **approximate the distribution of all tasks**

KL-divergence

$$\min_{\{n'_k\}} \sum_{k=1}^m \frac{n'_k}{n'} \cdot \ln \frac{n'_k \cdot n}{n' \cdot \sum_{i=1}^n r_k^{t_i}}$$

s.t. $\sum_{k=1}^m n'_k = n'$ and $n'_k \in \mathbb{N}$ for $1 \leq k \leq m$.

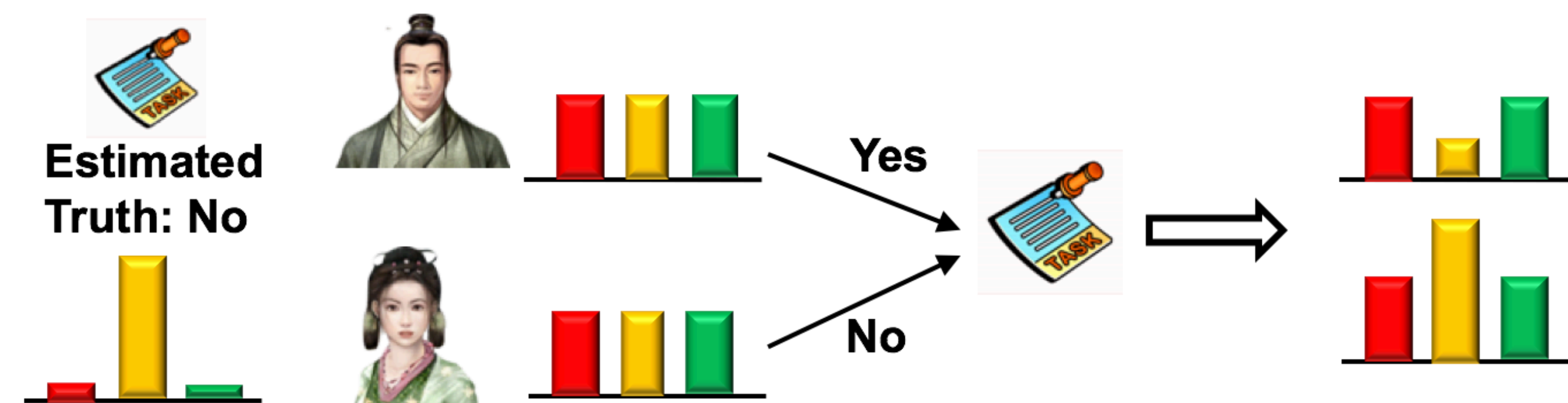
Truth Inference

- 1. **Quality for each worker** \Rightarrow **Truth for each task**

■ Sports ■ Politics ■ Entertainment



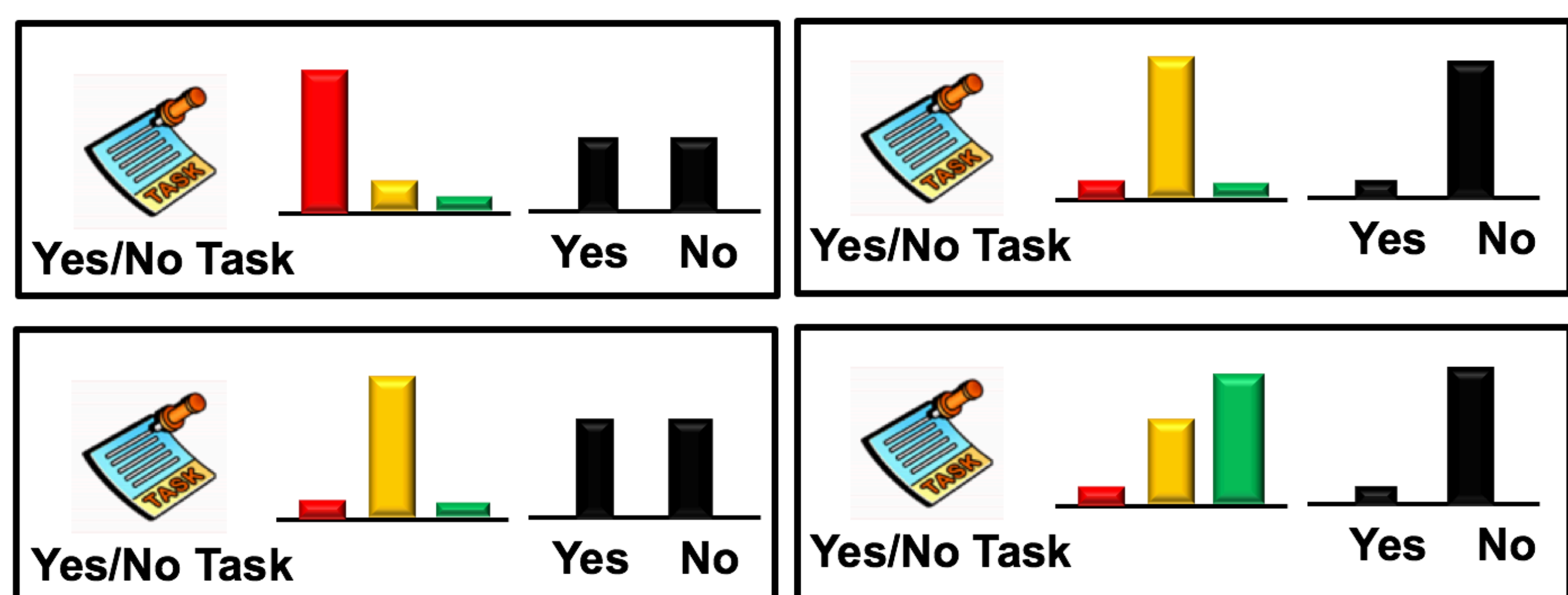
- 2. **Truth for each task** \Rightarrow **Quality for each worker**



Task Assignment

- Select the most suitable tasks for assignment

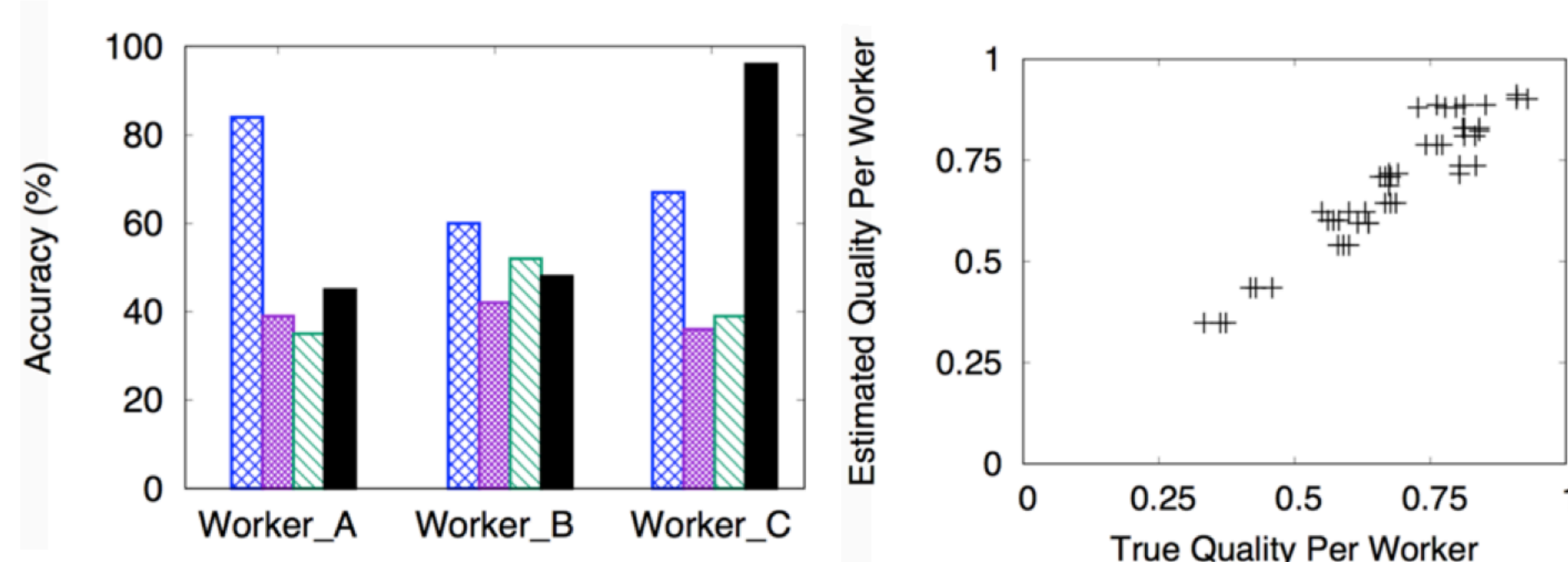
■ Sports ■ Politics ■ Entertainment



- (1) Matching Domains
- (2) Answer Uncertainty

Experiments

- Worker Characteristics on Dataset D_Item



- System Comparisons

