Augmented intelligence: the Web and human intelligence

Luis von Ahn

School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

1. CAPTCHA

Prof. von Ahn began his discussion of the possibilities of augmented intelligence with a description of his invention (with colleagues) of the Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) [1], often encountered as a random sequence of distorted letters, which human Web users have to repeat to prove that they are indeed human, rather than a computer program capable of filling in a form millions of times per second. This is a task that visually unimpaired humans can do very easily, and much better than computers; more broadly a CAPTCHA is a program that can generate and grade tests that most humans can pass and most computers cannot pass [1].

2. reCAPTCHA

From this, von Ahn moved on to describe reCAPTCHA [2], which began as a project at Carnegie Mellon University, and became a spinout company which was acquired by Google in 2009. reCAPTCHA uses the realization that over 200 million CAPTCHAs were typed daily; each CAPTCHA takes about 10 s, and so about 500 000 h of time is wasted every day typing CAPTCHAs. The question is whether this time could be used more productively, given that during those 10 s one is doing something that computers are unable to do. Is there a large problem that cannot currently be solved by computers, and that can be split into small chunks for humans to solve piece by piece as a by-product of which they authenticate themselves as humans?

Von Ahn suggested the problem of digitizing books, a problem for organizations such as Google, the Internet Archive and Amazon. A book is scanned, and optical character recognition (OCR) technology is used to decipher the page. However, OCR is imperfect, especially on older books (30% mistakes at the word level in books over 100 years old). reCAPTCHA takes the words that OCR cannot recognize (as determined by the OCR
confidence rating) from a scanned book, and presents each word (further distorted) as a CAPTCHA. It is presented alongside a known (and distorted) word; the human’s recognition of the second word authenticates him or her as a human, and then the answer to the original word can be taken as a good indicator of what it actually is.

reCAPTCHA is available as a Web service by companies which need to protect themselves from spammers; the service can be offered free because the human not only authenticates himself or herself, but also contributes to the task of digitizing books. Two hundred thousand sites are now (September 2010) using reCAPTCHA, resulting in 85 million words recognized per day (equivalent to about two million books per year). reCAPTCHA takes about the same amount of time as a CAPTCHA, as it is easier to type recognized words than random characters.

3. Duolingo

Seven hundred and fifty million distinct people (about 10% of the world’s population) have helped digitize at least one word using reCAPTCHAs. This compares with large-scale projects, including building the pyramids, the Panama Canal, or putting a man on the Moon, all of which used something of the order of 100 000 people. The reason this limit was reached was that coordinating (and paying) more than this number of people was impossible prior to the Web. Von Ahn posed the question of what large-scale project could be solved with a three orders of magnitude increase of potential manpower.

The particular project von Ahn focused on was translating the Web into every major language (for free). Machine translation is wrong too often, and its output is not smooth enough to read. Professional translators would be too expensive; translating Wikipedia from English into Spanish would cost $50 million at the cheapest possible rates.

However, getting 100 million people to cooperate in translating the Web has two major hurdles. First, there is a shortage of bilingual people; and second, people lack motivation without money. But von Ahn pointed out that both of these problems could be addressed simultaneously by noting that 1.2 billion people are learning a foreign language. These people are certainly motivated; in the USA alone there are five million people who have paid over $500 for language learning software.

Von Ahn explained how he has transformed the problem of language translation into language learning. He is working on a project called Duolingo, in which people learn a language (for free) while simultaneously translating text (for free). The learning comes about by ensuring that the translations that learners do are not too hard, but still stretching. The project is at the testing phase (at the time of speaking, September 2010), revealing that (i) users learn as well with Duolingo as with highly rated language learning software and (ii) they translate as well as professional translators, if the right sentence is given to the right person. Under such a regime, Wikipedia could be translated into Spanish in five weeks with 100 000 users, and in 80 h with one million users.

Language learning has four components: reading, writing, listening and speaking. Duolingo addresses these as follows.

— **Reading**. The user is asked to translate from the object language into the native language. The output of this exercise, correct or incorrect, will always make sense.

— **Writing**. The user is asked to translate from the native language into the object language. The user understands the input, but may produce nonsensical output. But combining these first two exercises over a number of users cooperatively produces output as good as professional translation.

— **Listening**. Rather than translate, to learn listening the user is asked to subtitle video in the object language. This again produces accurate output, augmented by automated processes such as spellchecking.

— **Speaking**. The user is asked to speak the object language in order to train a speech recognition algorithm (which currently do not work very well with people with strong accents).
All of these are important computing problems, and Duolingo provides an environment in which people are motivated to help solve them.

4. Conclusion

Von Ahn concluded that the human brain is an extremely advanced processing unit that can solve problems that computers cannot yet solve. And because of the Web, we can consider humanity as an extremely advanced and large-scale distributed processing unit that can solve large-scale problems that computers cannot yet solve. Currently, we have a very parasitic relationship with computers; von Ahn advocated a more symbiotic relationship, where humans solve some sub-problems, computers others, and the two are combined to address large-scale problems.

References