Human Computation: Papers from the 2011 AAAI Workshop (WS-11-11)

PulaCloud: Using Human Computation to Enable Development at the Bottom of the Economic Ladder

Andrew W. Schriner[1], Daniel B. Oerther[2], James G. Uber[1]

[1]University of Cincinnati
Environmental Engineering and Science Program
Cincinnati, OH 45221

[2]Missouri University of Science and Technology
Department of Civil, Architectural, and Environmental Engineering
Rolla, MO 65409
{schrinaw@mail., jim.uber@}uc.edu, daniel.oerther@mst.edu

Abstract

This research aims to explore how Human Computation can be used to aid economic development in communities experiencing extreme poverty throughout the world. Work is ongoing with a community in rural Kenya to connect them to employment opportunities through a Human Computation system. A feasibility study has been conducted in the community using the 3D protein folding game Foldit and Amazon's Mechanical Turk. Feasibility has been confirmed and obstacles identified. Current work includes a pilot study doing image analysis for two research projects and developing a GUI that is usable by workers with little computer literacy. Future work includes developing effective incentive systems that operate both at the individual level and the group level and integrating worker accuracy evaluation, worker compensation, and result-credibility evaluation.

Introduction

Many human computation tasks require no specific expertise and simply take advantage of the fact that almost anyone can perform certain basic perceptual tasks far better than the most sophisticated of existing algorithms. Consequently, an opportunity exists to increase the size of the labor markets in human computation systems vastly by including some of the 3 billion people in the world living on less than $2.50 per day.

Some of these 3 billion people are, as Jeffrey Sachs describes in *The End of Poverty*, unable to reach even the bottom rung of the economic ladder. The benefits of participation in the global knowledge economy through crowdsourced human computation work should be clear: such employment would provide the step onto the ladder, where the beneficial cycle of saving and investment in one's future (through education, the provision of healthcare, infrastructure improvements, business ventures, etc) becomes possible.

The logistical obstacles should not be overlooked though: computers, electricity, an internet connection, and computer skills are all required in some form or another to participate in a human computation market. It may be possible, however, to lower the barrier to entry with a carefully designed program of interventions including technological and social resources.

Feasibility Study

In the rural Kenyan village of Kamuga near Lake Victoria, 7 people were recruited to participate in a feasibility study. Participants ranged in age from 19-46 and previous computer experience ranged from none to a single high school class. Using two low-power-consumption netbooks, a solar panel and a car battery, the participants played through the introductory protein puzzles in the protein folding game Foldit, completing an average of 18 puzzles over a total of 8 hours of gameplay. An evaluation of Amazon's Mechanical Turk presented some of the same difficulties as reported by Khanna et al., including a confusing interface and a very limited number of suitable tasks.

PulaCloud

Two image analysis projects are ongoing on a new, simplified web framework for human computation, called PulaCloud (the pula is the currency of Botswana, and means, literally, "rain"). In the first phase of the first project, images are searched from journal articles and then classified as to whether they depict biochemical pathway information; in the second phase, the pathway will be mapped using network mapping software. The second project requires workers to identify microspheres in noisy microscope images in order to calculate concentrations of microspheres. Both tasks are examples of "low-hanging fruit" of human computation tasks, chosen to demonstrate that the approach works in Kamuga with real tasks.

Project Orientation

PulaCloud is oriented toward projects, rather than tasks, to leverage the inherent clustering of the labor market (explained below). Here a "project" refers to the overall problem to be solved, and "task" refers to the subproblem which constitutes a single cycle of human computation. Thus the two-phase project above consists of two deliverables requested by a particular researcher for some payment amount, and the deliverables, or projects, will be completed by one group of people, who will be each be paid a fraction of the total project payment according to their contribution.

The issue arises then of how to calculate contributions in both a fair and beneficially incentivizing way. The fraction of payment should increase with quantity of work completed and also with the accuracy of a user's work. This then requires a way to calculate user accuracy without comparison against a gold standard (and perhaps that is transferable across tasks). Such a measure of compensation might be thought of as a contribution to the information content of the final deliverable. An additional requirement of the accuracy measure is that it should prevent "spammy" work from being compensated at all, and gradually eliminate untrustworthy "workers" from the system. Work is ongoing to develop these methods.

An additional reason for the focus on projects rather than tasks is that is opens the door for the group of people providing the human computation to have a stake in the intellectual property of the project, whether by formal (e.g. patents) or by informal, voluntary means.

Market entry and worker clustering

One significant question to be asked is, "Given the extreme poverty in much of the developing world, how do we enable people to participate in a human computation system in sustainable and cost-effective ways?" Several methods are being explored to allow easier entry into the human computation market, which generally requires a computer or smartphone, a power source, and an internet connection. These methods include 1) microfinance loans for capital purchases, 2) pooling of capital by interested businesspeople in the community, 3) a franchise model, and combinations thereof. For example, a meeting was held with a group of 7 businesswomen from Kamuga regarding their interest in creating a business in which they would invest together to buy computers and then hire community members to do human computation tasks, and their response was very positive (Q: "Would you be interested in creating a business like this?" A: "Of course!"). Their major concern was having someone in the community who was computer-savvy enough to work through problems that might arise. Note also that this is a community where a daily wage of $3.00 is considered a good wage. Another possible model would be to affiliate a human computation "company" with a cyber café.

All of the above models would result in clusters of co-located workers. Such clustering could, with the right incentives, result in higher quality work (due to social accountability and shared responsibility for the deliverable) and over the long term, encourage repeat partnerships between requesters and specific groups of workers.

Tasks

What tasks are appropriate for such a human computation system? Even "simple" language tasks proved to be fairly difficult in the feasibility study, even for those Kenyans who spoke English well, because such tasks often require a high degree of verbal subtlety that is extremely difficult for the non-native speaker to achieve. Fortunately, it is possible (these authors would argue that it is even probable) that the scientific research community as a whole has a tremendous amount of one-time odd jobs with a visual component that would not only provide employment in human computation, but free researchers' time for more productive activities. The great challenge here is increasing awareness of human computation services and making it easier for researchers to post requests.

Future work will include applying human computation to bioinformatics problems in metagenomics.

Acknowledgements

This research is supported by a National Science Foundation Graduate Research Fellowship.

References

J. Sachs, *The End of Poverty: Economic Possibilities for Our Time* (Penguin Press HC, 2005).

S. Khanna, A. Ratan, J. Davis, W. Thies, *Proceedings of the First ACM Symposium on Computing for Development*, ACM DEV '10 (ACM, New York, NY, USA, 2010).

S. Cooper, *et al.*, *Nature* **466**, 756 (2010).