Structures for Creativity: The crowdsourcing of design

Jeffrey V. Nickerson
Stevens Institute of Technology
jnickerson@stevens.edu

Yasuaki Sakamoto
Stevens Institute of Technology
ysakamot@stevens.edu

Lixiu Yu
Stevens Institute of Technology
lisayu1103@gmail.com

ABSTRACT

Crowdsourcing can be used for design. Our work explores the forms of coordination that facilitate creative activity. In particular we have explored the combining of ideas through the use of a sequential combination system, a kind of human based genetic algorithm. We think that even more complex structures of coordination are possible, in which participants specialize their design activity by organizing into fluid hierarchies. In sum, new structures for creativity are possible, and the crowd itself will discover these structures.

Author Keywords

Crowdsourcing, Creativity and conceptual combination, Human based genetic algorithms, Design sketches, Social computing.

ACM Classification Keywords

H5.3. Group and Organization Interfaces.

General Terms

Design.

EXPERIENCES IN CROWDSOURCING

The crowd can be used to design. We have performed several experiments, in which crowds are used to generate designs for open-ended problems [10, 18]: this work has followed our set of papers that explored how implicit coordination happens through web technology [9, 15, 16], and how diagrammatic design representations can be elicited from the crowd [11, 12, 19, 20]. Specifically, we have been studying the crowd's design creativity by implementing an infrastructure to perform human based genetic algorithms.

How can the crowd be organized to take part in large design projects? The ideas of many participants need to be aggregated: one promising method is combination. In a genetic algorithm, part of one genome is combined with part of another to create a new genome. Genetic algorithms have been productive in optimizing many tasks, but they are limited to those situations in which there are clear computable objective functions, and clear unambiguous solution representations. To overcome such limitations, humans might be used as the computing nodes [7]. Human based genetic algorithms, however, need large numbers of people; such crowds were, in the past, difficult to assemble. Crowdsourcing marketplaces such as Amazon's Mechanical Turk have changed this, leading to a variety of large scale experiments in which humans perform tasks almost as if they were computers processing information (e.g. [1, 6, 8 13, 14]). In a similar way, web-enabled human based genetic algorithms become possible: the computer is used to manage the workflow, while the crowd is used to perform the actual combining of two ideas, the crossover step of artificial evolution. These algorithms can be applied to many domains; several will now be discussed.

CHI 2011 Workshop on Crowdsourcing and Human Computation: Systems, Studies, and Platforms, May 8, 2011, Vancouver, BC, Canada.

Solutions for a large scale engineering problem

In our first project, we asked members of the crowd to address the oil-spill problem in the gulf, while the crisis was happening. In all, 1853 participants created, combined, and evaluated ideas over three generations. In Generation 1, a crowd produced text ideas. Next, another crowd evaluated the ideas. We then used tournament selection [5] based on the crowd's evaluation to choose parent pairs of ideas. In Generation 2, another crowd combined the parent pairs to form offspring. This process was repeated to collect Generation 3's ideas.

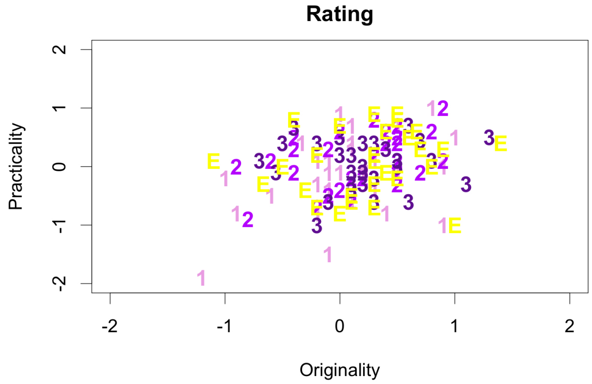

Figure 1. Three generations of ideas, and expert ideas, rated by the crowd with respect to originality and practicality. Purple numbers depict ideas from the crowd: Generation 1, Generation 2, and Generation 3. A yellow E depicts an idea from an expert.

The scatter plot of Figure 1 shows 30 randomly sampled ideas from each of the three generations created by the crowd, overlaid by a set of ideas from experts posted online. The 120 ideas are plotted with respect to their originality and practicality rating: Originality and practicality ratings are the standard measures of creativity [3]. For originality, participants rated whether an idea is more or less novel and surprising than an anchor idea using a 5-point scale, where 2 indicates *more*, 0 indicates the two ideas are about the same, and -2 indicates *less*. With respect to practicality, different participants rated whether an idea is more or less sensible than an anchor idea using an identical scale. The same anchor idea from the first generation was used for all comparisons to let subjects use the same reference point.

As Figure 1 shows, the most original ideas are from the experts and the third generation crowd. The least original ideas are from the first generation, and interestingly, the experts. The ideas from the later generations shift toward originality. Although the least practical ideas are from the first generation, there is no clear shift toward practicality in the last generation. Furthermore, expert solutions are not rated higher in terms of practicality.

Although more work is needed to understand crowd behavior, it is promising that the best idea from the crowd is rated as good as the best idea from the experts. Moreover, the crowd judged that the ideas from the later generations were in general more creative, using three methods: Likert scale rating tasks, two alternative forced choice tasks, and prediction tasks in which participants predict which idea is likely to be judged more creative.

Design sketches for consumer products

In our second project, participants were asked to present their design ideas through sketches. One crowd solved the design problem by producing sketches independently, another crowd evaluated the sketches, and yet another crowd combined the design sketches generated by the previous crowd. Two design problems were tested. The first was an open-ended problem, designing a chair for children [18]. Solutions to this problem are shown in Figure 2. The second problem, the design of alarm clocks, was more tightly specified: constraints of cost and safety were added.

Both of the design problems were run through three generations – the chairs problem with 1047 participants, the clocks problem with 540 participants. The creativity of the designs from Generation 3 was compared to the creativity of the initial designs from Generation 1, using the method explained in [3]. The number of creative designs in Generation 3 was significantly greater than that in Generation 1 for both the chairs and clocks problems.

In the clocks experiment, a contrasting control condition was added, in which the same number of designers as in the combination condition generated new designs instead of combining old designs. Designs created in the combination condition were judged significantly more creative than those in the control condition.

Thus, we have performed three experiments in which large numbers of individuals can work on design problems. Using a combination technique, the outputs become more creative as judged by the crowd. Next, we discuss future research related to the crowdsourcing of design.

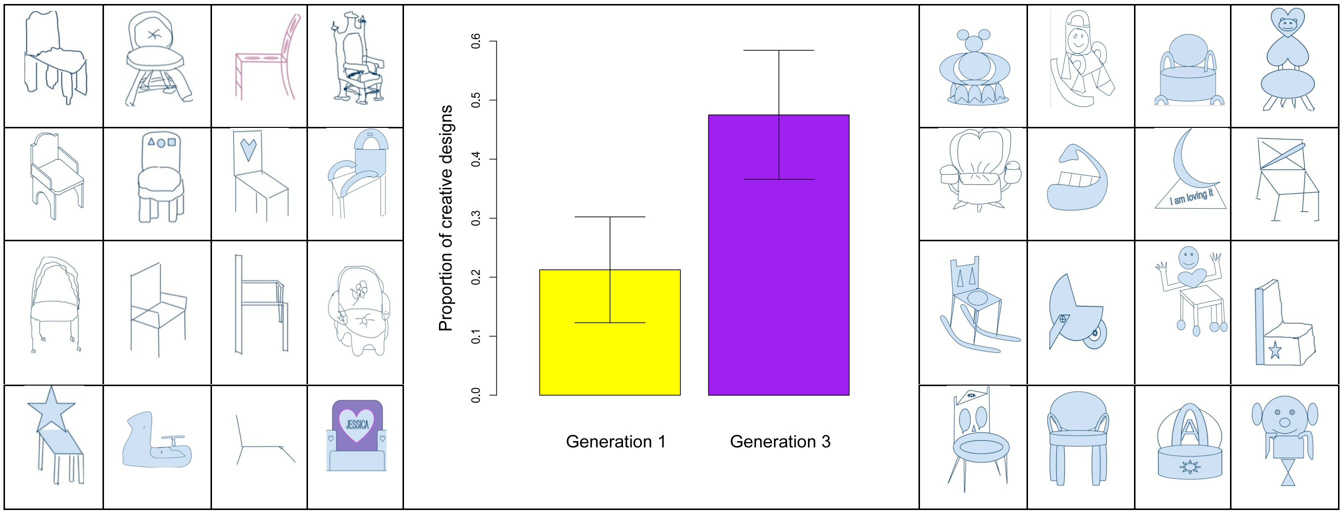

Figure 2. On the left and right are randomly chosen chairs from the first and third generation. The bar chart shows the proportion of creative designs in the first and third generations. The error bars indicate 95% confidence intervals.

FUTURE RESEARCH DIRECTIONS

Our work to date has for the most part used crowds made up of novices to evaluate ideas. In many environments, however, experts may be better sources of evaluation, at least with respect to design criteria related to costs, materials, and production processes. Then, the logical next step is to invite experts into crowd-based experiments. This presents challenges: there are fewer experts than crowd members, and it is not immediately obvious how to use their knowledge most effectively. Some have suggested weighting expert evaluations more heavily [14], but it is probably also important to let experts affect the earlier stages of the design process. They might, for example, be asked to take impractical ideas and find practical, but still original analogs. Experts might contribute criticism, and the crowd might be asked to address this criticism. Also, experts might recommend which specific ideas should be paired for combination.

This presupposes that we as researchers are in the best position to structure the crowd's interaction. But often technologies are appropriated by users in interesting ways. For example, programmers in the Scratch community formed user-initiated contests in order to motivate each other [9]. It seems possible that the crowd could appropriate the sequential combination technology we have built in different and better ways than we can imagine. Consequently, we plan to build a flexible infrastructure to allow for a hierarchy of crowdsourcing: participants will be able to crowdsource to others, and will be invited to define their own coordination structures. We call this *design for appropriation*. The design of the coordination structures might themselves be crowdsourced: just as the crowd has combined problem solutions, the crowd could combine coordination structures using a method similar to the one used in our previous experiments. That is, one member of the crowd may combine a structure that uses voting with a structure that uses hierarchical decision making to form a crowdsourcing process in which the crowd filters a set of solutions through votes, and a few of the filtered set are then selected by the crowdsourcer. Genetic programming [2] treats computer procedures as genomes; by letting the crowd combine workflow processes this activity becomes *human based genetic programming*.

A Future Scenario

Five years from now we imagine that we have been sent this entry from the experiment diary of a professor, who tells us she was a student who attended the 2011 CHI Workshop on Crowdsourcing and Human Computation:

I walked into our laboratory, a room with lots of displays of various sorts that we use to monitor a set of ongoing experiments. We can see our own experiments, and also the experiments of several collaborators around the world, linked through a crowdsourcing collaboratory infrastructure. Ever since we created a mechanism to turn control of crowd organization over to the crowd, we have been frequently surprised.

There are two interesting things emerging right now at the same time, experiments that are operating on different time scales. First, an oil spill happened off the coast of Alaska three days ago, and a latent crowd we had established to volunteer to help in natural emergencies has assembled and organized. They have split themselves up into several groups: (1) Those who are connected to governmental and non-governmental agencies who make sure the volunteer effort is helpful rather than distracting to the professionals involved, (2) those who have contacts in Alaska and can assemble people on the ground who have knowledge of how to cope with the arctic climate, (3) those who will organize the raising of money, (4) those who will help with the dispersing of money and survival goods, (5) those who will coordinate with other crowds, and (6) a group that looks for solvable design problems, like handling oil clean up on a rocky beach, and farm these problems out to other crowds to design. There are still other crowds outside our control, and we can see which organizational structures are effective by noticing how they gain traction, cooperate, and integrate. For example, we found out from a previous experiment that the most effective crowds will quickly redirect excess resources to other problems, to prevent the newly formed organization from being slowed down by over-communication.

In another experiment, a crowd has decided to work on the design of novel forms of water filtration in emergencies, long the domain of experts. A crowd of novices and students are now handling this problem, with some experts providing guidance to the process. They have organized in such a way that ideas are quickly selected and combined with other ideas. Promising original but impractical ideas are refined by experts who see value in them. Expert ideas are morphed and combined by novices, using a process that evaluates at each step the position of the idea in a large design and evaluation space.

As part of this process, we are observing that the crowd as a whole seems to learn, even as individuals join and drop. This may be because the crowd's performance is influencing and molding the organizational structure: the crowd adopts the structures that address the challenge at hand, and these structures solidify learning. We are making use of tools from other researchers that help us measure this collective learning. We are discovering new forms of organization structure, while at the same time creating infrastructure that helps citizens participate in solving large-scale social problems.

BIOGRAPHY OF THE PRESENTING AUTHOR

Jeffrey V. Nickerson is an Associate Professor in the Wesley J. Howe School of Technology Management at Stevens Institute of Technology. He is the Director of the Center for Decision Technologies, an interdisciplinary group of researchers with backgrounds in information systems, computer science, and cognitive psychology. Professor Nickerson has an M.F.A in Graphic Design from Rhode Island School of Design, as well as a Ph.D. in Computer Science from New York University. He is currently the principal investigator on three NSF-funded projects, two involving the use of diagrams in the design of software, and one focused on crowdsourcing design.

ACKNOWLEDGMENTS

This research was supported by two grants from the National Science Foundation, IIS-0855995 and IIS-0968561. We wish to thank Florent D'Halluin and Jin Bao for their participation in running and analyzing the oil spill experiments.

REFERENCES

1. von Ahn, L., and Dabbish, L. Labeling images with a computer game. In *Proc. CHI 2004*, ACM Press (2004), 319-326.
2. Cramer, N. L. "A representation for the Adaptive Generation of Simple Sequential Programs" in *Proceedings of an International Conference on Genetic Algorithms and the Applications*, Grefenstette, John J. (ed.), Carnegie Mellon University 1985.
3. Finke, R., Ward, T., and Smith, S. *Creative Cognition: Theory, Research, and Applications*. MIT Press Cambridge, MA, USA, 1996.
4. Fischer, G. Social Creativity: Turning Barriers into Opportunities for Collaborative Design. In *Proc. PDC2004*, (2004), 152-161.
5. Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Kluwer Academic Publishers, Boston, MA, USA, 1989.
6. Kittur, A., Chi, E. H., and Suh, B. Crowdsourcing user studies with Mechanical Turk. In *Proc. CHI 2008*, ACM Press (2008), 453-456.
7. Kosorukoff, A. Human based genetic algorithm. In *Proc. IEEE Conference on Systems, Man, and Cybernetics*, (2001), 3464-3469.
8. Little, G., Chilton, L. B., Goldman, M., and Miller, R. C. Exploring iterative and parallel human computation processes. In *Proc. of the ACM SIGKDD Workshop on Human Computation* (2010), 68-76.
9. Nickerson, J. V., and Monroy-Hernandez, A. Appropriation and Creativity: User Initiated Contests in Scratch. *Hawaii International Conference on System Sciences*, IEEE Press (2011).
10. Nickerson, J.V., and Sakamoto, Y. Crowdsourcing Creativity: Combining Ideas in Networks, *Workshop on Information in Networks*, NYU, (2010).
11. Nickerson, J. V., Tversky, B., Corter, J.E., Yu, L. and Mason, D., Thinking with Networks, *Proceedings of the 32nd Annual Conference of the Cognitive Science Society,* (2010).
12. Nickerson, J.V., Zahner, D., Corter, J.E., Tversky, B., Yu, L., and Rho, Y.J. Matching Mechanisms to Situations through the Wisdom of the Crowd, *Proceedings of the International Conference on Information Systems*, (2009).
13. Quinn, A. J., and Bederson, B. B. Human Computation: A Survey and Taxonomy of a Growing Field, In *Proc. CHI 2011,* ACM Press (2011).
14. Raykar, V. C., Yu, S., Zhao, L. H., Valadez, G. H., Florin, C., Bogoni, L., and Moy, L. Learning from crowds. *Journal of Machine Learning Research*. 11, 7, (2010) 1297-1322.
15. Sadlon, E., Sakamoto, Y., Ma, J., Barrett, S., Nickerson, J. V. The Ecology of Digg: Niches and Reciprocity in a Social Network Landscape, *ACM CHI Workshop on Social Mediating Technologies: Developing the Research Agenda* (2009).
16. Sakamoto, Y., Ma, J., and Nickerson, J. V. 2377 people like this article: The influence of others' decisions on yours. In N. Taatgen, H. van Rijn, L. Schomaker, and J. Nerbonne (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, (2009).
17. Voiklis, J. *A Thing Is What We Say It Is: Referential Communication and Indirect Category Learning*. PhD thesis, Columbia University, New York, 2008.
18. Yu, L. and Nickerson, J.V., Cooks or Cobblers? Crowd Creativity through Combination, *CHI 2011*, ACM Press (2011).
19. Yu, L., Nickerson, J.V., Corter , J.E., and Tversky, B. The Shifting Shape of Collaboration: The Effect of Hierarchy on the Topology of Communication Plans, *The Ninth Annual SIG IS Cognitive Research Exchange Workshop*, (2010).
20. Yu, L., Nickerson, J.V. and Tversky, B., Discovering Perceptions of Personal Social Networks through Diagrams*, Seventh Annual International Conference on the Theory and Application of Diagrams*, Springer, LNAI 5223, (2010).