Games with a Purpose

Luis von Ahn
Carnegie Mellon University

Through online games, people can collectively solve large-scale computational problems.

Each year, people around the world spend billions of hours playing computer games. What if all this time and energy could be channeled into useful work? What if people playing computer games could, without consciously doing so, simultaneously solve large-scale problems?

Despite colossal advances over the past 50 years, computers still don't possess the basic conceptual intelligence or perceptual capabilities that most humans take for granted. If we treat human brains as processors in a distributed system, each can perform a small part of a massive computation. Such a "human computation" paradigm has enormous potential to address problems that computers can't yet tackle on their own and eventually teach computers many of these human talents.

Unlike computer processors, humans require some incentive to become part of a collective computation. Online games are a seductive method for encouraging people to participate in the process. Such games constitute a general mechanism for using brain power to solve open problems.

In fact, designing such a game is much like designing an algorithm—it must be proven correct, its efficiency can be analyzed, a more efficient version can supersede a less efficient one, and so on. Instead of using a silicon processor, these "algorithms" run on a processor consisting of ordinary humans interacting with computers over the Internet.

"Games with a purpose" have a vast range of applications in areas as diverse as security, computer vision, Internet accessibility, adult content filtering, and Internet search. Two such games under development at Carnegie Mellon University, the *ESP Game* and *Peekaboom*, demonstrate how humans, as they play, can solve problems that computers can't yet solve.

LABELING RANDOM IMAGES

Several important online applications such as search engines and accessibility programs for the visually impaired require accurate image descriptions. However, there are no guidelines about providing appropriate textual descriptions for the millions of images on the Web, and computer vision can't yet accurately determine their content.

Current techniques used to categorize images for these applications are inadequate, largely because they assume that image content on a Web page is related to adjacent text. Unfortunately, the text near an image is often scarce or misleading and can be hard to process.

Manual labeling is traditionally the only method for obtaining precise image descriptions, but this tedious and labor-intensive process is extremely costly.

The *ESP Game* (www.espgame.org) accomplishes the same task through a simple online game that randomly pairs players together. Players don't know who their partner is, nor can they communicate with each other. The only thing partners have in common is an image they can both see.

The games' goal is to guess what label your partner would give to the image. Players type a word or phrase and then press the enter key to submit it to the game. Once both players have typed the exact same string, a new image appears; they don't have to type the string at the same time, but each must type the same string at some point while the image is onscreen.

Players can submit as many words or phrases as they want; in fact, the more strings they submit, the better their chance of getting a match.

The process of typing the same string is called "agreeing on an image," illustrated in Figure 1. Partners strive to agree on as many images as they can up to a total of 15 in two and one-half minutes, receiving a certain number of points for each match as well as a bonus for matching all 15. A meter at the bottom of the screen indicates the partners' progress.

Some images have "taboo words" that players can't use; the more taboo words an image has, the more points a string match is worth. Players can also choose to pass on difficult images.

To increase the chances of coming up with the same string for images over the course of a game, partners must learn to "think like each other"—thus the name ESP. It turns out that the string on which the two players agree is typically a good label for the image.

The *ESP Game* is extremely popular, with many people playing more than 40 hours per week. Within a few months of initial deployment on 25 October 2003, the game collected more than 10 million image labels; if hosted on a major site like MSN Games or Yahoo Games, all images on the Web could be labeled in a matter of weeks.

LOCATING OBJECTS IN IMAGES

While the *ESP Game* can determine what objects are in an image, it can't determine where in the image each object is. Such location information is necessary for training and testing computer vision algorithms.

The online game *Peekaboom* (www.peekaboom.org) improves on the data collected by the *ESP Game* by obtaining precise location information for each object in a given image. More specifically, it identifies which pixels belong to which object in the image.

In this game, two randomly paired players are assigned the roles of "Peek" and "Boom." Peek starts with a blank screen while Boom sees an image and a related word, as shown in Figure 2. All image-word pairs in *Peekaboom* come directly from the *ESP Game*.

Peek's goal is to guess the associated word as Boom slowly reveals the image. Each time Boom clicks on the image, a circular area in a 20-pixel radius around that click appears to Peek, who can guess what the word is by typing it in the box below the image.

Boom can see Peek's guesses and indicate whether they are "hot" or "cold." To help Peek identify the word, Boom can also "ping" or point to part of the image as well as provide hints indicating whether the word is a verb or refers to a noun in the image, a noun related to the image, or text in the image.

When Peek correctly guesses the word, the players receive a certain number of points—using the hint buttons adds points—and then switch "booming" and "peeking" roles using a new image-word pair.

Players have four minutes to go through as many combinations as possible. As in the *ESP Game*, players can pass on difficult images.

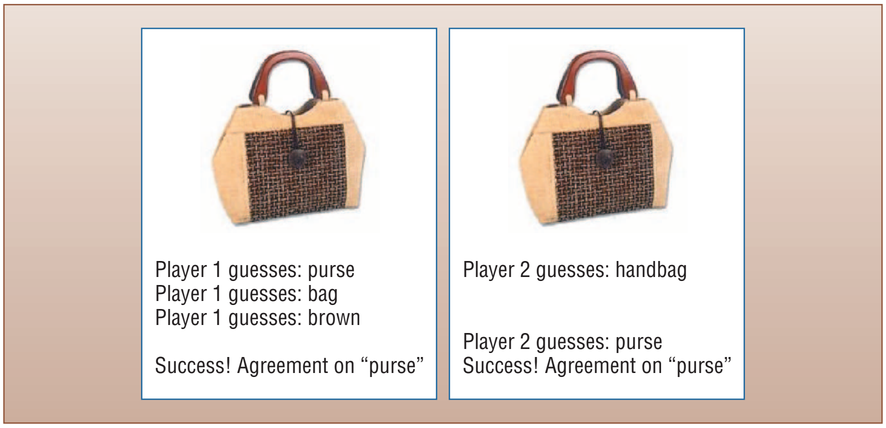

Figure 1. Partners agreeing on an image in the ESP Game. *Neither player can see the other's guesses.*

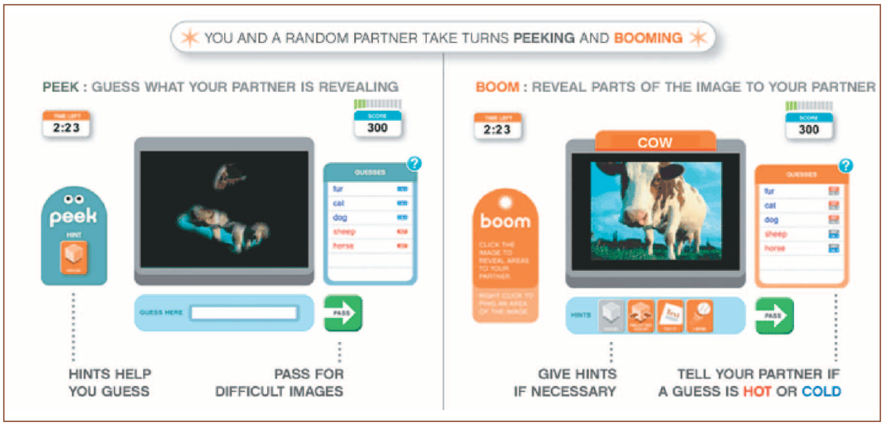

Figure 2. Peekaboom. *"Peek" tries to guess the word associated with an image slowly revealed by "Boom."*

Intuitively, to maximize points, Boom has an incentive to reveal only the areas of the image necessary for Peek to guess the correct word. For example, if the image contains a car and a dog and the associated word is "dog," Boom would reveal only those parts of the image that contain the dog. Thus, for a given image-word pair, data from multiple games yields the area of the image pertaining to the word.

Since the release of *Peekaboom* to a general audience on 1 August 2005, nearly 30,000 different people have played the game, generating roughly 2 million pieces of data—a "piece of data" is a successful round of *Peekaboom* in which Peek correctly guessed the word given Boom's revealed region.

POTENTIAL APPLICATIONS FOR FUTURE GAMES

Many other large-scale open problems can be solved using collective human brainpower in this unique way. Examples include:

- *Language translation.* A game could challenge two players who don't speak the same language to translate text from one language to the other.
- *Monitoring of security cameras.* The rapidly decreasing cost of digital video cameras is making it feasible to install security cameras almost everywhere. In the context of a game, players could monitor such cameras and alert authorities about suspected illegal activity.
- *Improving Web search.* People have varying degrees of skill at

searching for information on the Web. A game could be designed in which the players perform searches for other people.

- *Text summarization.* Imagine a game in which people summarize important documents for the rest of the world. Solving this problem would require an intelligent way to break up such documents into small "bite-size" chunks.

Any game designed to address these and other problems must ensure that game play results in a correct solution and, at the same time, is enjoyable. People will play such games to be entertained, not to solve a problem—no matter how laudable the objective.

For the first time in human history, hundreds of millions of people can, via the Internet, easily collaborate on the same problem. At CMU, we continue working on ways to solve complex computational challenges through the medium of online entertainment. Two additional "games with a purpose" under development include *Phetch*, which annotates images with descriptive paragraphs, and *Verbosity*, which collects commonsense facts to train reasoning algorithms. Log in soon!

***Luis von Ahn** is a postdoctoral associate at Carnegie Mellon University's Center for Algorithm Adaptation, Dissemination, and Integration (ALADDIN). Contact him at biglou@cs.cmu.edu.*

Editor: Bill Schilit, Intel Research Seattle; bill.schilit@intel.com