A Taxonomy of Data Types for Games with a Purpose

David Gundry[*]
Department of Computer Science
University of York
Heslington, UK
deg500@york.ac.uk

Abstract

Games with a Purpose collect specific kinds of data from their players. They generally reward the player for providing good data. I propose a taxonomy of four types of data that such games can collect. I discuss the challenges for each in providing suitable rewards for players.

Introduction

Game design is used to motivate engagement with Human Computation tasks such as *FoldIt* (Cooper et al. 2010), and the *ESP Game* (Ahn 2005), but also scientific experiments (eg. Oladimeji et al. 2012). Unlike games intended to inform or change behaviour, data collection tasks are primarily concerned with gathering high quality data. Poor design can undermine the validity of this data.

Motivation is often encouraged by using extrinsic rewards. Therefore, it seems obvious that player behaviours that yield good data should be rewarded. However, two difficulties arise with this policy:

1. How ‘good’ the data is might be dependent on the circumstances in which it is collected
2. Some data cannot be objectively (or automatically) evaluated to determine a reward.

Ahn (2005) proposed agreement-based game mechanics to infer the value of subjective data. Similarly, there have been studies (eg. Oladimeji et al. 2012) that make use of data that is dependent on the manner of collection, but only when this can be evaluated automatically by the game. The intersection of the two — dependent data that cannot be objectively evaluated — has been left, possibly because of the inherent problems in motivating it. Yet this sort of data is relevant for many studies, particularly in the social sciences and humanities.

In order to characterise this sort of data for future research, I propose a taxonomy of four data types that games with a purpose can collect, defined by two dimensions.

[*]This work was supported by the EPSRC Centre for Doctoral Training in Intelligent Games & Games Intelligence (IGGI) [EP/L015846/1].

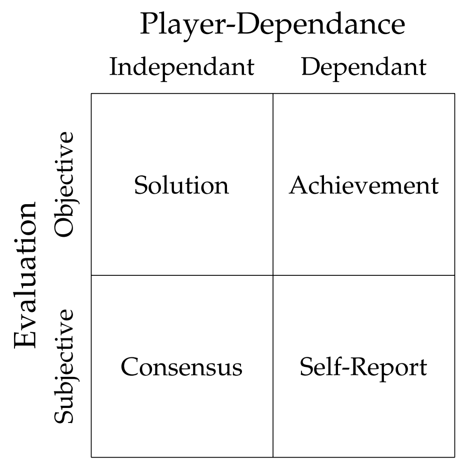

Figure 1: Data objectivity and dependance on the player give four types of data.

Dimensions

1. Player-Dependance In the game *FoldIt*, players engage in a 3D virtual protein-folding task, trying to achieve the most points by finding the most optimal solution. Here the value of a solution is independent of how it was discovered. I will call this *player-independent*. In contrast, a scientific experiment requires strict experimental controls to ensure that data has been collected in a specific manner. Where the data does not conform to these controls, it cannot be used. I will call this *player-dependent*.

2. Evaluation Certain problems, like protein-folding, are very difficult for computers to solve, but easy for computers to check. This means that once a player has provided a solution in *FoldIt*, it can be immediately scored by an algorithm. I will call this an *objective* evaluation. Sometimes there is no way to design an algorithm to check a solution, but it can be evaluated by a human. I will call this a *subjective* evaluation.

Data types

The four types of data that emerge from these two dimensions are shown diagrammatically in Figure 1. These are termed *solution*, *consensus*, *achievement*, and *self-report*.

Solution Games such as *FoldIt* gather solutions. The problem is independent from the individual providing the data, and the quality of a solution can be evaluated objectively, thus using the player base as an efficient search algorithm.

The design of the game is likely to affect the quantity of the data and the quality of solutions. As it does not matter how the player arrives at a solution, the reward mechanisms do not affect the validity. A close match between the rewarded behaviour and the desired data will help the players practice generating solutions and thus improve the data they provide.

Consensus However, for some problems the solution cannot be objectively evaluated. Yet appeal can be made to the 'wisdom of crowds'. Here the data gathered is not a solution to a problem, but a consensus from the players as a whole as to the quality of a given solution. For example, players of the *ESP Game* provide descriptive labels for images; however, a candidate image label has little intrinsic value, until it is known whether or not it is accurate. To design for consensus, agreement-based game mechanics are often used (Quinn and Bederson 2011).

Feedback is provided to the player on the basis of the consensus of other players, rather than an objective evaluation. It is also necessary that the ability for players to collude must be constrained. The circumstances of an individual player, however, need not be controlled. Similarly, the game may be seeded with initial solutions, or the players used to also evaluate machine-generated solutions.

Achievement Where the data is dependent upon the context and behaviour of the player, but can be objectively evaluated, there is arguably always some degree of achievement involved. This achievement is in how *well* the player performed the task, as provided by the objective evaluation. For example, in a reaction time study, there is no confound in rewarding speed of reaction; nor is there a confound in rewarding accuracy in a number-entry study. If there is no sense in which attempts differ in how well they performed, then the evaluation is subjective, at least for the purposes of this taxonomy.

Oladimeji et al. (2012) use a game about monitoring patients in a hospital to study number-entry error. The data of interest is not a solution to the number-entry problem, rather it is the error-making behavior of the players. The players were trying to enter the numbers as quickly and accurately as possible, therefore, their behaviour is also their achievement at that task. The outcome can be objectively valued, and there is no confound in using this to drive rewards.

Such experiments must control inconsistent presentation between players. This precludes incorporating feedback that is based on direct comparisons with other players, such as leaderboards.

It is not necessarily desired that players practice and refine their skills to produce higher quality solutions, so the quantity of data is usually fixed to avoid practice effects.

Self-report There are research contexts in which the data relates uniquely to an individual, or to game-external factors in the player's environment. These are cases where the player data provides measurements, not of behaviour during the game, but of qualities that the game cannot control. For example, if the game collects players' heights, this cannot be validated. As the data in a self-report game cannot be evaluated, either automatically or through agreement, rewards for data quality are not possible.

Yet, because the data often relates directly to the player, another form of reward is possible: the data itself. Collected data may be visualised in a way that is most relevant or interesting when it is correct. As a simple example, a self-report measure of height and weight could supply the associated Body Mass Index (BMI) value. While a player could provide inaccurate values, there would be little reason to.

Similarly, communication with other players can provide a motivation for providing good data, particularly when the data is linguistic. Here the motivation is to play collaboratively, and communicating meaningful data is an intrinsic part of that.

This motivation for the data itself suggests that it is crucial for self-report games to establish *relatedness*, as described in Self-Determination Theory (Deci and Ryan 2002). By establishing characters and stories in the game that the players want to engage with genuinely, the motivation of providing good data becomes intrinsic to the activity. The players can enjoy the game more because, in some way, it is about them.

Discussion

The taxonomy presented here describes four kinds of data that games could collect. Solution and consensus are well explored in Human Computation, as is achievement in HCI.

However, there are few examples of games designed for self-report data; perhaps because this is seen as a less suitable target for games due to its restrictions. Certainly motivation with extrinsic rewards is not applicable. I have suggested that relatedness is the key to motivating this data.

References

Ahn, L. V. 2005. *Human computation*. Ph.D. Dissertation, Carnegie Mellon University.

Cooper, S.; Khatib, F.; Treuille, A.; Barbero, J.; Lee, J.; Beenen, M.; Leaver-Fay, A.; Baker, D.; Popović, Z.; and Players, F. 2010. Predicting protein structures with a multiplayer online game. *Nature* 466(7307):756–60.

Deci, E., and Ryan, R. 2002. Overview of self-determination theory: An organismic dialectical perspective. In *Handbook of self-determination research*. Rochester, NY: University Rochester Press. chapter 1, 3–33.

Oladimeji, P.; Thimbleby, H.; Curzon, P.; Iacovides, I.; and Cox, A. 2012. Exploring unlikely errors using video games : An example in number entry research. In *Workshop on Safety-Critical Systems and Video Games: Contradictions and Commonalities, Held at Fun and Games*, 3–7.

Quinn, A., and Bederson, B. 2011. Human computation: a survey and taxonomy of a growing field. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1403–1412. New York: ACM.