Optimizing human computation using a game with a purpose

Main Applicant: Marisa Ponti, 1959-03-22

Purpose and Aims

The creation of more intelligent machines with increasing abilities opens up new ways of collaboration between humans and machines. An arena for the design of complex human-machine systems is "citizen science" or "scientific crowdsourcing" (Lintott & Reed, 2013), an approach to solving complex problems that creates novel opportunities for accelerating scientific progress by involving members of the general public.

The purpose of this project is to optimize a hybrid configuration, involving both human and machine intelligence, which can make the most of the capability of both intelligences, in a unique citizen science game called Quantum Moves. Quantum Moves (Sørensen et al., 2016) was launched in 2012 and by now more than 200,000 members of the general public have participated in the task of building a quantum computer. Currently, there is a great interest in quantum technology throughout the world. Major investments are being made in Sweden, USA, Canada, Japan and China and the EU is launching a Quantum Technology Flagship in 2019. Policy-makers and business managers are starting to realize that quantum technology has the potential to change our society significantly, through improved artificial intelligence, secure encryption and more efficient design of medicines and materials.

This project will unravel the opportunities provided by a more personalized human-computer interaction to solve computationally intractable problems in this game. The ultimate goal is to contribute to better games to solve these problems, meaning that in the end we hope to contribute to the quantum physics field and solve optimization problems and advanced machine learning, for example.

The current optimization in Quantum Moves is based on a deterministic algorithm. For example. currently, there is only one step in Quantum Moves, that is, a player input and then a single step of computer optimization. The simulations in which the Quantum Moves algorithms have tried to optimize the process bear a remarkable similarity to side scrolling casual games, therefore the notion of human-quantum optimization has come up and called into question the characterization of human participants as "computational nodes", resulting in using machine computation or abstract algorithmic processes as a design analogy for human activities. Humans act differently from the logical step-by-step nature of algorithms and, arguably, their intuitive understanding of the counterintuitive quantum problems could help solve challenging problems more creatively in and out of science. We argue that this is the premise of human-computer hybrid optimization: helping algorithms learn from people's blooming, buzzing confusion of solutions, when problems are represented as engaging and counterintuitive game puzzles, and helping machines make sense of and display information in ways that can support human intuition. While the processing power and the sophistication of algorithms have improved at previously unimaginable levels and some computer programs already outplay humans in abstract games (Gibney, 2016a), in quantum physics humans still perform better than computers at solving certain problems. Indeed, Sørensen et al. (2016) showed in a previous study of Quantum Moves that human players are able to find solutions to difficult problems associated with the task of quantum computing. These solutions can be further optimized by a computer algorithm, resulting into a class of solutions that significantly improve state-of-the-art computer optimizations. Whether players succeeded by fluke, by building in-game skills, or perhaps through a simple theoretical understanding of the quantum physics principles represented in the game is still an open question. It is, however, noteworthy that the vast majority of the players managed to find solutions yielding an improvement over the computational results.

This project will address two important scientific challenges for achieving an optimized hybrid configuration in Quantum Moves. The **first scientific challenge** is how to develop a personalized human-computer interaction that reflects and responds to players' sense-making and decision-making more dynamically. The **second scientific challenge** involves how to develop an algorithm that learns the best combinations of strategies by employing mechanisms that appropriately account for the history of players' engagement and solutions.

In order to address these two scientific challenges, the project will pursue the following objectives:

1) Evaluate the interactions between individual players and the existing game version by conducting a cognitive ethnographic study on how players use the features of the game interface to turn information into sense-making and decision-making.
2) Based on the findings of this evaluation, develop a new and more interactive version of the game.
3) Evaluate the interactions between individual players and the new more interactive versions of the two games by conducting cognitive an ethnographic study on how players use the new features of the game interfaces to turn information into sense-making and decision-making.

This project offers a great opportunity to study this human-computer interaction process in the complete development stage. The starting point of this project is the ongoing development and research work on Quantum Moves done at the AU Ideas Center for Community Driven Research (CODER) at Aarhus University in Denmark, which has developed the platform ScienceAtHome, currently featuring several scientific games. By cooperating with them, we will bring our expertise in ethnographic studies and social aspects of human-computer interactions into a group with experience in game development and computational and statistical techniques. We will use our respective competencies to integrate different methods and data generated both on a larger and smaller scale and to achieve the project's objectives, which will ultimately contribute to the development of the game that might be able to solve one of todays' computational challenges.

Survey of the Field

This project situates itself within the ongoing debate over the mind vs machine competition. This debate takes place not only within the scientific community in human computation (HC) (Michelucci, 2013) and other disciplinary areas, but also in popular magazines, and raises questions about the role of technologies and whether they can act "more human" than a person. Several scholars (e.g., Law & van Ahn, 2011) have argued that despite the increasing processing power of today's computers, many problems are still computationally intensive for machines but quite well mastered by humans. Law and van Ahn (2011) argued that part of the reason for this is that we do not yet understand the human intuition that goes into solving these problems, and "as a result, many algorithms resort to a "brute force" approach that exhaustively searches the space of solutions – a strategy which is feasible for small search spaces, but prohibitive when the problem instance is large (p. 10)". Although numerous groups work to develop software that could replace humans in activities like image recognition (e.g., Shamir et al., 2014), there are some scientific problems that are considered computationally intractable. To solve these problems, some successful citizen science projects have developed and used "games with a purpose" (GWAPs) (von Ahn, 2006) to integrate what humans and computers, respectively, can do well. These GWAPs include Phylo (Kawrykow et al., 2012), EteRNA (Lee et al., 2014), Foldit (Cooper, 2014) and Quantum Moves (Maniscalco, 2016; Sørensen et al., 2016). These projects recruit crowds that are usually large and diverse, and crowdsource tasks involving computationally intensive problems, such as protein structure prediction and genomic sequencing, which require creativity, judgment and experience. These complex problems entail a steep learning curve because players need to learn the game mechanics and become skilled before their contributions become significant (Good & Su, 2011). However, little is still known about how players with little or no scientific background learn to solve these computationally intensive problems.

HC in Relation to Human-Computer Interaction

Arguably, a main reason for the success of these GWAPs lies in the knowledge that *interaction* between human intelligence and machine intelligence provides a greater computational power than non-interactive algorithmic systems (Wegner, 1997). The centrality of the interaction between humans and machines in HC, exemplified by early influential work by von Ahn and Dabbish (2004), links this area to a significant body of work done in human-computer interaction (HCI) (Reeves, 2013). Reeves (2013) noted that the HC literature published at HCI venues has tended to focus on how to design interactive HC systems which are correct (i.e., "producing the correct answer in the presence of noise") and efficient (Law & von Ahn, 2011) in terms of 'quality control', or managing issues of 'cheating' or 'gaming the system' through input and output agreement systems (Law & von Ahn, 2009).

This focus on efficiency and correctness tends to characterize humans as "computational nodes" and results in using machine computation or abstract algorithmic processes as a design analogy for human activities (Reeves & Sherwood, 2010). Drawing this analogy tends to obscure the differences between designing for machine algorithmic components and designing for human 'components'. This is a main design issue because human actions are not algorithmic but interpretive and socially organized within an HC system (Reeves, 2013). HC as it has featured so far in HCI "has remained generally unaffected by the significance to the "turn to the social" in HCI" (Reeves, p. 413), which has implied a move from individualistic cognitive conception to a social conception of "user". Embracing a social turn means to examine how humans perform their tasks not in isolation - in terms of input and output transactions - but as a coordinated accomplishment with the features of interface and task design. For this reason, we will adopt distributed cognition (DCog) as a theoretical framework (Hollan, Hutchins & Kirsh, 2000).

Project Description

Theory

Algorithm

Computer algorithms used to optimize the type of problems addressed in games like Quantum Moves typically exist in two variants. In *local algorithms*, an initial solution is slightly perturbed and if the change proves beneficial, further developed in that same direction. This step – although seemingly simple – can be very sophisticated in state-of-the-art algorithms. In particular, in quantum physics the gradient of the score function with respect to small changes can often be estimated analytically, making local search tremendously efficient in many cases (Glaser et al., 2015). Unfortunately, this iterative *deterministic improvement* often causes it to get stuck in what is called a local maximum. In contrast, algorithms with a *random component* such as genetic optimization or machine learning allow for crossing the barriers from one solution type to the next. This randomness ensures a more thorough search of the entire solution space that may eventually locate the best solution – called the global maximum. The disadvantage with this type of optimization is the fact that the steps are random and therefore only beneficial in a small fraction of the times. This makes algorithms with random components exceedingly slow, which means that for certain problems with a sufficiently complex solution space they will in practice never converge to the optimal solution.

In the current version of Quantum Moves there is no active link between players and the optimization algorithms, therefore providing a framework for offering such an interaction will be one of the major milestones of this proposal.

Distributed Cognition to Understand Interactions

Distributed Cognition (DCog) (Hollan, Hutchins & Kirsh, 2000) provides a good descriptive and explanatory power for understanding the interaction between players and technologies, which supports human intuition. Following DCog, intuition is conceptualized as an emergent property of such interactions which can be studied empirically by examining how players' actions are accomplished in cooperation with technologies. DCog posits that cognitive processes should be viewed as including elements outside the physical individual. In DCog, artifacts are scaffolds for cognition, for example, they make cognitive tasks easier or more efficient, and can provide a means for accomplishing cognitive tasks that could not be performed without the tools. This notion posits that cognition is distributed over social groups, the material artifacts in the environment where cognition takes place, and the time dimension, which means that earlier events influence later ones through a process that is itself integral to cognition.

Description of Quantum Moves

The game we will study is Quantum Moves in which players have to learn how to manipulate and control single atoms with high precision and under time pressure. Traditionally, this problem requires simulations of abstract and complex physical systems, but by transforming the physics problem into a game, players can develop an intuitive understanding of the physics and find new and improved solutions to the problem. When playing, individual players interact with the game through their mouse or finger movements and get real-time feedback on their actions by continuously sending their input to the physics model. Once the game is completed, the score is communicated to the player. Then, the entire mouse-path during the game is sent to the researchers at ScienceAtHome and they can subsequently use players' solutions as seeds for a deterministic local algorithm. The successful results from Quantum Moves have recently been published in *Nature* (Gibney, 2016; Maniscalco, 2016; Sørensen et al., 2016).

(Source: http://www.nature.com/news/human-mind-excels-at-quantum-physics-computer-game-1.19725)

Research Design

We will use a design science approach in this project. Design science is an approach to scholarly study that combines traditional research methodologies with the development of an IT artifact to address natural science or social-psychological research questions coupled with design-related problems (Prestopnik & Crowston, 2012). Therefore, the CODER group in Aarhus will gradually incorporate more interactivity in the software code of the two games. Thereupon, we will examine how the performance of players changes in terms of identifying what features are most relevant for sense-making and decision-making.

Our research design includes three main steps: an initial qualitative phase based on small exploratory lab studies using cognitive ethnography and focus groups followed by a quantitative phase based on a large-scale A/B testing, followed by another qualitative stage based on small exploratory lab studies.

In the initial phase, we will conduct exploratory lab studies using cognitive ethnography (Hollan, Hutchins & Kirsh, 2000) to observe how small groups of individual players interact with the existing versions of the two game interfaces to support sense-making and decision-making. In the following section, we describe how we will conduct these ethnographic lab studies.

After completing these studies, we will also run focus groups with players, in which we will tell them about the whole data processing flow used in the existing versions of the two games. For example, they will be told that score matters but is not the only important element to consider in their gameplay because, scientifically, it is more important what their score turns into after it has been given to the local optimization algorithm. In this phase, the main goal is to gather ideas on how players envision the development of a new game in which they can interact directly with the optimization algorithm. Collected data will be transcribed, coded and analyzed to identify patterns, emerging interesting stories, and influence of players' environments or past experiences. The colleagues at CODER will use the findings of the focus groups to develop the first new more interactive version of the two games. More specifically, they will make subtle design changes to the interface and interactivity which will allow better catering better to the needs of the different players' behaviors and levels of progression in the game. These subtle changes will be tested in the second stage of our research, by conducting a large-scale A/B testing (Myers, Ko, LaToza, & Yoon, 2016) over a period of two months. Using the results of focus groups, in combination with other analytics, we will identify the experiment variables likely to provide meaningful variations during the testing. For example, two groups of players in both Quantum Moves and Turbulence will have everything enabled, while other two groups will have a set of conditions toggled. Options can include:

- single conditions for single games (*one game without a certain feature, the other game without a certain feature);*
- a limited amount of variations (*half of the groups in each game with everything enabled, half of the groups with nothing enabled*);
- some other combination (*one game with single variations, bigger groups with certain conditions*).

Two-sample hypothesis tests will be used for comparing data from the tested groups. Since the data from A/B testing does not provide insights in the interaction between players and technologies, in the third phase of our research we will conduct a second round of exploratory lab studies using cognitive ethnography to observe again how small groups of individual players interact with the features of the two new game interfaces to support sense-making and decision-making.

Cognitive Ethnographic Lab Studies

We will use a concurrent Think Aloud Protocol (Ericsson & Simon, 1980) as a primary tool to make players' thinking processes as explicit as possible during their gameplay and find out about the perceived "hints" provided by the interfaces. Adopting DCog, the unit of analysis will not be the individual player, but a cognitive system composed of individual players and the artifacts they use to play the game. We will conduct two rounds of exploratory lab studies, each one over a period of about a week, and will try to recruit two small samples of online players (one for each game) who participated in the A/B testing. We estimate a sample of 27 players of each game is adequate, since according to G*Power calculation, for a one tailed paired t-test 27 participants is adequate to ensure a power of 0.80 at the significance level of 0.05 with a medium effect size (0.50) (Faul, Erdfelder, Buchner, & Lang, 2009). Players are sparsely distributed and it can be a challenge to invite them to participate in the lab studies. However, although out of the 200,000 players of Quantum Moves only 10% are located in Denmark, it can be easier to recruit participants based in this country because of geographical proximity. If we will not manage to recruit a sufficient number of online players, we will use various recruitment strategies, including inviting some participants from physics courses at the graduate and undergraduate levels at Aarhus University.

We will use the Cognition and Behavior (COBE) Lab located at Aarhus University, where Dr. Sherson is associated researcher, to record players while playing Quantum Moves and Turbulence. This lab features 24 computer stations and is adequately equipped for this type of studies. Using a video mixer, both the participants' faces and the computer screen actions will be recorded. This is a common "picture in picture" technique wherein a smaller picture of the participant's face is displayed in one corner of the screen, with the rest of the screen being filled with a recording of the participant's screen activity and most movements. Recordings of the participant's voice while thinking aloud will be gathered on the videotape as well. In addition, we will take notes for later data triangulation. After the participants have completed the games, we will transcribe and code verbalizations and video recordings using a grounded theory approach (Muller & Kogan, 2012) to establish categories and analyze data for patterns of interaction.

Tentative Timeline

	2019 - 2020			
Task	Jul-Sep	Oct-Dec	Jan-Mar	Apr-Jun
X – Conduct exploratory lab studies and focus groups on the existing version of the game – Data collection and analysis. XX – Start to develop the first new more interactive version of the game.	X	X	XX	XX
	2020 - 2021			
	Jul - Sep	Oct-Dec	Jan-Mar	Apr-Jun
XXX – Setup and conduct A/B testing XXX – Analysis of the collected corpus of data from A/B testing XXXX – Conduct exploratory lab studies and focus groups on a new version of the game – Data collection and analysis.	XXX	XXX	XXXX	XXXX
	2021-2022			
	Jul-Sep	Oct-Dec	Jan-Mar	Apr-Jun
XXXX – Conduct exploratory lab studies and focus groups on new versions of the two games – Data collection and analysis. D – Writing papers and dissemination	XXXX	XXXX	D	D

Research Team and Roles

This project is situated in the area of human computation, which is a conceptually distributed field. Therefore, the project is interdisciplinary and builds on the knowledge and methods of more than one field of research, including human computer interaction, computer science and social science. The project will be conducted in close collaboration with the AU Ideas Center for Community Driven Research (CODER) led by Dr. Jacob Sherson at Aarhus University in Denmark, which features an interdisciplinary team of research from physics, computer science, cognitive and social science as well as both backend and frontend software developers. CODER also contributes with cofounding the project. At the Department of Applied IT, University of Gothenburg, Dr. Marisa Ponti will be responsible for the scientific coordination of the project, contributing to setup the A/B testing and taking primary responsibility for conducting cognitive ethnographic research in collaboration with Dr. Wolmet Barendregt, who is an expert in using the Thinking Aloud method, as well as interaction design for games. This project will require extensive game and interaction-design development for the various new iterations of the two games. This process will be funded locally by the CODER center. Dr. Sherson will supervise the overall iterative developmental and computational research activities that will be conducted at CODER (both domain specific quantum physics and mathematics algorithmic development and general data science research on the massive amounts of player data generated within this project), and Mads Kock Pedersen, a data scientist at CODER with appropriate skills in physics, mathematics and computer science, will provide the daily link between the developers and scientists at CODER and the Gothenburg project partners.

Dissemination of Research

We will use multiple vehicles to disseminate our research findings, including peer-reviewed international journal articles (e.g., Nature, Human Computation, International Journal of Human–Computer Interaction) and conferences (e.g., CHI); the ScienceAtHome existing blog to share findings with the people who participated in our research and without whose participation our research would not have been able to take place; social media- Facebook and Twitter - to disseminate with a larger audience; the ScienceAtHome website will also be used to upload and disseminate all the outputs of the project.

Significance for the research area and societal relevance

The use of scientific GWAPs is still in its infancy, thus this project would contribute to the development of knowledge of this human computation genre. Furthermore, we decided to start this collaborative effort because spanning computational and human centered approaches and perspectives is the best way to work at the forefront of human computation.

This project falls into the scope of digitalization and its societal consequences, which the Swedish government identified as a priority research area last year. The most important expected outcome is a contribution to the design of cooperative human-computer systems able to solve outstanding scientific problems. Therefore, our results are expected to widen the knowledge about human computation and contribute to developing this emergent interdisciplinary field. Our results will also have important implications for configuring complex partnerships of humans and machines in other societal areas, for example in work environments where harnessing human computation can offer great opportunities to make more informed and effective decisions to tackle critical problems. This project will also be significant to demonstrate the role citizen science can play in informal learning of physics and math. Involving the general public in the world of hard sciences, starting by their interest in the natural environment, can shed light on the role of participation, social ethics and creative imagination that we nowadays need as a society rich in contradictions. The use of different digital instruments and the role of education to find something in common, despite any cultural and personal diversity, can have an important implication also for those who cannot immediately get the beauty of physics and mathematics. Our results are also expected to be significant for digital science and contribute to the emergence of novel scientific practices to respond to new challenges through global distributed collaborations where members of the general public participate as contributors and direct beneficiaries of scientific knowledge, as advocated by DG Connect to develop European capacity for further research.

International and National Collaboration

Dr Ponti is vice-chair of the COST Action "Citizen Science to promote creativity, scientific literacy, and innovation throughout Europe", a collaborative research network of scientists, which counts 32 countries. Dr. Barendregt is one of the founders of IDAC (Gothenburg Working group on Interaction Design and Children) and the (shared) research leader for the Applied Robotics group at the Interaction Design Division at Chalmers University of Technology.

References

Allen, J. & Chudley, J. (2012). *Smashing UX design: Foundations for designing online user experiences* (Vol. 34). Chichester, UK: John Wiley & Sons.

Cooper, S. (2014). *A framework for scientific discovery through video games*. New York: ACM and Morgan & Claypool.

Ericsson, K., & Simon, H. (May 1980). Verbal reports as data. *Psychological Review*, *87* (3): 215–251.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41* (4), 1149–1160.

Glaser, S. et al. (2015). Training Schrödinger's cat: quantum optimal control. *Eur. Phys. J. D* 69, 279.

Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Trans. Comput.-Hum. Interact.* 7(2), 174-196.

Law, E. & von Ahn, L. (2011). Human computation. In R.J. Brachman, W.W. Cohen, & T. Dietterich (Eds.), *Human computation* (pp. 1-11). San Rafael, CA: Morgan & Claypool.

Law, E. & von Ahn, L. (2009). Input-agreement: a new mechanism for collecting data using human computation games. *Proceedings of the 27th international conference on human factors in computing systems* (CHI) (pp. 1197–1206). New York: ACM.

Lee, J., Kladwang, W., Lee, M., Cantu, D. et al. (2014). RNA design rules from a massive open laboratory. *Proc. Natl. Acad. Sci. USA, 111*, 2122-2127.

Lintott, C. & Reed, J. (2013). Human computation in citizen science. In P. Michelucci, P. (Ed.), *Handbook of Human Computation* (pp. 153-162). New York: Springer.

Kawrykow, A., Roumanis, G., Kam, A. et al. (2012). Phylo: A citizen science approach for improving multiple sequence alignment. *PloS One*, *2*(7):e31362.

Gibney, E. (2016). Human mind excels at quantum-physics computer game. *Nature, 532*, 160–161 (14 April 2016).

Gibney, E. (2016a). Google AI algorithm masters ancient game of Go. *Nature 529*, 445–446 (28 January 2016).

von Ahn, L. (2006). Games with a purpose. *IEEE Computer Magazine*, 96-98, June 2006.

Good, B. M., & Su, A. I. (2011) Games with a scientific purpose. *Genome Biology*, *12*(12):135

Maniscalco, S. (2016). Physics: Quantum problems solved through games. *Nature*, *532*, 184-185.

Michelucci, P. (2013). Human computation: a manifesto. In P. Michelucci (ed.) *Handbook of Human Computation* (pp. 1021-1038). New York: Springer.

Myers, B. A., Ko, A. J., LaToza, T. D., & Yoon, Y. (2016). Programmers are users too: human-centered methods for improving programming tools. *Computer*, *49*(7), 44-52.

Muller, M.J. & Kogan, S. (2012). Grounded theory method in HCI and CSCW. Technical Report. In J. A. Jacko (Ed.), *Human–Computer Interaction Handbook. Fundamentals, Evolving Technologies, and Emerging Applications* (pp. 1003–1024), Third Edition. Boca Raton, FL: CRC Press.

Prestopnik, N. & Crowston, K. (2012). Purposeful gaming & socio-computational systems: a citizen science design case. In: *Group'12*, Sanibel Island, FL, USA, October 27–31, 2012.

Reeves, S. (2013). Human-computer interaction issues in human computation. In P. Michelucci (Ed.), *Handbook of Human Computation* (pp. 411-418). New York, NY: Springer.

Reeves, S. & Sherwood, S. (2010). Five design challenges for human computation. *NordiCHI '10: Proceedings of the 6th Nordic conference on human-computer interaction* (pp. 383–392). New York: ACM

Shamir, L., Yerby, C., Simpson, R., von Benda-Beckmann, A. M., Tyack, P., Samarra, F., et al. (2014). Classification of large acoustic datasets using machine learning and crowdsourcing: Application to whale calls. *Acoustical Society of America*, *135*(2), 953–962.

Sørensen, J.J.W.H et al. (2016). Exploring the quantum speed limit with computer games. *Nature*, *532*, 210-213.

von Ahn, L. & Dabbish, L. (2004). Labeling images with a computer game. *Proc ACM CHI*, pp. 319-326.

Wegner, P. (1997). Why interaction is more powerful than algorithms. *Commun ACM, 40*(5), 80–91.