

Exploration Session 2: Human Computation

Melissa Winstanley
mwinst@cs.washington.edu

(based on slides by Roy McElmurry)

The Situation


- I am a spammer
- I want to send millions of spam emails
- I want to seem legitimate
- Several web services offer “legit” free email accounts
- I’ll create many of these accounts



The Problem

- Spammers want many email addresses
 - Use computers to automate signing up
- Email providers (Google, Microsoft, Yahoo) want real people, not spam accounts
- How to prevent computers from signing up?

THANK YOU!!

 **nhy nhy** [show details](#) Aug 4 (9 days ago) [Reply](#)

I am an Account Manager, A client of mine died in 2005 in a horrible boat mishap and he has a lodgment fund to claims US\$14.7M. And no other person knows about this account. I am contacting you because both of you have the same last name. If you are interested forward to me your full names, cell, Phone/fax, profession, age and country so I can contact you for further clarifications.
Thank You,
Mr.Kelly Khumalo

Possible Solutions

- What are most computers bad at?
 - Image recognition
 - Context recognition
 - Natural language processing
 - Character recognition

Character Recognition

1. Clear background noise
 2. Chop text into blocks, each one a character
 3. Classify the character in each region
- Computers aren't very good at this yet

CAPTCHA

- “Completely Automated Public Turing test to tell Computers and Humans Apart”
 - Distort a random string randomly
- Term from Carnegie Mellon University in 2000
- You can pass, a computer can't



spidsees



numencar



masterated

Humans are also good at wasting time

- 9 billion human-hours of solitaire in 2003
- Empire State Building build in 7 million human-hours
- Panama Canal built in 20 million human-hours
- What if we could harness wasted human power?

reCAPTCHA

- Luis von Ahn
 - When people identify a CAPTCHA, they do useful work
 - Don't waste that work
- Computers can't recognize some text with modern technology
- Use humans to do that work

Scanned type

This aged portion of society were distinguished from

OCR reads as

"niis aged pnkm at society were distinguished frow."

steamboat train, from New
this **morning** ran off the track
New-London. Four cars plunged!

morning

UPON

Type the two words:



reCAPTCHA™

stop spam.
read books.

Human Computation

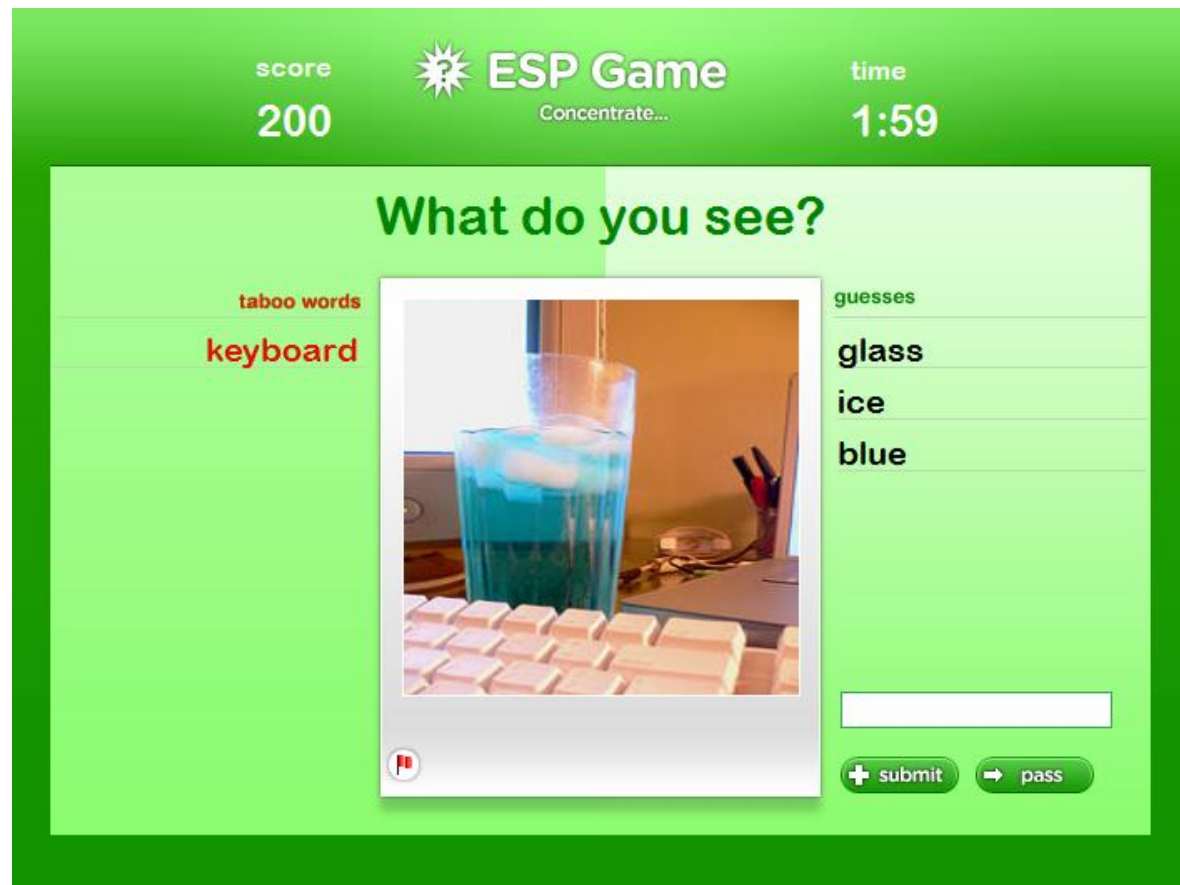
- Outsourcing some computations to people
- Balance cost of computation with cost of labor
- Helps solve otherwise unsolvable problems
- (Even spammers are learning: CAPTCHA sweat shops)

Another Idea: Games

- People like to play online games
- What if we could come up with a useful game?
- Luis von Ahn: “games with a purpose”

ESP Game

- Multiplayer
- Labeling images
- Taboo words for more variety
- Verify over many trials
- Acquired by Google to improve image search quality



GWAP

- ESP Game
 - Image labeling
- Tag-a-tune
 - Identify/label music
- Verbosity
 - Acquire common-sense knowledge
- Squigl
 - Trace an object within an image

FoldIt

- Game developed at UW in part by Zoran Popović
- Solve protein-folding problems from chemistry and biology
- Point system encourages minimizing free energy of a protein
- In multiple cases, folded proteins from the game beat biochemist predictions

Amazon Mechanical Turk



- People in need of human power come up with tasks
- People can complete these tasks for a specified price
- Affordable results for researchers
- “Artificial artificial intelligence”