Crowdsourcing & Human Computation
Today

- mainly driven by **Mechanical Turk** and **Social Networks**
Today

~1000

~660

~455

- mainly driven by Mechanical Turk and Social Networks
Taxonomy of Social Computing
[Quinn and Bederson, CHI 2011]

What are the differences?
ABSTRACT

The rapid growth of human computation within research and industry has produced many novel ideas aimed at organizing web users to do great things. However, the growth is not adequately supported by a framework with which to understand each new system in the context of the old. We classify human computation systems to help identify parallels between different systems and reveal holes in the existing work as opportunities for new research. Since human computation is often confused with crowdsourcing and other terms, we explore the position of human computation with respect to these related topics.

Author Keywords

Human computation, crowdsourcing, taxonomy, survey, literature review, social computing, data mining

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI).

General Keywords

Theory

INTRODUCTION

Since the birth of artificial intelligence research in the 1950s, computer scientists have been trying to emulate human-like capabilities, such as language, visual processing, and reasoning. Alan Turing wrote in 1950: “The idea behind digital computers may be explained by saying that these machines are intended to carry out any operations which could be done by a human computer.”

Turing’s article stands as enduring evidence that the roles of human computation and machine computation have been intertwined since the earliest days. Even the idea of humans and computers working together in complementary roles was envisioned in 1960 in Licklider’s sketch of “man-computer symbiosis.” Only recently have researchers begun to explore this idea in earnest [21,50,53].

In 2005, a doctoral thesis about human computation was completed [64]. Four years later, the first annual Workshop on Human Computation was held in Paris with participants representing a wide range of disciplines [28]. This diversity is important because finding appropriate and effective ways of enabling online human participation in the computational process will require new algorithms and solutions to tough policy and ethical issues, as well as the same understanding of users that we apply in other areas of HCI. Today, the field of human computation is being advanced by researchers from areas as diverse as artificial intelligence [35,38,58], business [41,56,29,72], cryptography [64], art [16,31], genetic algorithms [32], and HCI [2,3,5,etc.].

As this area has blossomed with an ever-expanding array of novel applications, the need for a consistent vocabulary of terms and distinctions has become increasingly pronounced. This paper presents a classification system for human computation systems that highlights the distinctions and

Figure 1:

Human computation is a means of solving computational problems. Such problems are found only occasionally in crowdsourcing and social computing applications.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.


Copyright 2011 ACM 978-1-4503-0267-8/11/05...$10.00.
Collabio

[Bernstein et al., UIST 2009]

- Guess the tags that describe your friend best
- Points for common tags
**PhotoCity**

[Tuite et al., CHI 2011]

Task: reconstruct the geometry of buildings from photos

- **PhotoCity** is a competition between two universities to get the better model

- The author describe this as “Game with a purpose”
Players get points for photos with new feature vectors
Better photos have more new features
Website guides the players with flags placed at problematic areas
Human Computation: A Survey and Taxonomy of a Growing Field

Alexander J. Quinn, Benjamin B. Bederson
University of Maryland, College Park
Human-Computer Interaction Lab
:: Computer Science
:: Institute for Advanced Computer Studies
College Park, Maryland 20740
{aq, bederson}@cs.umd.edu

ABSTRACT
The rapid growth of human computation within research and industry has produced many novel ideas aimed at organizing web users to do great things. However, the growth is not adequately supported by a framework with which to understand each new system in the context of the old. We classify human computation systems to help identify parallels between different systems and reveal holes in the existing work as opportunities for new research. Since human computation is often confused with crowdsourcing and other terms, we explore the position of human computation with respect to these related topics.

Author Keywords
Human computation, crowdsourcing, taxonomy, survey, literature review, social computing, data mining

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI).

General Keywords
Theory

INTRODUCTION
Since the birth of artificial intelligence research in the 1950s, computer scientists have been trying to emulate human-like capabilities, such as language, visual processing, and reasoning. Alan Turing wrote in 1950: "The idea behind digital computers may be explained by saying that these machines are intended to carry out any operations which could be done by a human computer." Turing’s article stands as enduring evidence that the roles of human computation and machine computation have been intertwined since the earliest days. Even the idea of humans and computers working together in complementary roles was envisioned in 1960 in Licklider’s sketch of “man-computer symbiosis” [37]. Only recently have researchers begun to explore this idea in earnest [21,50,53].

In 2005, a doctoral thesis about human computation was completed [64]. Four years later, the first annual Workshop on Human Computation was held in Paris with participants representing a wide range of disciplines [28]. This diversity is important because finding appropriate and effective ways of enabling online human participation in the computational process will require new algorithms and solutions to tough policy and ethical issues, as well as the same understanding of users that we apply in other areas of HCI. Today, the field of human computation is being advanced by researchers from areas as diverse as artificial intelligence [35,38,58], business [41,56,29,72], cryptography [64], art [16,31], genetic algorithms [32], and HCI [2,3,5,etc.].

As this area has blossomed with an ever-expanding array of novel applications, the need for a consistent vocabulary of terms and distinctions has become increasingly pronounced. This paper presents a classification system for human computation systems that highlights the distinctions and

Figure 1: Human computation is a means of solving computational problems. Such problems are found only occasionally in crowdsourcing and social computing applications.

• Social computing facilitates relatively natural human behavior mediated by technology
• Crowdsourcing replaces experts with undefined, usually large groups of people
• Collective intelligence: groups of individuals that are doing things that seem intelligent

reCaptcha [von Ahn et al., Science 2008]

PhotoCity [Tuite et al., CHI 2011]

Collabio [Bernstein et al., UIST 2009]
• Prove to be human by deciphering text that OCR cannot
• 200 million CAPTCHAs are solved by humans around the world every day
• Leverage this “processing power”!

The New-York State Yacht Squadron, on its annual cruise to Newport, came into the harbor yesterday afternoon. The following are the names of the boats that came to anchor here: Jessie, Geralda, Evelyn, Anna, Mannering, Julia, Bonita, Magic, Widgeon, Rambler, Fleur-de-Lys, Henrietta, Sea-Drift and Maria, with the steamer America as a tender. On anchoring, each boat fired a gun, according to custom. The reports were heard distinctly in the city, causing considerable inquiry as to “what was up,” and quite a number of sanguine individuals came into our office to inquire if the guns were not annunciantory signals of the successful laying of the Atlantic Cable. We invariably replied in the negative. The squadron will leave to-day for Newport. The yachts Washington and Rambler, of this city, start with it, with parties of New Haven people.
[von Ahn et al., Science 2008]
Human Computation

• Von Ahn, 2005
  “... a paradigm for utilizing human processing power to solve problems that computers cannot yet solve”

• Human as a processor for a larger (computer directed) algorithm
Human Computation

- Von Ahn, 2005

“...a paradigm for utilizing human processing power to solve problems that computers cannot yet solve”
In-class Exercise: Mechanical Turk

• Visit http://www.mturk.com/ click “find hits now”

• What kind of jobs are posted there?

• In what manner are they posted?

• How much are they paid?

• How much time do they take?

• What kind of jobs are missing?
The original Mechanical Turk
VizWiz: Nearly Real-time Answers to Visual Questions

1. Blind user takes a photo with his smartphone and formulates a question
2. Turker has a look at the photo and listens to question
3. Answer is transmitted back to the phone and blind user

[16]

Bigham et al., 2010

Mechanical Turk has made outsourcing small paid jobs practical if Mechanical Turk has been used for a wide variety of tasks including gathering data for user studies, labeling image data sets used in Computer Vision research, and determining political sentiments in blog snippets. The Amazon Remembers feature of its iPhone application lets users take pictures of objects and later emails similar products that Amazon sells. It is widely suspected that Amazon outsources some of these questions to Mechanical Turk. The TurKit library enables programmers to easily employ multiple turk workers using common programming paradigms.

2.3. Connecting Remote Workers to Mobile Devices

Some human-powered services provide an expectation of latency. ChaCha and KGz employees answer questions asked via the phone or by text message in just a few minutes. Other common remote services include relay services for deaf and hard of hearing people which requires trained employees. The retroactive nearly real-time audio captioning by dedicated workers in ScribelMe user study of ScribelMe found that participants felt waiting the required 6 minutes was too long because it “leaves one as an observer rather than an active participant.”

Existing Use of Photos and Video for Assistance:

Several of the blind consultants whom we interviewed mentioned using digital cameras and email to informally consult sighted friends or family in particularly frustrating or important situations, checking one’s appearance before a job interview. Remote reading services for the blind were proposed using low cost fax equipment and sighted remote readers. Compressed video technology allowed very low frame-rated higher resolution video transmission over ordinary telephone lines. oMoby is an iPhone application similar to Google’s, but instead of an automated database lookup, human computation is used. The Soyлent Grid CyPTCH image labeling system requires remote human annotation for CyPTCH images then included in a searchable database.

LookTel is a soon-to-be-released talking mobile application that can connect blind people to friends and family members via a live video feed. Although future versions of VizWiz may similarly employ video, we chose to focus on photos for two reasons. First, mobile streaming is not possible in much of the world because of slow connections. Even in areas with 4G coverage, our experience has been that the resolution and reliability of existing video services like UStream and knocking is too low for many of the questions important to blind people. Second, using video removes the abstraction between user and provider that VizWiz currently provides. With photos, questions can be asked quickly, workers can be employed for short amounts of time, and multiple redundant answers can be returned.

3. VizWiz::LocateIt

Here we present our work on VizWiz::LocateIt, a prototype system that combines remote human work with automatic computer vision to help blind people locate arbitrary items in their environments (Figure kaf). To support object localization we created the following two components: a web interface to let remote workers outline objects, and the VizWiz::LocateIt mobile interface consisting of the Sensor zoom and filter and Sonification modules.
VizWiz: Nearly Real-time Answers to Visual Questions

[Bigham et al., 2010]
Who are the Crowdworkers?

[Ross et al., CHI 2011]

- Mechanical Turk was launched in 2005
- 400k registered workers in 2010
- 338k Hits this morning
- Average pay per hour: 2$
- Typical worker is well-educated, Indian or American, and young

### Table 1: Demographics of Crowdworkers

<table>
<thead>
<tr>
<th>Month</th>
<th>US</th>
<th>India</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mar 08</td>
<td>76%</td>
<td>8%</td>
<td>16%</td>
</tr>
<tr>
<td>Nov 08</td>
<td>66%</td>
<td>5%</td>
<td>12%</td>
</tr>
<tr>
<td>May 09</td>
<td>57%</td>
<td>12%</td>
<td>11%</td>
</tr>
<tr>
<td>Aug 09</td>
<td>56%</td>
<td>8%</td>
<td>11%</td>
</tr>
<tr>
<td>Nov 09</td>
<td>56%</td>
<td>5%</td>
<td>11%</td>
</tr>
</tbody>
</table>

### Figure 2: Nationality of MTurk workers over time.

- Countries with more than 1% of respondents include Canada, the UK, and the Philippines.
- We do not have country data for Feb 2009.

### Figure 3: Gender of MTurk workers over time.

- Around 2/3 of U.S. Turkers are female, 2/3 of Indian Turkers are male.

### Figure 4: Age of MTurk workers over time.

- Indian workers are substantially more likely to be male than U.S. Turkers.
- Indian workers have significantly lower annual incomes than U.S. workers.
**TurKit**

[Little et al., UIST 2010]

---

```javascript
// generate a description of X // and iterate it N times
var text = ""
for (var i = 0; i < N; i++) {
  // generate new text
  var newText = mturk.prompt("Please write/improve this paragraph describing " + X + ": " + text)

  // decide whether to keep it
  if (vote("Which describes " + X + " better?", [text, newText]) == newText) {
    text = newText
  }
}
```

---

**Iteration 1:** Lightening strike in a blue sky near a tree and a building.

**Iteration 2:** The image depicts a strike of fork lightening, striking a blue sky over a silhouetted building and trees. *(4/5 votes)*

**Iteration 3:** The image depicts a strike of fork lightning, against a blue sky with a few white clouds over a silhouetted building and trees. *(5/5 votes)*

**Iteration 4:** The image depicts a strike of fork lightning, against a blue sky—wonderful capture of the nature. *(1/5 votes)*

**Iteration 5:** This image shows a large white strike of lightning coming down from a blue sky with the tops of the trees and rooftop peaking from the bottom. *(3/5 votes)*

**Iteration 6:** This image shows a large white strike of lightning coming down from a blue sky with the silhouettes of tops of the trees and rooftop peaking from the bottom. The sky is a dark blue and the lightning is a contrasting bright white. The lightning has many arms of electricity coming off of it. *(4/5 votes)*

---

http://turkit-online.appspot.com/
In-class Exercise: Quality Control

- 30% of open ended tasks tend to yield unusable results
- “Lazy” Workers do only the minimum amount required
- “Eager Beavers” might try to do more than your algorithm can handle
- Exercise: In groups, come up with strategies how one can ensure quality of the delivered work
Quality Control Mechanisms

- **Verify the work**: Verification is often easier than the task
- **Output agreement**: Have multiple workers agree on answer
- **Reputation**: Mechanical Turk records approval rate
- **Economic models**: Pay more for honest workers
- **Defensive task design**: Is it easier to cheat than to be honest?
- **Statistical filtering**: Discard outliers from an expected distribution
- **Multilevel review**: One worker reviews the previous stage of work
• Word processor with cloud inside

• Plug-in to MS Word with different modules
  • Crowdproof: spelling and grammar checking by asking Mechanical Turk workers
  • Shortn: asks workers to suggests ways to shorten a given text
  • The Human Macro: any word processing task

• Embed human computation in an everyday application
  • Wizard-of-Oz prototyping as part of running system

• Achieves complex tasks that would require expert users otherwise

• Multilevel review: Find-fix-verify pattern
  • Splits task into a series of generation and review stages

[ Bernstein et al., UIST 2010 ]
Soylent is a prototype crowdsourced word processing interface. It focuses on three main tasks: shortening the user’s writing, proofreading […]
Problems and Open Questions:

• Latency, real-time answers?
• Sustainability and reliability?
• Ethical questions
• Privacy
• Ownership
• Personalization
• Impact on the worker
  • poor compensation
  • no team interaction, no learning from peers
  • no perception of context, pride in work

[O’Neill et al., CHI 2013]
Using Mechanical Turk for HCI Studies

• E.g. Harrison et al., 2011, Kineticons asked users to rate meaning of the moving icon

• Mechanical Turk offers the researcher a large user pool

• Quick and relatively inexpensive results

• But workers will try to game the system

• Insert verifiable questions before subjective questions to judge the worker’s honesty (E.g., describe the animation of the icon) [Kittur et al., 2008; Komarov et al. CHI 2013]

• Make an dishonest answer as hard as an honest one

• Used in design classes for user evaluations [Dow et al., CHI 2013]
Summary

- Humans can act as processor in Human Computation schemes to solve algorithms that CPUs will not solve.
- Human Computation can solve simple and complex tasks.
- Many options for quality control and incentives.
- HCI researchers can use Mechanical Turk for their user studies or teaching.