Harnessing the Cognitive Surplus

James Scott-Brown

How do you spend your free time? If you were an average American, you would spend 20 hours a week watching television, and another 3 hours playing games [1]. Clay Shirky has written about how, after the Second World War, enormous changes in society occurred, so that "society forced onto an enormous number of its citizens the requirement to manage something they had never had to manage before--free time", creating a vast "cognitive surplus" [2]. How can this surplus be effectively exploited? One approach is to take advantage of people's leisure time and enjoyment of computer games to solve real problems, by creating games in which players complete tasks that computers cannot yet. This has been done for the problems of image tagging (the ESP Game/Google Image Labeler), locating objects in images (Peekaboom), collecting commonsense facts (Verbosity), predicting protein folding (Foldit) and improving the design of electronic circuits (funSAT). Alternatively, mundane but essential tasks can be modified to have useful side-effects (reCAPTCHA).

> **"By September 2008, reCAPTCHA... transcribed over 440 million words."**

The work of Luis von Ahn is a good example of harnessing the human "cognitive surplus" using both approaches. He began by developing the CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart), which distinguishes humans from computers by asking them to complete a task, usually typing letters from a distorted image. This allows websites to prevent computer programs from automatically creating multiple accounts, sending large numbers of spam messages, or making many attempts to guess a user's password. The initial paper describing CAPTCHA suggested that, as well as helping to distinguish between computers and humans, the tests could encourage work on improved text recognition algorithms, jokingly saying that this was 'how lazy cryptographers do AI'; however, the time humans spent completing them was wasted [3]. A subsequent refinement, reCAPTCHA, uses the process of completing a CAPTCHA to decipher words in scanned documents that are too distorted or smudged to be recognised by a computer [4]. This is done by presenting the user with images of two words simultaneously: an "unknown word", and a known "control" word selected randomly from a list of more than 100,000. If a user types the control word correctly, they are assumed to be human, and their attempt at typing the unknown word is recorded. Once three people have provided the same response to the "unknown word" image, they are assumed to be correct, and their transcription is added to the list of control words. Thus, the number of transcribed words increases as reCAPTCHAs

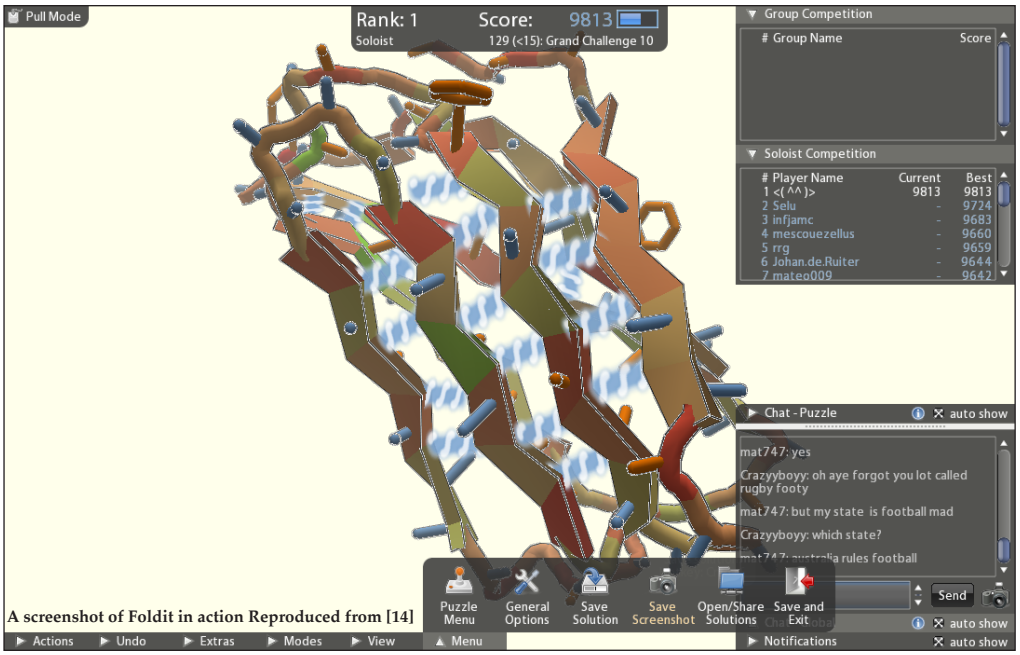
Reproduced from [13]

are completed. reCAPTCHA has been an enormous success: by September 2008, it was being used by more than 40,000 websites, and had transcribed over 440 million words [5].

> **"Both reCAPTCHA and the GWAP games are very easy to learn"**

Von Ahm has also developed a series of what he calls "Games With a Purpose" (GWAP) [6,7]. Several of them award points for agreeing with other players: in the ESP Game and popVideo, a group of players see the same image or video, and must submit keywords that could describe it; in Matchin, players are presented with a pair of images, and asked which one they prefer; in Squigl, players are shown an image and must draw an outline around a named object. Others rely on describing an object to a partner, and understanding their description: in TagTune, players describe what they hear to their partner and, based upon each other's descriptions, guess whether they are listening to the same tune; in Verbosity one player has to describe a word by what it "is" – "is a type of", "looks like", "is about the same size as", "is the opposite of" and "is a kind of" – while the other has to guess what it is [8]. By collecting human responses to questions, these games teach the computer specific facts, either about the meaning of words (Verbosity) or about particular images, videos, or pieces of music. All of them are useful in some way: tagging media with keywords allows search engines to produce more relevant results; the facts collected by Verbosity may be useful for Artificial Intelligence and Natural Language Processing projects; and by rating how attractive images are, Matchin could lead to computer systems that select the prettiest images to present to users.

The main reason that these games are fun is that players must try to guess what other players are thinking. In the ESP game, for example, users are told which of their suggested tags matched those of their partner, which can influence their future suggestions. By adding this social interaction, boring tasks like tagging images or music are made fun. Additionally, users are actively encouraged to refer their friends to the site by bonus points. The competitive element extends beyond recruiting friends, as at the end



A screenshot of Foldit in action Reproduced from [14]

of each game, players are told how many more points they need to match the day's high-score, encouraging them to start another game. Players who receive specific numbers of points are awarded skill levels, and those with the highest scores appear on a leaderboard. During a game, continual encouragement to play on is provided by a sound playing

> **"Humans are better than computers at selecting more promising starting points"**

(and, in some games, motivating text appearing) whenever a match occurs. Fixed time limits for each game maintain interest, forcing players to think quickly, so that the games are more challenging - and hence fun. One player apparently claimed that the ESP game was "like crack": arguably, with its flashing lights and beeping noises, the game has more in common with a casino [9].

Both reCAPTCHA and the GWAP games are very easy to learn - it takes just seconds to read and understand the rules - which has surely contributed to their success. However, such simplicity is not essential, and many players have been attracted to the much more complicated game Foldit, in which players attempt to manipulate predicted proteins structure to produce more likely (i.e. lower energy) structures. Since it is not immediately obvious how to interpret or alter the protein structures, help is provided by a series of in-game tutorials. A separate wiki explains the relevant biochemistry, in-game controls and strategies [10]. Most of the top players in Foldit have had no more than high-school teaching in biology, yet in a major protein-folding competition they predicted structures better than all fully automated programs, and about as well as professional biochemists [11,12]. In particular, humans are better than computers at selecting more promising starting points, and performing major restructuring to escape from energy local minima (e.g. rearranging the core of proteins to create space to bury a previously outwards-pointing hydrophobic residue, or changing hydrogen bonding in beta-sheets). Players can thus produce better predictions of protein structures, which are important not only to the understanding of basic biological processes, but also to the rational design of drugs targeting specific proteins in disease.

So why do people choose to play these games? Perhaps they are enticed by invitations to "contribute to science" (Foldit) or claims that "You're helping the world become a better place . . . you're training computers to solve problems for humans all over the world" (GWAP) - which the games do fulfill. When Foldit players were asked their motivation for playing, about 40% of answers referred to the scientific purpose of Foldit; 35% to a feeling of immersion in a task; 20% to a feeling of achievement; and the remainder to social involvement in the game [11]. Players of the more casual GWAP games are likely to be motivated less by the higher purpose of the games, and more by a sense of fun and competitiveness. Together, GWAP and Foldit have shown that people can be persuaded to perform useful tasks that cannot be automated, by presenting them in the context of a game, with rules, goals, and scores. In this way, otherwise idle minds can achieve what computers cannot. ■

James Scott-Brown is a first year student studying Natural Sciences at Trinity College.

References

1.US Bureau of Labor Statistics American Time Use Survey 2009, Tables A-2 and 11
2. Clay Shirky. Gin, Television, and Social Surplus. Talk given by Clay Shirky at the Web 2.0 conference, 2008 Apr 23. Transcript available from: http://www.herecomeseverybody.org/2008/04/looking-for-the-mouse.html
2. reCAPTCHA: Stop Spam, Read Books [Online Homepage]. http://www.google.com/recaptcha
3. Von Ahn et al. CAPTCHA: Using Hard AI Problems For Security. Proceedings of Eurocrypt 2003; 2656:294-311
4. reCAPTCHA: Stop Spam, Read Books. http://www.google.com/recaptcha
5. Von Ahn et al. reCAPTCHA: human-based character recognition via Web security measures. Science 2008; 321(5895):1465-8.
6. Gwap.com [Online Homepage]. Available from: http://gwap.com
7. Von Ahn and Dabbish. Designing Games With a Purpose. Communications of the ACM 2008; 51(8):58-67
8. This is similar to Burgener's 20Q game, in which users train a neural network, so that it is able to distinguish between objects by asking fewer than 20 questions. 20Q.net Inc. [Online Homepage] http://20q.net
9. Thompson, Clive. For Certain Tasks, the Cortex Still Beats the CPU. Wired Magazine. Issue 15.07. Accessible from: http://www.wired.com/techbiz/it/magazine/15-07/ff_humancomp
10. Foldit wiki. [Online Homepage] http://foldit.wikia.com/wiki/FoldIt_Wiki
11. Cooper et al. "Predicting protein structures with a multiplayer online game." Nature 466(7307):756. Supplementary Figure 3 shows the biochemical experience of players; Supplementary Figure 4 shows motivations for playing the game
12. CASP8 Results. Available from: http://fold.it/portal/node/729520
13. ©Google. Generated from: http://www.google.com/recaptcha/learnmore
14. CC-BY-SA, Ducey, Wikicommons. Available from: http://commons.wikimedia.org/wiki/File:Foldit_1236616893.png