In celebration of CS Education Week 2011, NSF is rolling out CS Bits & Bytes, a one-page biweekly newsletter for the classroom highlighting innovative computer science research; sign up at: www.nsf.gov/cise/csbytes/.

December 5, 2011

Volume 1, Issue 1

Human Computation

Did you know?! When you sign up for an email account or buy tickets to a concert online, you may be helping to digitize text from libraries across the world? In fact, people just like you helped to digitize 20 years of The New York Times in less than three months. Digitizing text allows computers to search all of the information in a written document, making it easier to find what you want.

MUST SEE!

http://www.pbs.org/wgbh/nova/tech/profile-von-ahn.html
Credit: WGBH Educational Foundation

You help out whenever you are asked by a website to decode more than one sequence of squiggly, distorted characters. **Human or computer?** One of the squiggly words is used as a security measure to make sure that a response is generated by a real person, rather than by automated software (people are good at decoding these things, computers are not). This practice is called "CAPTCHA" which stands for **C**ompletely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part. The other word is a "reCAPTCHA" and is an image of a word from text that automated software was unable to recognize. Lots of people decipher the reCAPTCHA and the word is translated by the wisdom of the crowd.

Making use of the wisdom of a crowd. Digitizing text is just one example of a problem that can be solved by harnessing the combined power of humans and computers that would be impossible for either to solve alone. This field of "Human Computation," often called "Crowdsourcing," was pioneered by Professor Luis von Ahn of Carnegie Mellon University. His work uses human skills and abilities in a novel way to solve large-scale computational problems.

Create your own reCaptcha at: https://www.google.com/recaptcha/admin/create.
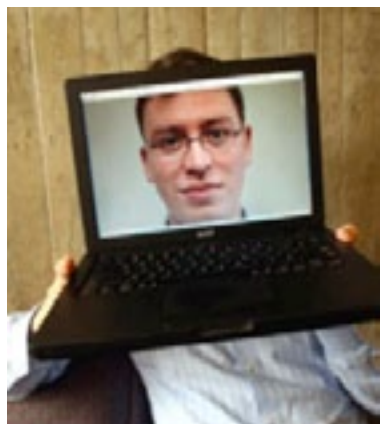

Image of Professor Luis von Ahn.

In addition to creating reCAPTCHAs, Professor von Ahn has also developed a number of "Games With A Purpose," or "GWAPs." These captivating games, available at http://www.gwap.com, pair random players to collect valuable information for training computer algorithms as a side effect of game play. Because the two players get points when their answers match, the accuracy and the fun increase with each game!

The basic idea relies on the wisdom of a crowd: if multiple people agree on a solution, then that solution is probably correct. A popular example is the ESP Game (http://www.gwap.com/gwap/gamesPreview/espgame/), where players are shown the same image and must independently generate tags; tags that match become labels for the image. ESP Game players have generated millions of labels that help improve image search engines.

Who thinks of this stuff? Professor von Ahn grew up in Guatemala, went to college at Duke University, and received a doctorate from Carnegie Mellon University. In his spare time, Luis likes to watch TV, play video games, and drive his Porsche around the streets of Pittsburgh, Pennsylvania. He sold reCaptcha to Google and continues to work for Google on new ideas while teaching at Carnegie Mellon University.

Links:

To learn more about reCaptchas, go to: http://www.google.com/recaptcha.

To see how Professor von Ahn describes reCaptcha, go to: http://googleblog.blogspot.com/2009/09/teaching-computers-to-read-google.html.

Activities:

Students can realize their abilities to be a part of the human computation crowd very easily. Here is a simple activity to show students how it works:

1. Place students into groups of 3 – 1 student is the "Picture" and 2 students are the "Taggers."
2. Each student should have a piece of paper in front of them. The "Taggers" need to keep theirs hidden from the others.
3. The "Picture" draws a quick sketch for the others to see (or shows pre-printed images). The "Taggers" then begin making a list of words to describe the picture.
4. After 30 seconds, the "Taggers" reveal their words. Any that they share in common are now part of the database of image tags.

Had they played this game on a computer, the computer would have captured their common tags and their work would have contributed to the work of the bigger "crowd."

5. If the entire class used the same images (pre-printed, rather than sketches), gather all of the tags and then have a discussion about the common words and the unique words found.

Teacher Links:

Watch for the release of Duolingo (http://duolingo.com), a free language-learning site, being developed by von Ahn with funding from the National Science Foundation. Duolingo teaches users new languages while also using the user to translate different language passages.