Scripts in a Gamified Crowdsourcing Approach to Do Science

Marisa Ponti, Igor Stankovic
University of Gothenburg, Sweden
marisa.ponti@ituniv.se, *igor.stankovic@gu.se*

Abstract. This work-in-progress presents early findings of a qualitative analysis of the way in which scripts influence the scope of gameplay and the human comprehensive knowledge of the game in Foldit, a citizen science project. Scripts are pieces of software code that allow players to automatically play the game. The analysis focuses on the circulation of skills between scripts and players. Findings suggest that while simpler actions can be "automated away" using scripts, human spatial-reasoning and creativity skills that are not easily embedded in scripts are still essential to become a top player. Experienced players know how to run scripts at proper stages of the gameplay, while beginners need to develop this competence. The findings suggest that the use of scripts, as game design mechanisms functional to a well-defined scoring system, allows competent Foldit players to strengthen their role of experts rather than becoming appendices of automated gameplay.

Humans and Technologies in Crowdsourcing Science

A growing amount of scientific research is done in an open collaborative mode, in projects referred to as "crowd science", "citizen science", or "networked science" (Franzoni & Sauermann, 2013). One main reason for crowdsourcing science is to address the problem of analyzing massive amounts of data, for example natural language corpora or image databases, while facing scarce human and financial resources. Using an outsourcing platform to delegate tasks, or part of them, to non-experts then represents a possible solution. One may wonder why delegating tasks to humans, rather than using computational methods. Increasingly fast and power computers can replace humans in many tasks. While fewer tasks related to data analysis are performed by people, humans are better than computers in lexical tasks, e.g., recognition of emotions, decoding synonyms, and assigning meanings (Venhuizen, Basile, Evong, & Bos, 2013), as well as in visual tasks, e.g., recognizing similarities, extracting figures from the background, and decoding meanings (Darg *et al.,* 2010).

Crowd science is seen here as a heterogeneous assemblage of entities including humans, technologies, and ideas, each of them with "an identity, interests, a role to play, a course of action to follow, and projects to carry out" (Callon, 1986, 24). In this short paper, the relationship between humans and technologies is addressed by examining the role scripts are brought to bear on gameplay in a crowdsourced scientific project called Foldit. According to Good and Su (2013), Foldit represents a model of crowdsourcing involving people for solving intellectually challenging 'megatasks'. In Foldit, scripts can be seen as "automatic players", as they are pieces of software code that allow players to automatically play the game. The research interest is in examining how scripts influence the scope of gameplay and the human comprehensive knowledge of the game. This paper reports early findings from an ongoing ethnographic study about the circulation of skills between players and scripts.

Setting: Foldit

Created by the University of Washington, Foldit is a collaborative serious game designed as a citizen science project, in which the public is invited to help researchers predict the structure of proteins by using their puzzle-solving intuitions and to play competitively to fold the best proteins. (http://fold.it/portal/info/science) (see Figure 1). In Foldit, not only do players predict possible protein structures, but also design brand new proteins that can help prevent or treat important diseases.

Designed to maximize the combination of powerful computing, and human creativity and spatial reasoning, Foldit presents a visual or spatial puzzle to players, who need to arrange an online protein into the most compact form it can fold into, following the game's rules. Although more than 400,000 people have registered since the start of the game in 2008, at present, only few hundreds are active. Foldit is a complex and intellectually challenging game and players strive to get high scores on each puzzle. Since properly folded proteins are in the lowest possible energy state, players' scores are the opposite of the energy of the proteins they have created.

Players start as beginners and a key task is to score high as solos and/or as evolvers who increase their teams' ranking score. Levels and score are obtained via hand folding, using scripts, and/or a combination of the two at different stages of the gameplay.

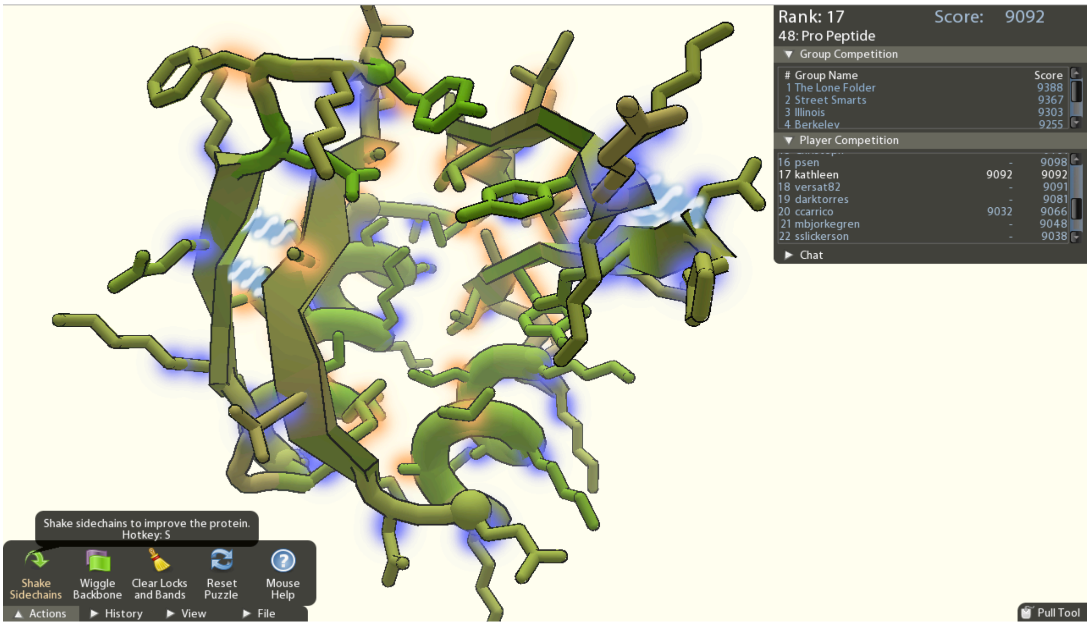

Figure 1. A screen shot of Foldit (Source: Fold.it).

Role of Scripts

Players can use two ways to fold a protein and solve a puzzle: they can hand fold and/or use scripts. In Wikia Foldit, a site maintained by experienced players, a script is described as a kind of cookbook recipe, as it contains a series of commands written in the programming language Lua telling Foldit what to do with a protein. Players themselves develop scripts, thus contributing to Foldit in a twofold way: by solving complex scientific problems and developing scripts. Given that the game has become increasingly complex over the years, the Foldit development team let the players use scripts to help them manage such complexity. The development team built a Cookbook feature into the game client (Figure 2) to help players systematize their strategies into scripts and integrating them into the game (Cooper, 2013). More than 100,000 scripts are shared by players, out of which 1000 scripts are publicly shared on the Foldit portal

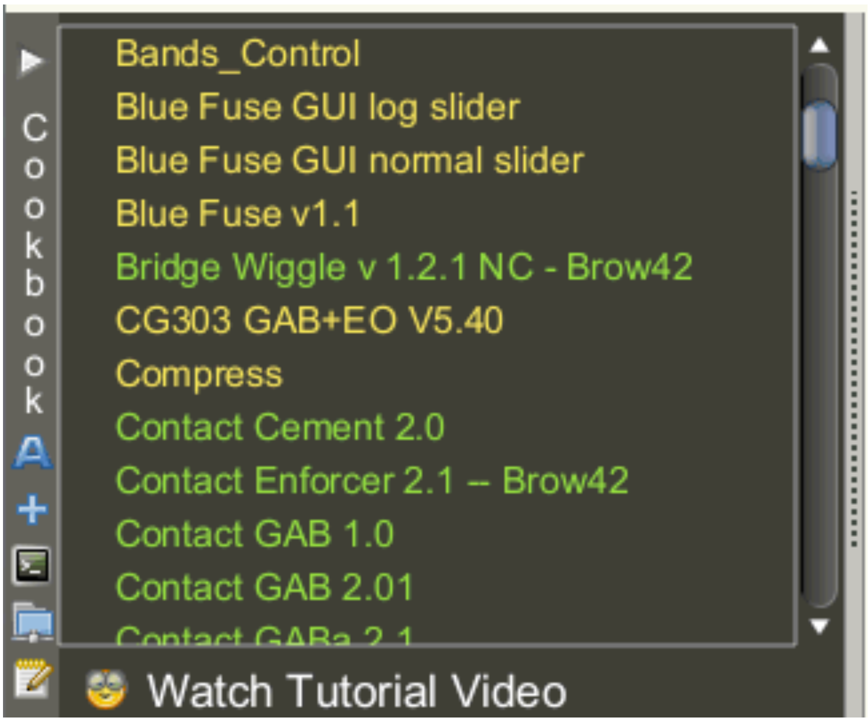

Figure 2. A screen shot of the cookbook containing different types of scripts (Source: Bruno Kestemont).

Gathering Empirical Data: A Case Study

Method

The study involves active participation in the game, use of content found in online documents (e.g., forum discussion and Wikia Foldit), semistructured interviews with players, and an online survey linked to the forum. 11 interviews were conducted with players (two of which no longer active) geographically located across Europe and the Unites States, who described their experiences with using scripts. English was used as lingua franca. Interviewees included eight men and three women and were conducted from February to May 2015. All interviewees achieved mid-to-high scores on games statistics, and several were present on player ranking leagues. Nine players undertook email interviews, which varied in the amount of content and in the depth of responses, and two players chose to talk via Skype. Some interviews involved several rounds of exchange via email and made it possible to gain a better insight into individuals' perspectives. Interviews were analyzed using the early stages of thematic analysis, such as open coding and collation of data into potential themes (Braun & Clarke, 2006). In parallel with working with codes and themes, Actor-Network Theory Analysis Diagrams (AADs) was used. Developed by Payne (2014), this analytical device helps focus attention on the actors involved in a process and the kind of relationships between them, revealing connections that might go unnoticed otherwise.

Finding: Relationship between Scripts, Score and Human Intervention

10 out of 11 players use scripts, as they think they are necessary to increase scores and become top solos. As one player put it, " *I would say that there are only a select few players that will only use hand folding, and they rarely play anymore. It would be like someone who always used a slide rule and refused to switch over to a calculator. This person eventually fades away as a participant*". Players reported their use of scripts to perform actions that are tedious, repetitive, and time consuming. Scripts are adaptations of manual moves. Scripts written by good handfolders are more likely to "contain" scoring actions and make player level up. Although almost all interviewees indicated that scripts are necessary to score high and become top players, they are also cautious about using only scripts. As reported by one player, participants using only scripts are a select few, whose *"strategy sometimes pays off and they manage to offer up a protein that scores well, but is not particularly useful to science"*. This quote clearly suggests that levelling up does not imply to replace human-play with automatic-play. Just running the script is not sufficient to level up consistently and become a top player. Players referred to scripts as being simply "macros" that are not "smart". While they can help detect good shapes and allow players to make rapid progress to the solution of puzzles, scripts were said to do very little to fix fundamental flaws, "*and that requires intelligent intervention – something that is notoriously difficult to do with software*". Thus levelling up is an activity that requires players to develop several skills to perform a good high-scoring fold. As another player noted, "*sometimes computers are able to use the contact data to correctly predict the native structure, and sometimes not*". Therefore, if certain automatic moves do not provide high score, players need to use their spatial reasoning and problem solving skills to make decisions about different actions. While experienced players know how to run scripts at proper stages of the gameplay – for example, they may spend time to handfold at the beginning of the puzzle and then run local optimize scripts as the game progresses (Cooper, 2013) – and have a good idea of what a natural protein would look like, beginners need to develop this competence. As one top-ranked player noted "*A typical beginner, when they come in our team, I see what they do. They try a recipe and then they share something which is awful*". Thus the results suggest that experience of players is an intervening condition bearing upon their competence to use scripts effectively.

The circulation of skills between players and scripts is described in Figure 3, using the AADs notation (a full description of how to make AADs is in Payne, 2014).

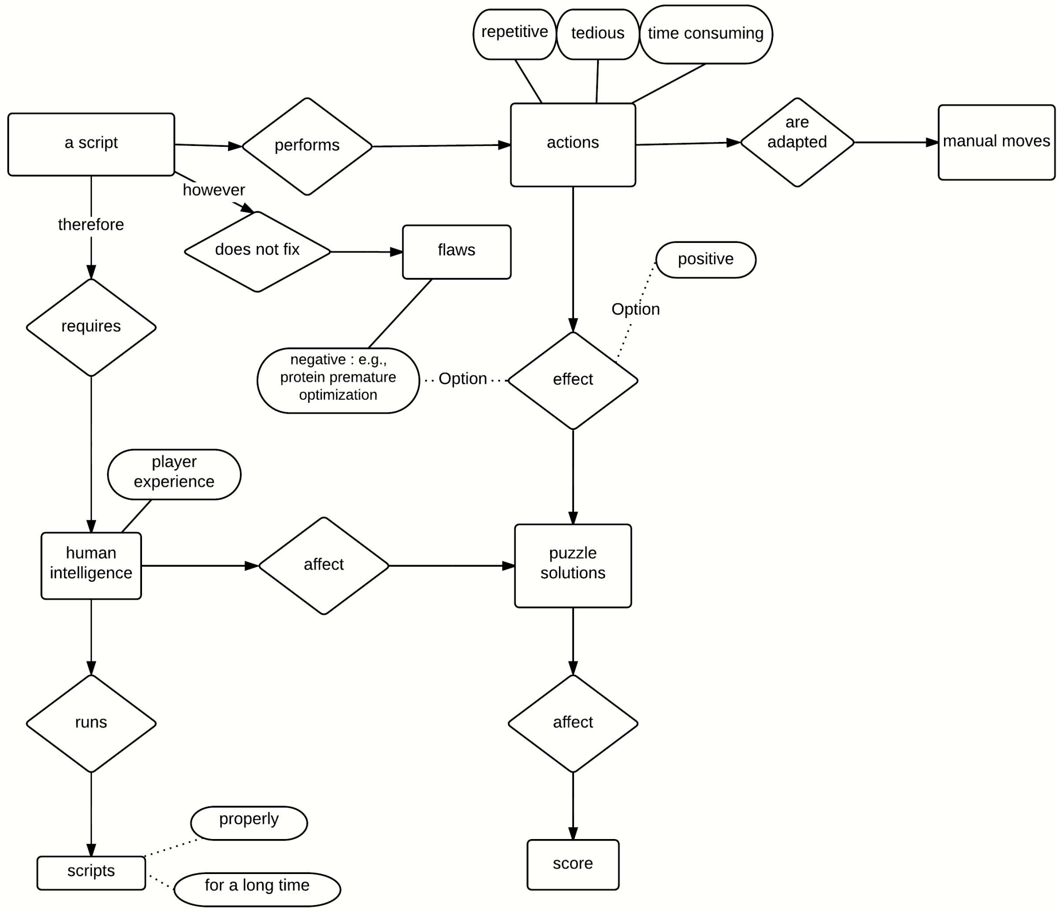

Figure 3. The circulation of skills between players and scripts

Discussion

In the context of digital games, Che Pee (2011, p. 22) defines competence as a ‘‘combination of knowledge, skills, and behavior leading to an individual being able to perform a certain task to a given level.’’ Carmeli and Tishler (2006) define skills as the ‘‘ability to do something in an effective manner’’ (p. 13). Foldit has been designed to offer a place where both computational power and human spatial reasoning and creativity for problem solving can be best applied (Cooper, 2013).

When playing Foldit, skill is thus the ability to solve puzzles in an effective manner. The findings suggest that the use of scripts, as game design mechanisms functional to a well-defined scoring system, allows competent Foldit players to strengthen their role of experts rather than becoming appendices of automated gameplay. The scoring system allows Foldit to identify the relatively few players that can make great contributions to the scientific goals (Good & Su, 2013). As rightly pointed out by Good and Su, the game is used as a mechanism to identify players with exceptional skills within a large number of participants. Scripts can be said to enforce this mechanism, because players with exceptional skills can use them strategically to gain in productivity, while achieving scientific goals. Therefore, scripts seem to strengthen the position of skilled handfolders and contribute to the creation of a community of game-developed experts, as the Foldit team aimed for. On the opposite, inexperienced players can use scripts to play blindly. However, just running scripts is not enough for levelling up, because scripts can have serious limitations in relation to designing and predicting the structure of a protein. This is why it is necessary to develop competence about what scripts can and cannot do. The challenging nature of protein folding and design requires competence development, because players need to learn when to use scripts and when to handfold. Arguably, the "enskilling" argument could be considered here. The advocates of enskilling posit that increased technology leads to more human skill and not less (Bratton et al., 2010). Following this thesis, it can be suggested that the use of scripts configures a redistribution of skills in the gameplay, in ways that enable a relatively small number of amateurs to take on the challenging tasks involved in the game. In fact, it is difficult to embed in a script the human capabilities that make a player score high and become top ranked. While simpler actions can be "automated away", human spatial-reasoning and creativity skills are still essential to design a good protein. Redistributing competences to scripts, through adaptations of manual moves that have proved to score well, does not make scripts able to outperform good players in the design of a protein, because scripts do not display yet the kind of creative skills that belong to skilled hand-folders. It can also be argued that the use of scripts does not attempt to engage a wide range of players at different levels of game competence, but it is more interested in maximizing the activity of "talented" players, helping them become "highly ranked" expert players.

Future Work and Theoretical Framework

The *circulation of skills* becomes a central theme around which to organize further the analysis of the relationship between scripts and players. Drawing from Latour (2005), the idea of symmetrical circulation of skills between humans and non-humans will be used to interpret the circulation of skills, and the concept of hybridity will be used to understand scripts as tools useful in appropriate combinations with other tools (e.g., players' strategies), and with the structure of the game itself (e.g., initial state of a protein).

Acknowledgments

We are very grateful to the Foldit players who help make this study possible. We want to especially thank Bruno Kestemont for his comments on this manuscript. This research is supported through funding from Marcus and Marianne Wallenberg grant no. 2013.0020.

References

Carmeli, A., & Tishler, A. (2006): 'An empirical analysis of the relative importance of management team', *International Journal of Manpower,* vol. 27, pp. 9–36.

Bratton, J., Sawchuk, P., Forshaw, C., Callinan, M., & Corbett, M. (2010): *Work and Organizational Behaviour: Understanding the Workplace*. Palgrave Macmillan, New York.

Callon M. (1986): 'Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of St Brieuc Bay', in J. Law (ed.): *Power, Action and Belief: A New Sociology of Knowledge?* London, Routledge, 1986, pp. 196-223.

Cooper, S. (2014): *A framework for scientific discovery through video games*. Association for Computing Machinery and Morgan & Claypool, New York, NY, USA.

Darg, D.W., Kaviraj, S., Lintott, C. J., Schawinski, K., Sarzi, M., Bamford, S., & Vondenberg, J. (2010): 'Galaxy Zoo: the fraction of merging galaxies in the SDSS and their morphologies'. *Monthly Notices of the Royal Astronomical Society,* vol. 401, pp. 1043–1056.

Che Pee, N. (2011): *Computer games use in an educational system*. Unpublished PhD thesis, University of Nottingham, Nottingham, England.

De Paoli, S. (2013): 'Automatic-play and player deskilling in MMORPGs', *Game Studies: The International Journal of Computer Game Research*, vol. 13, no.1. Retrieved March 23, 2015 from http://gamestudies.org/1301/articles/depaoli_automatic_play.

Good, B. M., & Su, A. I. (2013): 'Crowdsourcing for bioinformatics'. *Bioinformatics* vol. 29, no. 16, pp. 1925-1933.

Latour, B. (2005. *Reassembling the social: An Introduction to Actor-Network Theory*. Oxford University Press, Oxford.

Franzoni C., & Sauerman, H. (2014): 'Crowd science: The organization of scientific research in open collaborative projects', *Research Policy*, vol. 43, pp. 1–20.

Payne, L. (2013): *Why do students choose computing?: Influences, perceptions and engagement*. Unpublished PhD thesis. Coventry University, UK.

Venhuizen, N.J., Basile, V., Evong, K., & Bos, J. (2013): 'Gamification for word sense labeling', *Proceeding of the 10th International Conference on Computational Semantics* (IWCS-2013), pp. 397–403.