A Survey of Human Computation Systems

Man-Ching Yuen[1], Ling-Jyh Chen[2], and Irwin King[1]

[1]Department of Computer Science and Engineering
The Chinese University of Hong Kong, Hong Kong

[2]Institute of Information Sciences
Academia Sinica, Taiwan

http://www.cse.cuhk.edu.hk/~king

Playing/Having Fun ⟷? Work/Computation

Idea of Human Computation

- Take advantage of people's desire to be entertained and perform useful tasks as a side effect

Motivations

- To describe the categorization of Human Computation Systems (HCS)
- To describe each category of HCS and present the previous work on each category
- To summarize the current state-of-the-art HCS

Outline

- Motivation and Background
- Types of Human Computation
  - Initiatory Human Computation
  - Distributed Human Computation
  - Social Game-based Human Computation with volunteers or paid engineers
  - Social Game-based Human Computation with online players
- Properties of Social Games
- Future Work and Final Remarks

Background

- Human Computation Systems (HCS) aim to solve Artificial Intelligence (AI) problems through the human human interactions

- In order to ensure the collected information to be useful, we have to:

1. guarantee the quality of collected information
2. attract more people to contribute information

Types of HCS

- The categories of the human computation systems are:

1. Initiatory Human Computation
2. Distributed Human Computation
3. Social Game-based Human Computation with volunteers or paid engineers
4. Social Game-based Human Computation with online players

Initiatory Human Computation (1)

- Objective: To complete some tasks that are natural for humans but difficult for computers even computation power increased rapid recently

- Example (1): CAPTCHA
  - A computer generated challenge-response test
  - Objective: To distinguish humans from computers using a common sense problem

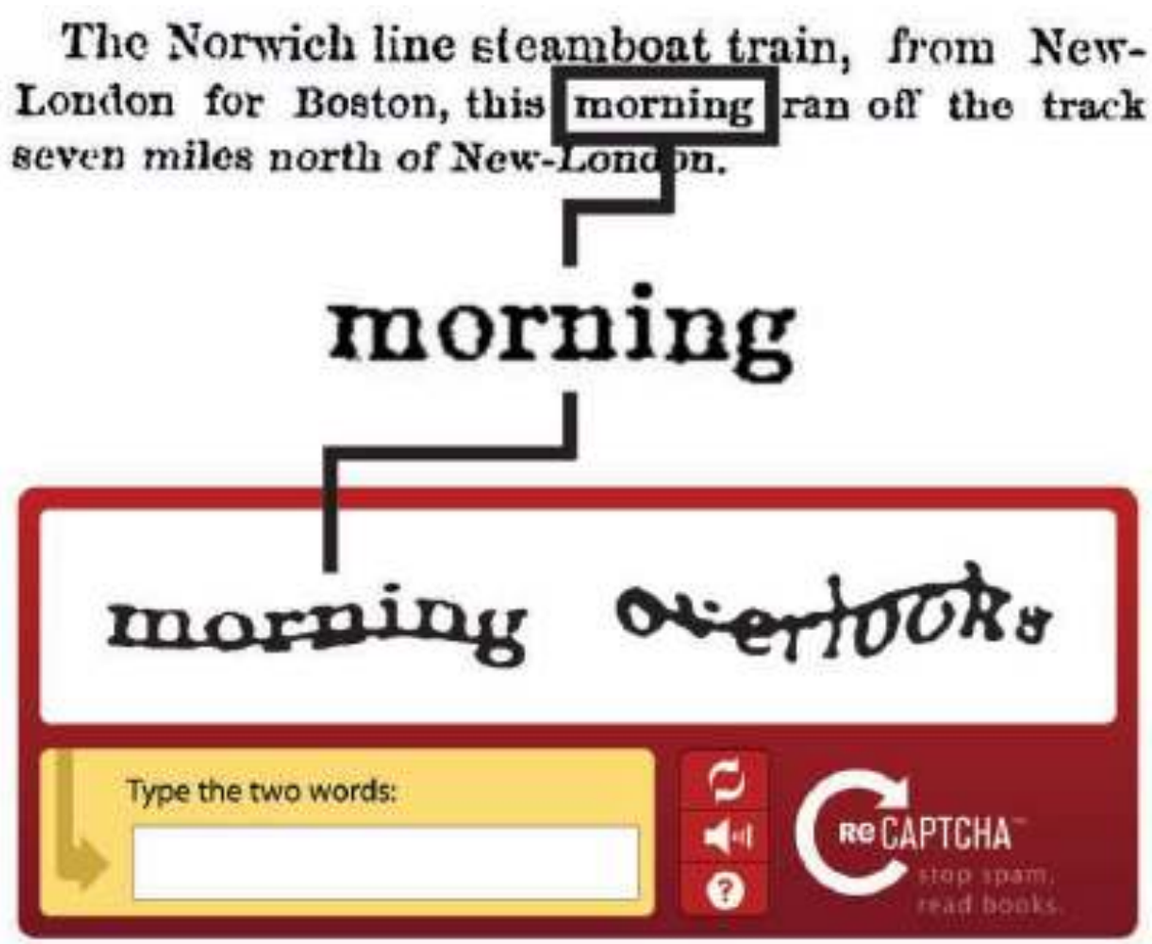
The Yahoo! CAPTCHA.

Initiatory Human Computation (2)

- Example (2): reCAPTCHA
  - Objective: To produce valuable common sense knowledge to improve the OCR quality in digitizing books
  - Combining two words: one identified word; and one unidentified word
  - If a user recognizes the identified word, the answer to the unidentified word is assumed to be correct

Initiatory Human Computation (3)

- Example (2): reCAPTCHA

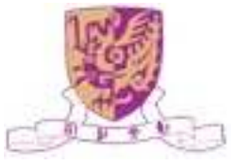


Initiatory Human Computation (4)

- Example (3): KA-CAPTCHA
  - Objective: To collect every correct answer submitted by humans to the CAPTCHA test as a solution to a problem that computers are unable to solve
  - CAPTCHA solvers are highly interested in providing a valid response to the CAPTCHA test (because they want to access the protected resource)
  - Knowledge acquisition mechanism: To strategically asking for a solution to a particular open problem that is of interest to the CAPTCHA designer.

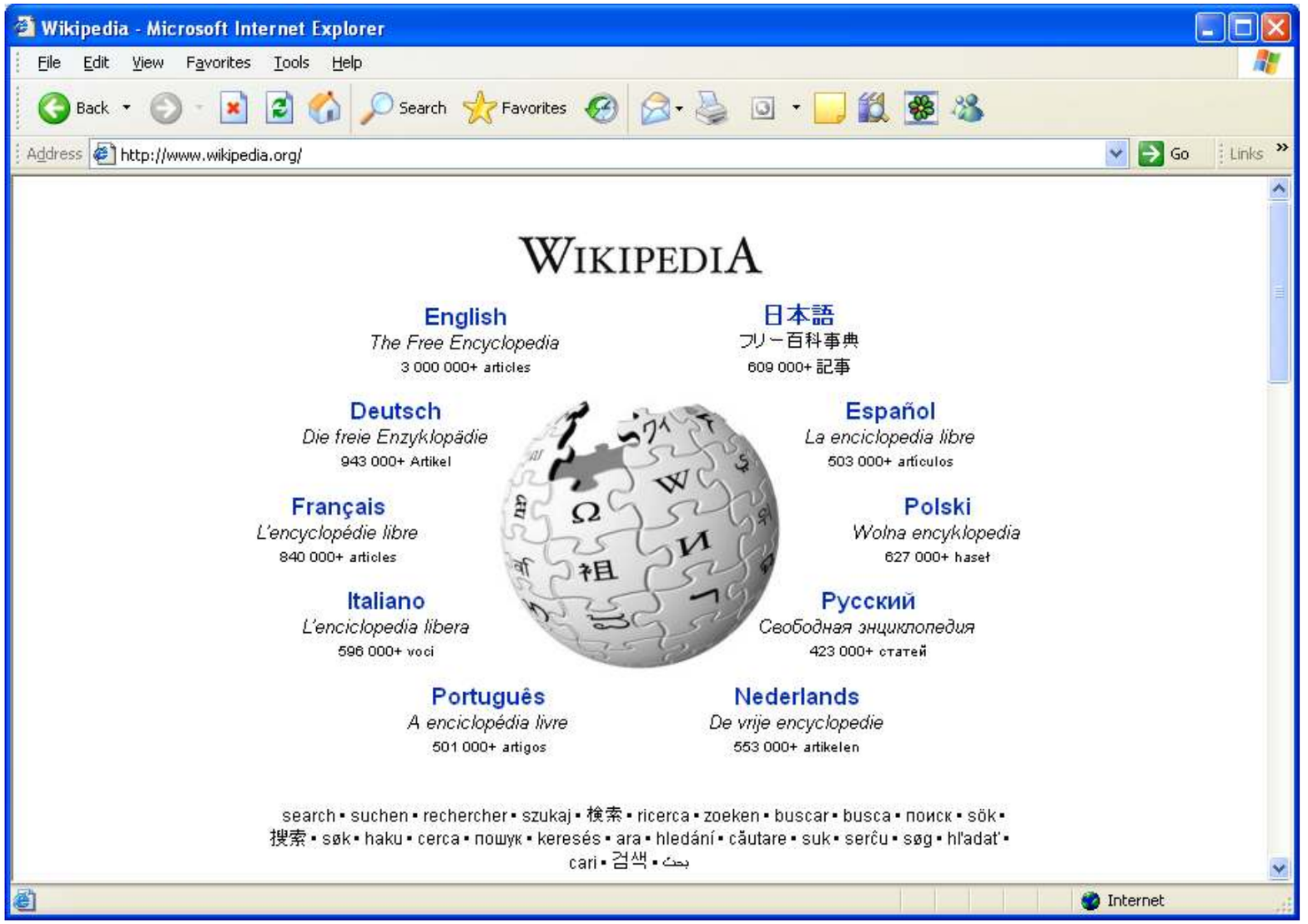
Distributed Human Computation (1)

- Objective: To encourage a huge population of Internet users to contribute to solve the difficult AI problems
- Example (1): Razor
  - To use human votes to determine if a given email is spam (anti-spam mechanism)
- Example (2): Proofreader
  - To give a (small) portion of the image file and corresponding text (generated by OCR) side-by-side to a human proofreader



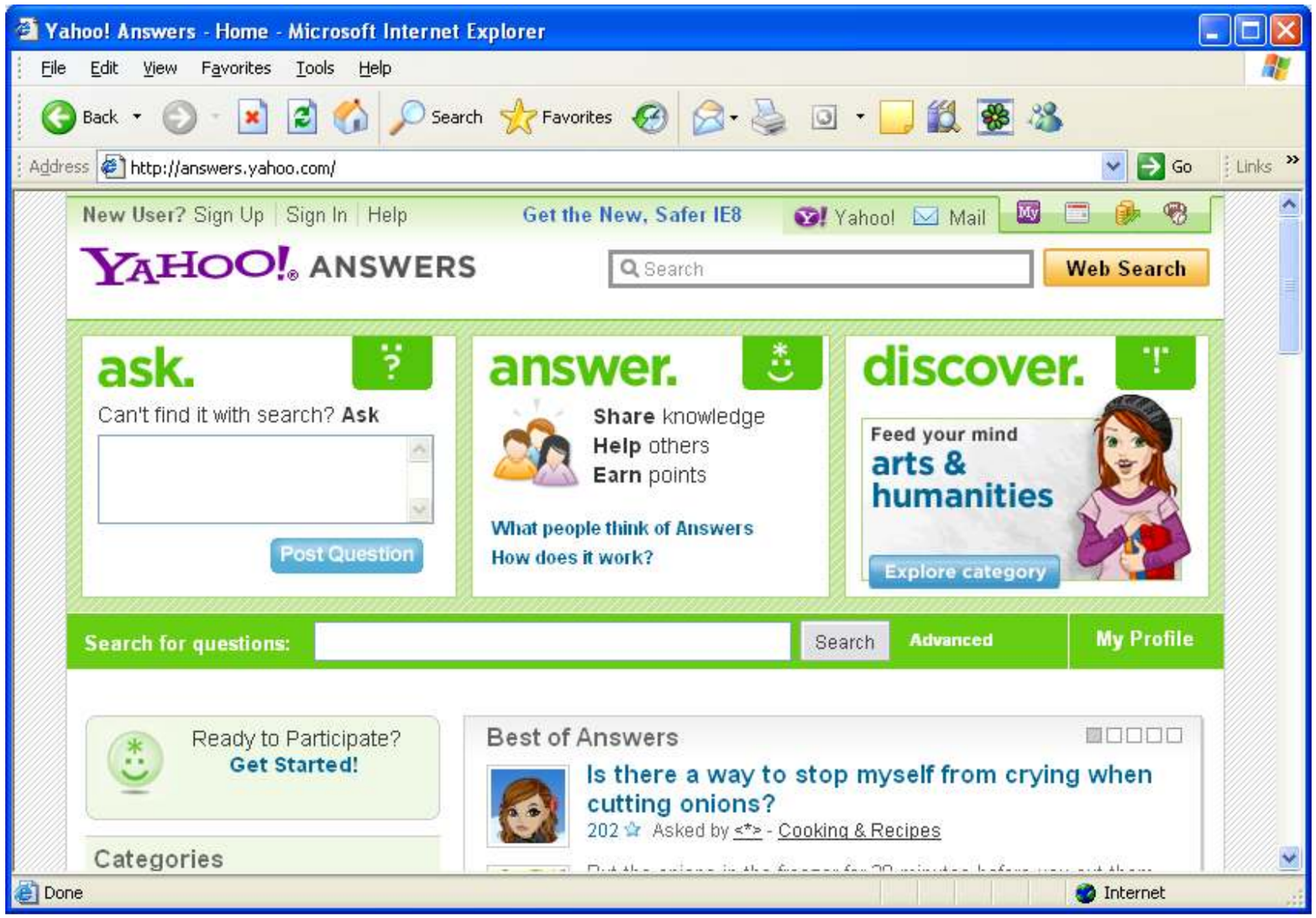
Distributed Human Computation (2)

- Example (3): Wikipedia
  - The collective knowledge is distributed in that essentially almost anyone can contribute to the Wiki



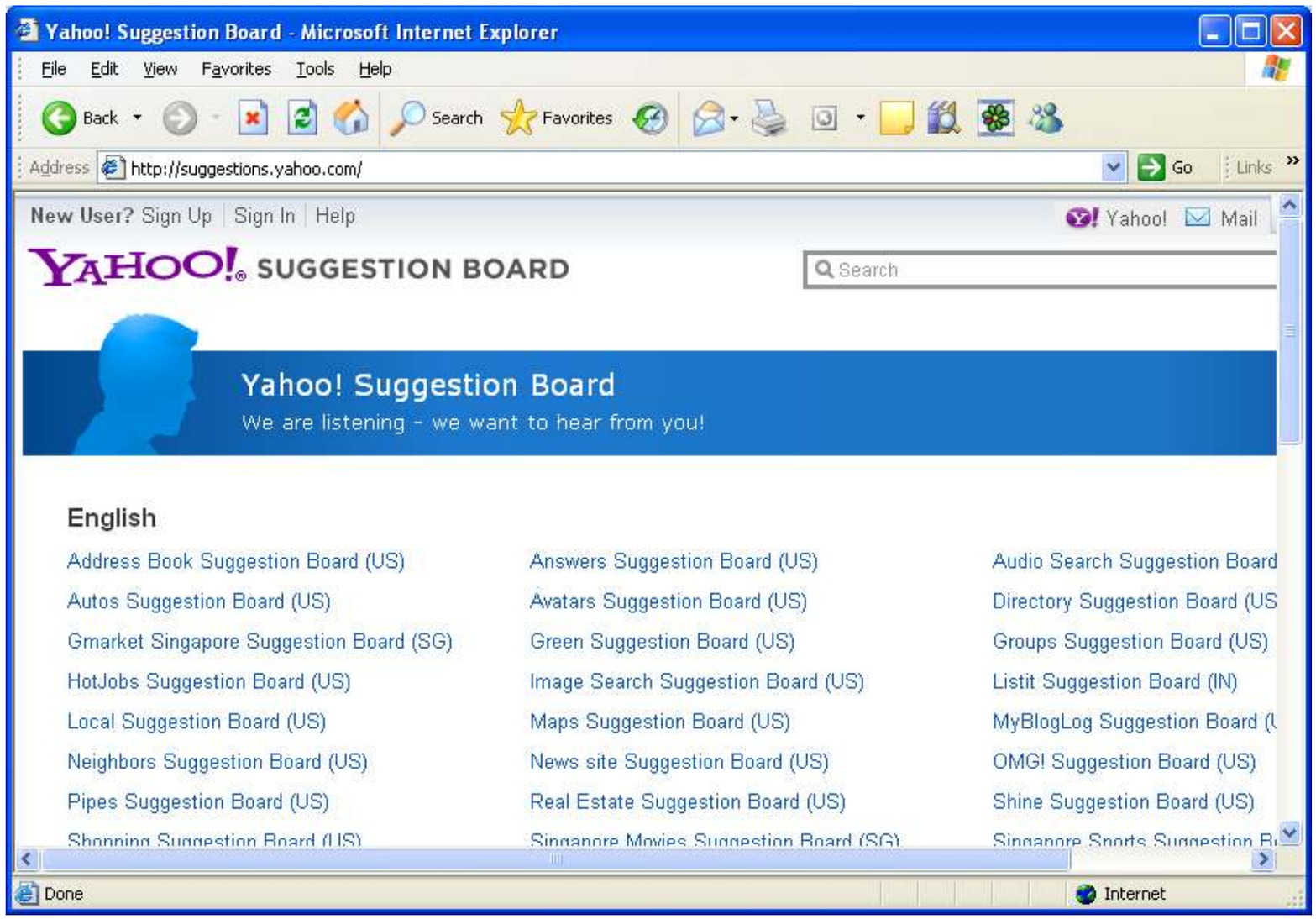
Distributed Human Computation (3)

- Example (4): Yahoo! Answers
  - To provide automated collection of human reviewed data at Internet-scale



Distributed Human Computation (4)

- Example (5): Yahoo! Suggestion Board
  - An Internet-scale feedback and suggestion system

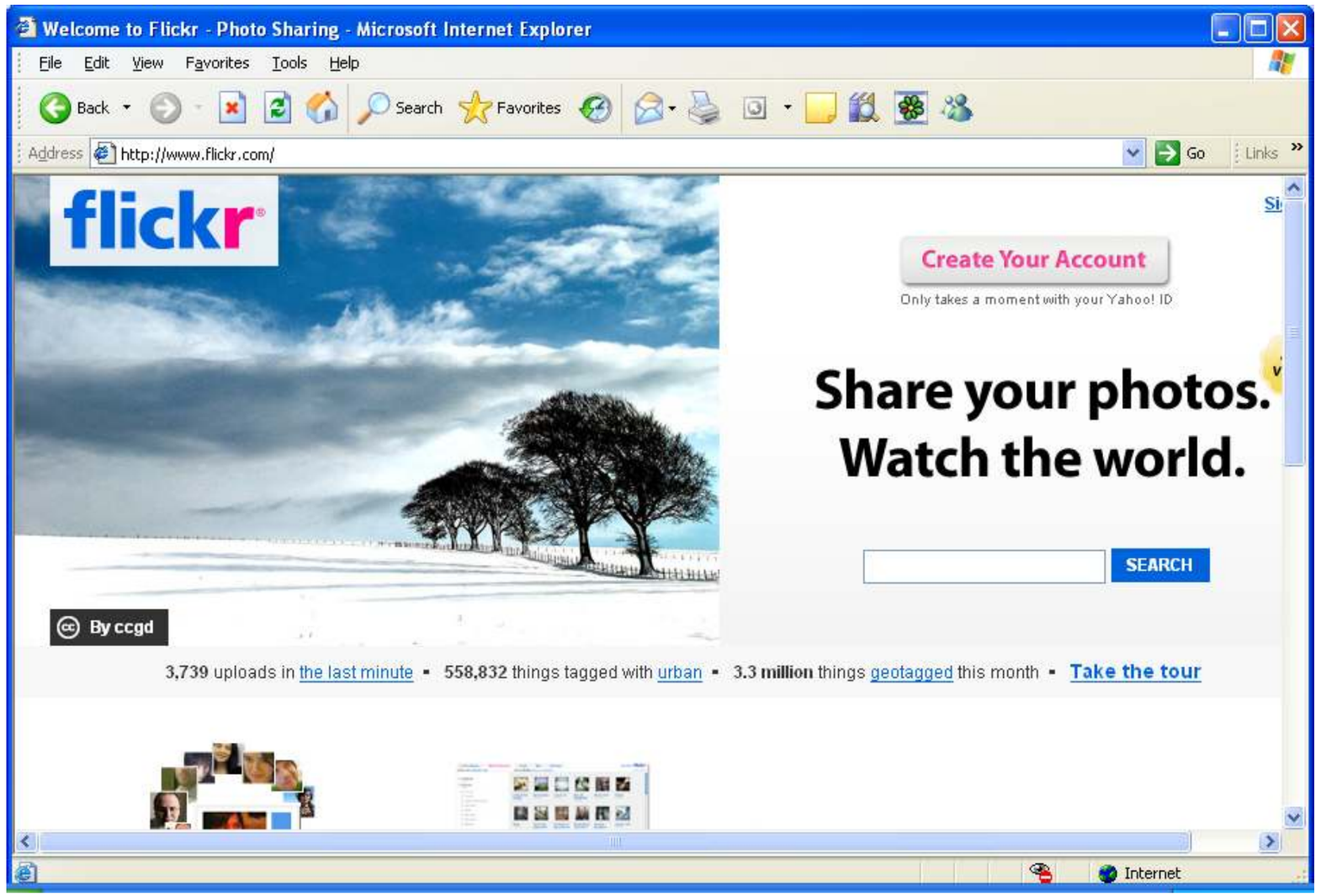


Distributed Human Computation (5)

- Example (6): Amazon Mechanical Turk
  - It provides monetary rewards for tasks
- Example (7): LabelMe
  - A web-based tool for image annotation
  - Anybody can annotate image using it. You can only have access to the database once you have annotated a certain number of images.
- Example (8): 43Things
  - To collect goals from users and help them to find other users who have similar goals
- Example 9: MajorMiner
  - Music annotation game

Distributed Human Computation (6)

- Example (10): Yahoo's flickr
  - It is a photo-sharing site with captions being used as photo tags



Social Game-based Human Computation with volunteers or paid engineers (1)

- Recently social games were proposed to collect accurate information from players as a side effect of their playing
- The players are volunteers or paid engineers
- Disadvantages:
  - Rely on online volunteers or paid engineers to enter information explicitly
  - Unable to scale up the system due to high cost
  - No validation mechanism to guarantee that the information collected is accurate

Social Game-based Human Computation with volunteers or paid engineers (2)

- Most of the games at early stage aimed to collect commonsense knowledge.
- Example (1): Cyc
  - To collect information from the input by paid knowledge engineers
- Example (2): Open Mind
  - To collect common sense knowledge from people to develop intelligent software
  - Shortcoming: was too reliant on the unpaid volunteers to donate their time to contribute information

Social Game-based Human Computation with volunteers or paid engineers (3)

- Example (2): Open Mind

Social Game-based Human Computation with volunteers or paid engineers (4)

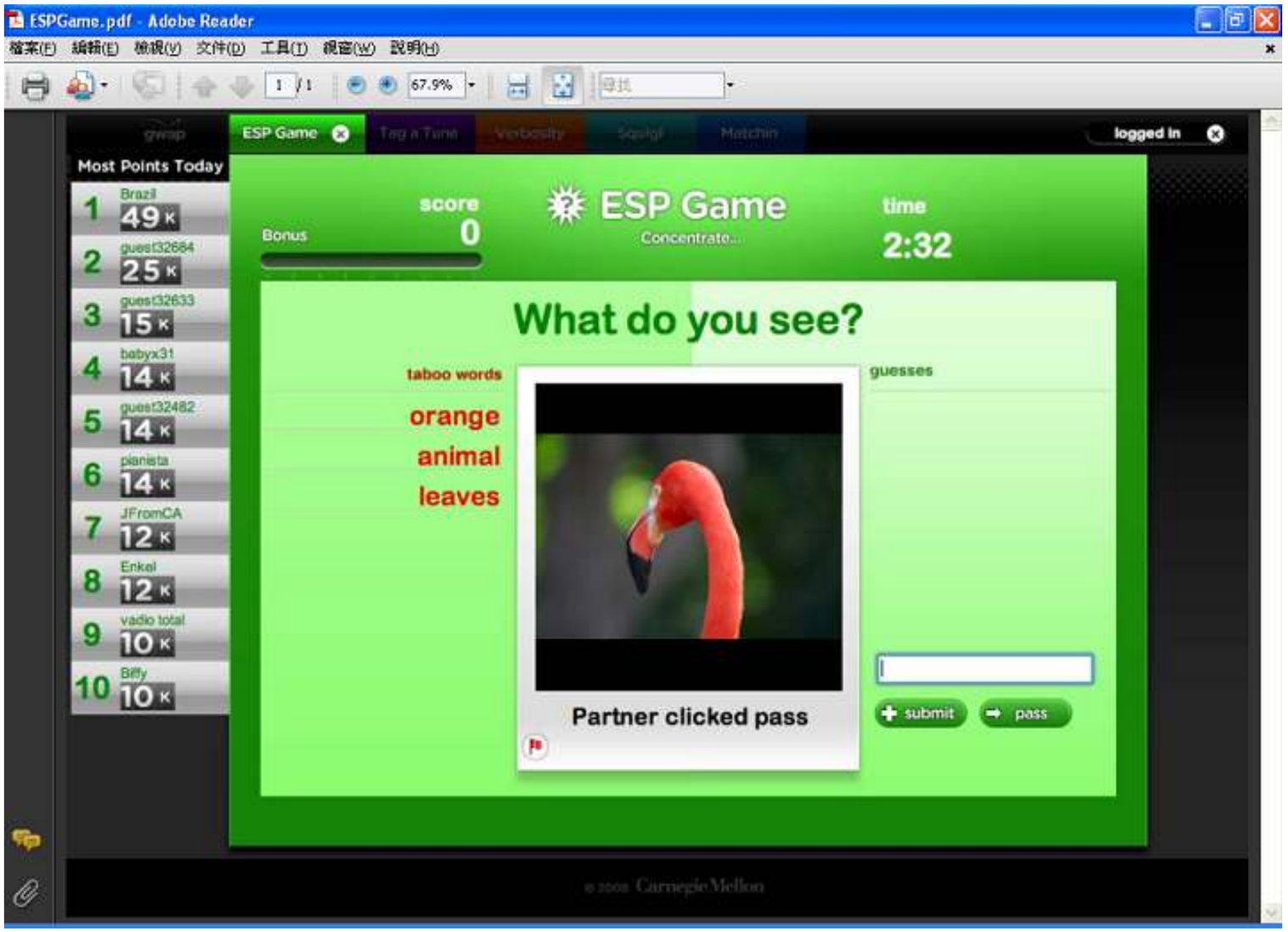
- Example (3): Mindpixel
  - Reward those Internet users who consistently validate a fact inline with the other users
  - Shortcoming: the cost is high!
- Example (4): Wildfire wally
  - To solve the maximum clique problem
  - Shortcoming: rely on unpaid volunteers to donate their time to contribute information

Social Game-based Human Computation with online players (1)

- Later, social games were proposed to collect information from the players as a side effect of their playing
- Advantage:
  - It encouraged more Internet users to contribute information to solve the AI problems because of the increasingly popularity of online game
- TWO important factors for collecting information effectively from players through a social game:
  - Guarantee the quality of collected information
  - Maintain the enjoyment of players in the game

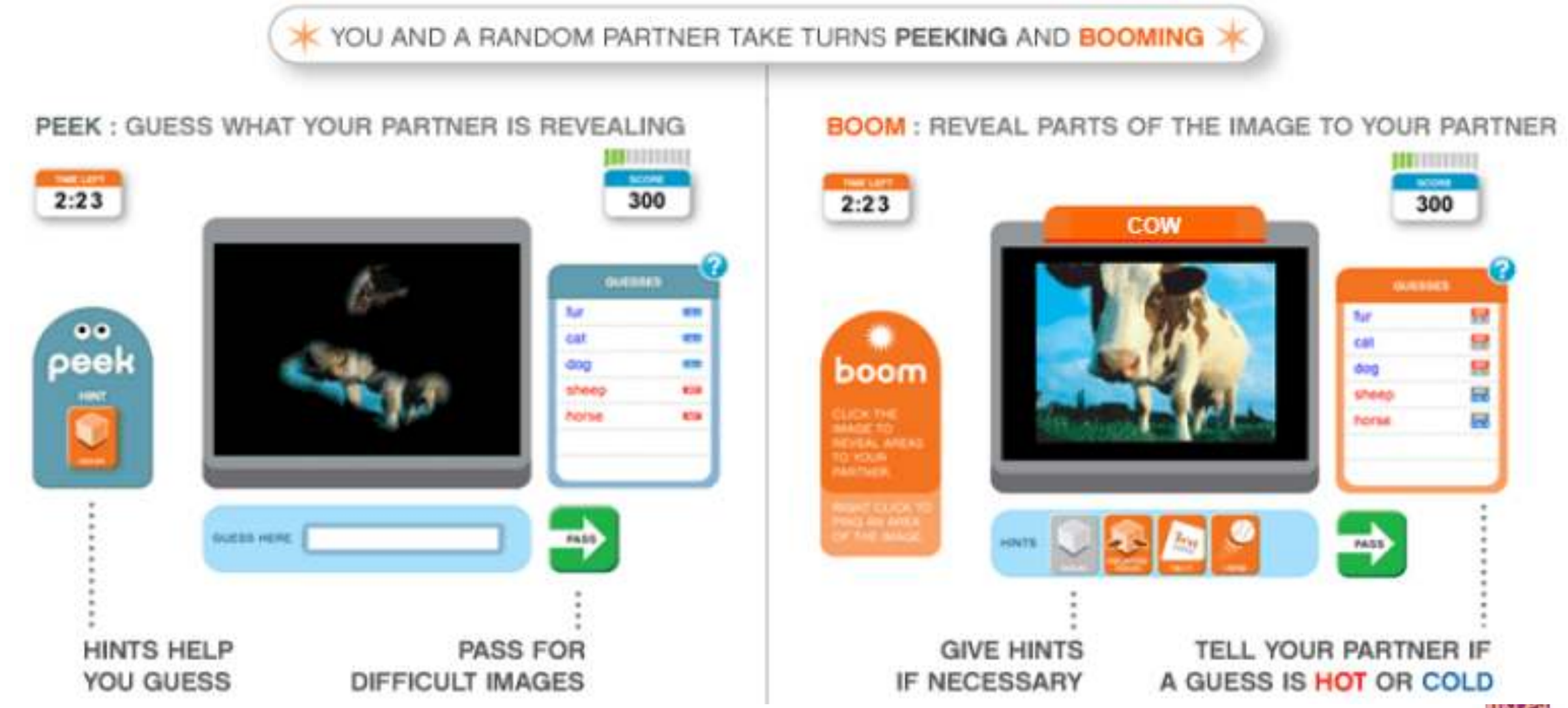
Social Game-based Human Computation with online players (2)

- To collect text information from images
  - Examples (1): ESP game



Social Game-based Human Computation with online players (3)

- To collect text information for images:
  - Examples (2): Peekaboom



Social Game-based Human Computation with online players (4)

- To collect commonsense knowledge:
  - Examples (3): Verbosity

Figure 1. Part of the Narrator's screen.

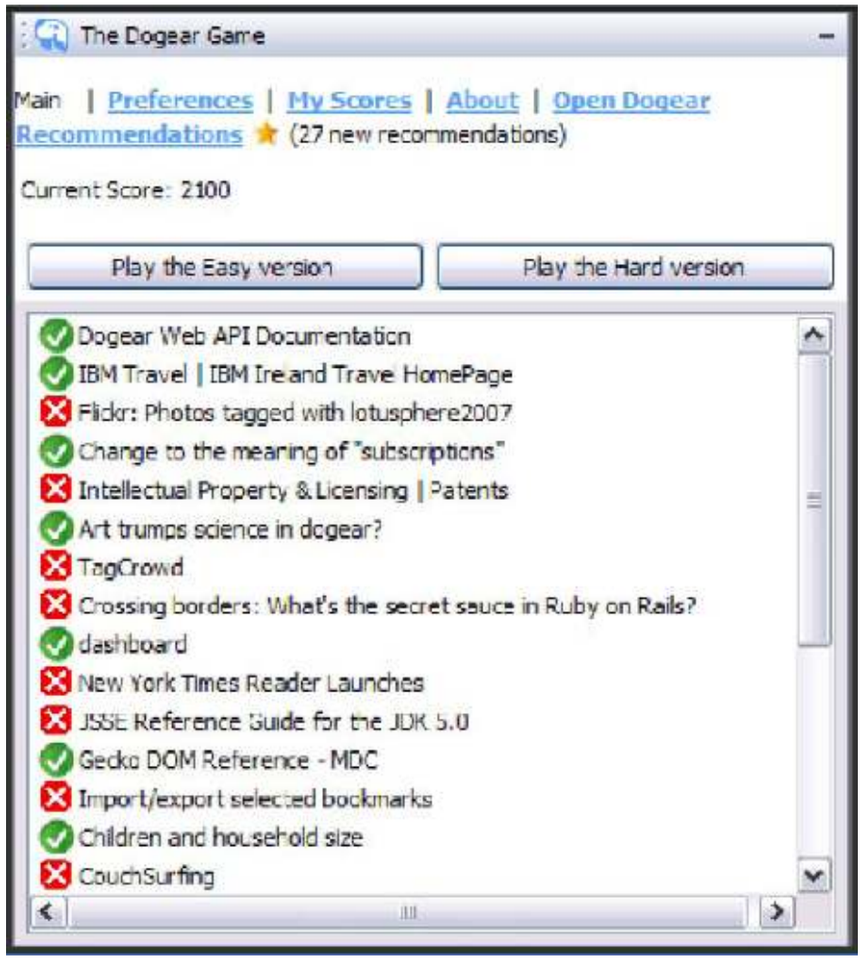
A Survey of Human Computation Systems, Irwin King, SCA2009, August 30, 2009, Vancouver, Canada

Social Game-based Human Computation with online players (5)

- To collect subjective descriptions of sounds and music:
  - Example (4): Tagatune

Social Game-based Human Computation with online players (6)

- To learn colleagues' bookmarks in an organizational goal:
  - Example (5): Dogear Game



Social Game-based Human Computation with online players (7)

- To tag locations in the real world through gameplay in mobile social games:
  - Example (6): Gopher guessing game

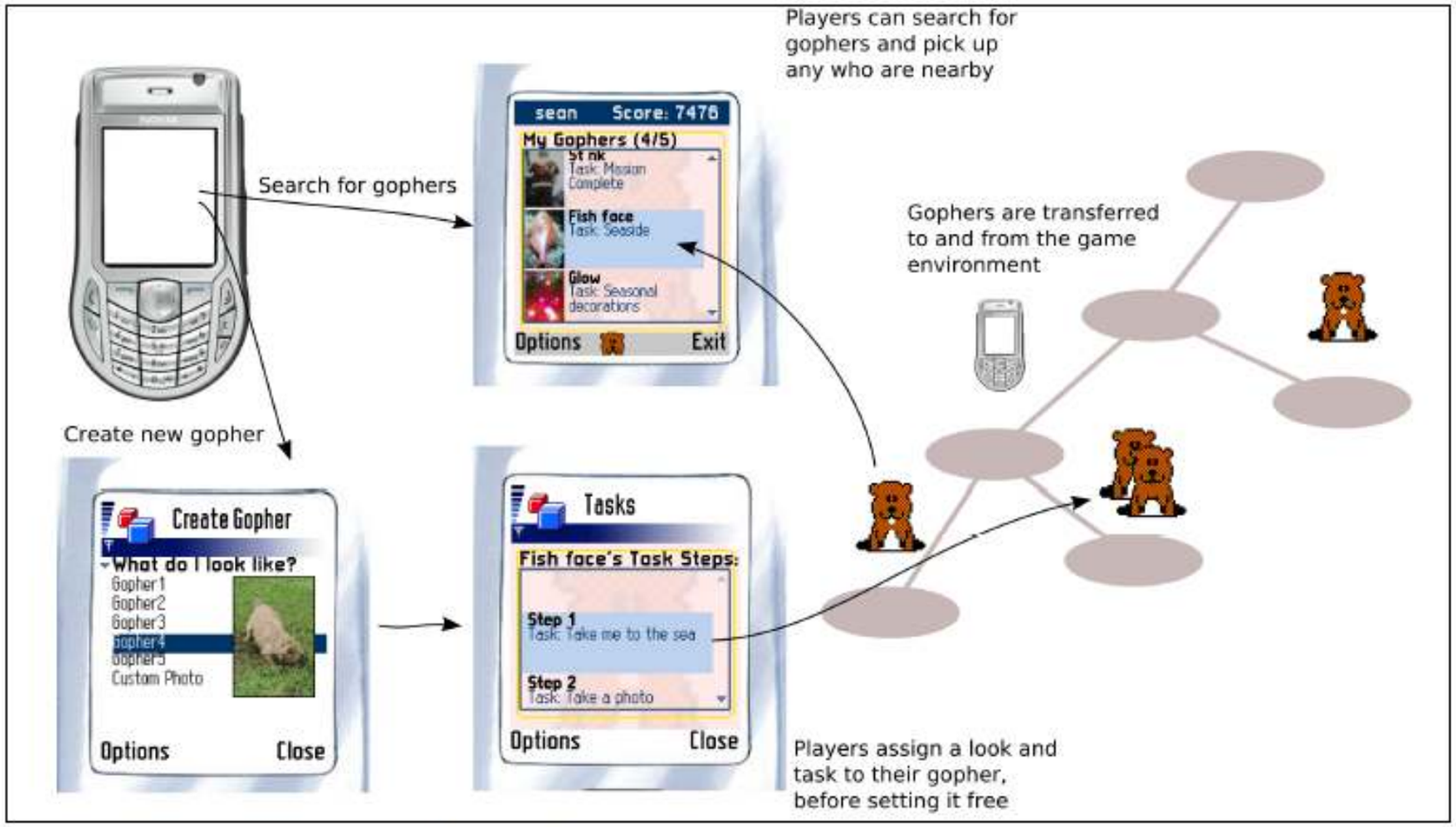


Figure 1. Real world experience, acquiring gophers

Social Game-based Human Computation with online players (8)

- To tag locations in the real world through gameplay in mobile social games:
  - Example (7): Gopher guessing game

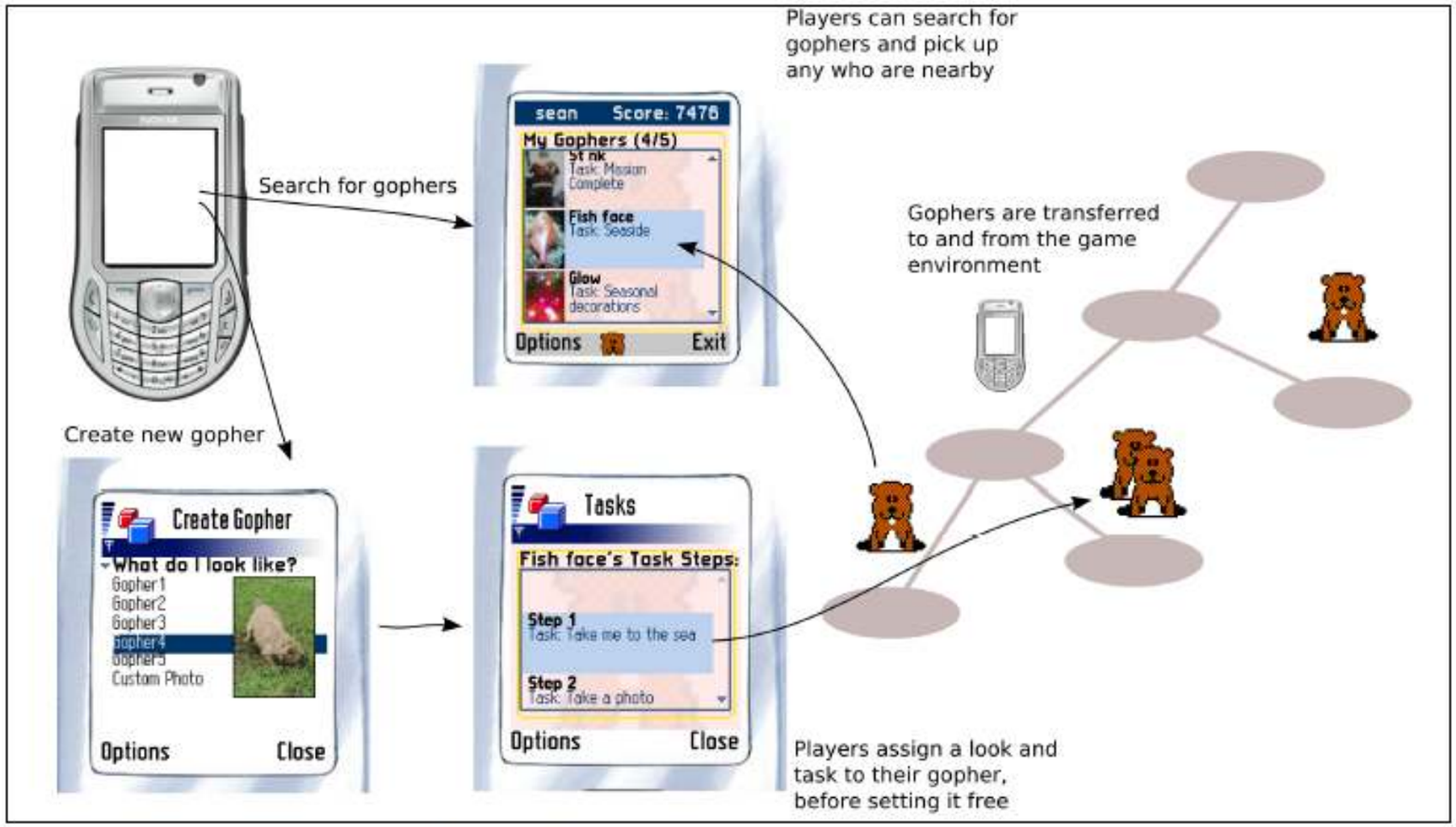
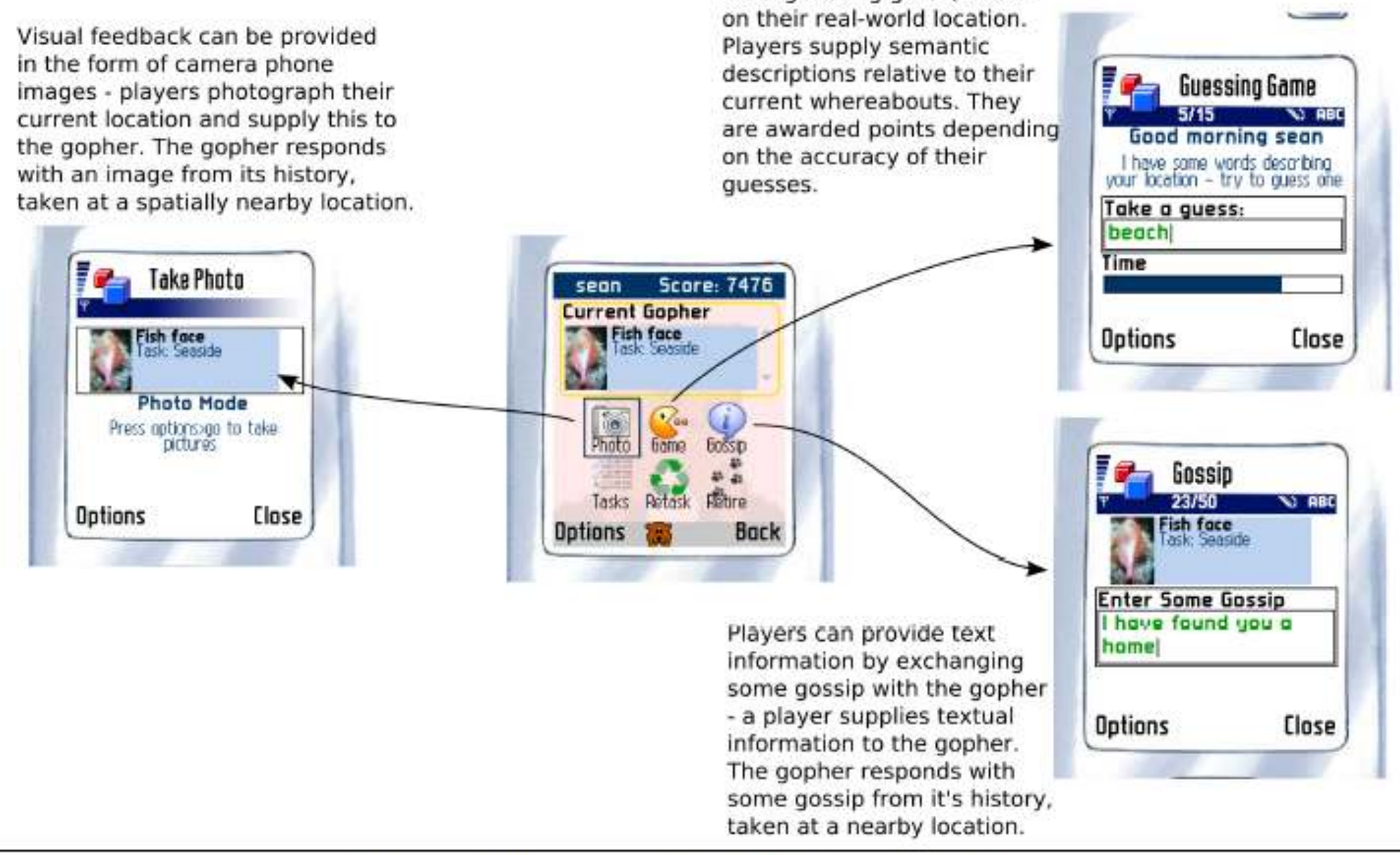


Figure 1. Real world experience, acquiring gophers

Visual feedback can be provided in the form of camera phone images - players photograph their current location and supply this to the gopher. The gopher responds with an image from its history, taken at a spatially nearby location.

Gophers can participate in a word guessing game, based on their real-world location. Players supply semantic descriptions relative to their current whereabouts. They are awarded points depending on the accuracy of their guesses.



Players can provide text information by exchanging some gossip with the gopher - a player supplies textual information to the gopher. The gopher responds with some gossip from it's history, taken at a nearby location.

Figure 2. Real world experience, interacting with gophers

Properties of Social Games

1. Type of information to be collected
2. Game Structure
    1. Output-agreement Game
    2. Input-agreement Game
    3. Inversion-problem Game
    4. Output-optimization Game
3. Verification Method
    1. Symmetric
    2. Asymmetric
4. Game Mechanism
    1. Collaborative
    2. Competitive
    3. Hybrid
5. Player Requirement

Categorization of Social Games

TABLE I
CATEGORIZATION OF SOCIAL GAMES

Game Structure	Verification Method	Game Mechanism
Output-agreement	Symmetric	Collaborative or Hybrid
Input-agreement	Symmetric	Collaborative or Hybrid
Inversion-problem	Asymmetric	Collaborative or Competitive or Hybrid
Output-optimization	Symmetric or Asymmetric	Collaborative or Competitive or Hybrid

Subjective vs. Objective Information

- For subjective information, the information presented for the same subject is affected by users because of different choices of vocabularies for the same subject.
  - lower probability on players' correct outputs being the same
- For objective information, the information presented for the same subject is NOT affected by users because of same choices of vocabularies for the same subject.
  - higher probability on players' correct outputs being the same

Game Structure (1)

- Game structure defines the key elements of a game including players' input, players' output, the relationship among the input and output of players, and the winning condition

- Four types of game structure
  1. Output-agreement Game
  2. Input-agreement Game
  3. Inversion-problem Game
  4. Output-optimization Game

Game Structure (2)

- Output-agreement Games: All players are given the same input and must produce outputs based on the common input
  - An output-agreement game should be used to collect objective information
- Input-agreement Games: All players are given inputs that are known by the game (but not by the players) to be the same or different. The players are instructed to produce outputs describing their input, so their partners are able to assess whether their inputs are the same or different. Players see only each other's outputs
  - An input-agreement game should be used to collect subjective information

Game Structure (3)

- Inversion-problem Games: The first player has access to the whole problem and gives hints to the second player to make a guess. If the second player is able to guess the secret, we assume that the hints given by the first player are correct.

- Output-optimization Games: All players are given the same input and their outputs are the hints of other players' outputs.

  - An output-optimization game should be used to collect subjective information, because the output pattern of players reflects outputs of players are strongly affected by others' outputs. It is subjective.

Verification Methods

- Verification method of a game defines the method to check the output accuracy of players by asking players to do the same task or different tasks
- Symmetric Verification Games: Either an output-agreement game or an input-agreement game is symmetric verification
- Asymmetric Verification Games: Players are assigned to one of the roles to do different tasks

Game Mechanism

- Game mechanism defines the relationship of all players in the game in order to achieve the winning condition
- Collaborative Games determine the winning condition of all players. The accuracy of output is guaranteed by collaboration of all players.
- Competitive Games determine the winning condition of a player. Output accuracy is guaranteed by information stored in a database. Players' enjoyment in the game can be increased in competition.
- Hybrid Game

Player Requirements (1)

- Player requirement defines the rules on accessing the game of all players.
- In Synchronous Games, players have to give real-time response to other players' action.
- In Asynchronous Games, players do not have to give real-time response to other players' action. The information collected from one player is stored in a database and will be used to determine the correctness of other players' output.

Player Requirements (2)

- Number of players define the following types:
- Single-player Games: It allows one player to play and the other's moves can be simulated from the prerecorded game. Only inversion-problem game can be a single-player game.
- Two-player Games: It allows two players to play together.
- Multi-player Games: It allows multiple players to play together. Only hybrid games can be a multi-player game.

Summary

TABLE II
CATEGORIZATION OF SOCIAL GAMES WITH EXAMPLES

Game Structure	Verification Method	Game Mechanism	Player Requirement		Examples
			Num of Player	Game Play	
Output-agreement	Symmetric	Collaborative	2	Synchronous	ESP, Matchi, Squigl, OntoGame
		Hybrid	Multi-players	Synchronous	Common Consensus, Social Heroes
		Hybrid	Multi-players	Asynchronous	Gopher Game
Input-agreement	Symmetric	Collaborative	2	Synchronous	TagATune
		Hybrid	N/A	N/A	N/A
Inversion-problem	Asymmetric	Collaborative	1 or 2	Synchronous	Peekaboom, Verbosity
		Competitive	2	Asynchronous	Dogear, CyPRESS, CARS
		Hybrid	1 or Multi-players	Synchronous	Phetch
Output-optimization	Symmetric	Collaborative	2	Synchronous	Restaurant Game
		Competitive	N/A	N/A	N/A
		Hybrid	Multi-players	Synchronous	Diplomacy

Final Remarks

- Future Work
  - Models, theories, etc.
  - Tools, platforms, etc.
  - Performance metrics, e.g., accuracy, complexity, etc.
- To provide a better understanding about Human Computation Systems (HCS) systematically
- To facilitate future research activities in the field of HCS

Q & A