COMBINATORIAL GAMES AS TESTBEDS FOR LEARNING PATTERNS RECOGNITION: THE NIM APP

Eugénio Rocha, Ana Breda
Department of Mathematics/CIDMA, University of Aveiro (PORTUGAL)

Abstract

In this work, we show the first results of a project where a combinatorial mobile application is used as a tool to gather users' data, allowing some understanding about the learning behaviour of users solving combinatorial tasks, in particular, related with the NIM game. To obtain the results, 6,514 games and 29,667 moves were analysed and players where considered as production units, that transform input resources, i.e. attempts, moves of game pieces, number of games played; into output results, e.g. number of victories. This approach may give some insights to the use of digital games as research tools in Education.

Keywords: Combinatorial games, optimal strategies, learning patterns, math skills, mobile apps.

1 INTRODUCTION

The idea of individualized learning styles has greatly influenced Education despite criticisms, see Coffield et al. [1], e.g. as there is no evidence that identifying an individual student's learning style produces better outcomes. However, well-characterized and strong focused activities, where a process of self-education exists without external tutoring, seem to be a good environment to recognize individual learning patterns. Combinatorial games (CG) can produce such environments where users are highly motivated and the topic, by itself, is mathematically rich. Many CGs have associated a so-called optimal winning strategy (OWS). In summary, a solved game is a game whose outcome (win or lose) can be correctly predicted from any position, assuming that both players play perfectly.

Traditionally, the NIM game is a CG board game for two players with quite simple rules: starting with any number of counters distributed in any number of piles, two players take turns to remove any number of counters from a single pile. The winner is the player who takes the last counter (i.e. classic rule version). In particular, the NIM game is a strong solved game, so we know an algorithm that can produce perfect moves from any position, even if mistakes have already been made on one or both sides of the players.

Regarding Education, games contribute to better peer relationships and turn learning a more fluid and fun process. They also draw on users a natural competitive instinct and thus encourage participation and engagement. Here, gamification of digital games is a key point on any modern application, where it is provided small rewards and positive reinforcement. They allow to assess the users own knowledge finding areas for personal improvement. Games motivate and interest users to continue repeating a certain activity to reach a goal. For such reason, they seem to be a valuable tool for assessment and data gathering of behaviors in solving tasks.

One of the main references in combinatorial game theory is the book series “Winning Ways for your Mathematical Plays”, see [2], and the mathematical foundations of the field are provided by Conway's wonderful book [3], with the title “On Numbers and Games”. Many papers from the more recent collections “Games of No Chance” [url:1] and “More Games of No Chance” [url:2] are also available online. Since the literature about uses of combinatorial games in Education is quite large, we just omit it from this brief introduction.

This work describes the results of the first steps of a project, where a mobile application was developed to implement a digital version of the NIM game and was used as a mean to gather information that we believe will be useful to understand some learning patterns of users playing a combinatorial game. The work was made in the scope of GEOMETRIX, which is a strand line of the Center for research and Development in Mathematics and Applications, with an interdisciplinary-oriented research and development focus, targeted at assorted target groups (running from primary to higher education level), committed to the study, use and creation of intelligent digital environments to promote knowledge and skills in mathematics, reflecting a transformation in the way they are grasped and applied.
In the next section, we describe the general methodology, the features of the mobile client application, the aim of the online game server, and the data and mathematical tools used in the analyses. Clustering, variable choice, efficiency scores and relative ranking is also discussed in Section 2. In Section 3, the results are presented where three views are considered: a general characterization of the games played; the study of players’ efficiency; and the study of games and move control. The paper finishes with some conclusions.

2 METHODOLOGY

We intend to use digital versions of combinatorial games, as NIM, to collect, measure and analyze learning patterns. Therefore, after collecting enough game log statistics, we aim to understand and somehow classify user’s learning profiles by finding the adequate mathematical techniques from data mining/deep learning, linear optimization, and/or graph theory. For such reason, we collect basic statistics about players’ performance and their way of playing, for which the players are alerted at installation time. Fortunately, for our aims, OWS are usually difficult to grasp in a short time, so players tend to use Heuristics, which work as personal rules to approach the problem solving, learning, or discovery and usually employs a practical method not guaranteed to be optimal, but sufficient for the immediate goals or which they believe is effective from previous experience.

2.1 The Mobile Client App

We develop a free available Android mobile application for the NIM board game, that has now more than 150K games played from all over the World. Besides its recreational nature it is also intended to be a tool to research how an anonymous player deal with a combinatorial problem and be a data gathering mechanism as source for studying Learning Patterns. The application is available in the Android Play Store [url:3].

Besides the standard and expected features that a mobile game should have, the application has the following specific features:

a) Four rule variations: Classic, Misère, NIM 21 and Fibonacci;

b) Each game can have more than two players (e.g. against computer players and/or other humans);

c) A local multiplayer mode and an online multiplayer mode;

d) A ‘power-up’ is available which allows a player to break the adversary winning strategy, based on a harmonic oscillator equation;

e) Several achievements may be unlocked by completing game related goals (i.e. gamification techniques).

![Figure 1: Screenshots of the NIM GAME application available on the Android Play Store.](image)

In Figure 1, the first image is the main menu where users may choose the playing mode (single, local multiplayer, or online multiplayer), can access their achievements and tasks missions, access general
options, and have some brief help about the game and its rule variations. In the second image we see a standard game going on between two players. In the case, there are more than two players we added a perturbing mechanism to turn the outcome of the game more unexpected. Each player can once in the game, click the dice button, which will present him with the screen of the last image. In their it appears a pendulum gadget which oscillates in increasing velocity and when selected will add or remove the number of pieces for which is pointing at. In this way, situations as when we are at the final rounds and we realize that we cannot win, it still can be completely changed, which creates a alive experience for the multiplayers mode.

2.2 The Online Game Server

As a support system to the mobile client applications, there is a server daemon, developed in java, which allows the online multiplayer mode synchronization and the data harvesting. When installing the application, the users are informed about the possibility of allowing the application to gather statistical information about the games they play, being of key importance to undertake studies about general players’ behavior. Such information is anonymously transmitted and archived in the server and treated afterwards. The analysis, on Section 3, is done with a subset of that information.

2.3 Data and Mathematical Tools

In this work, we consider each player as an entity with the following data: players id (PID); number of games played (NGP); number of games won (NWon); number of games lost (NLost); number of games aborted (NLeaved), number of moves where the strategy where under the control of other player (NimCtrl), number of moves where the player had the possibility of making other player loss (NimLoss); win/lost weighted index for the first 10 games (WA); win/lost weighted index for the games played for 10<NGP<=50 (WB); win/lost weighted index for the games played for 50<NGP<=100 (WC); and win/lost weighted index for the games played for NPG>100 (WD). To compute the win/lost weighted index a game win count with +1 and a lost game with -1. When there are just two players, the field NimLoss contains the number of moves where the player had the possibility of taking control of the game.

2.3.1 Clustering and Optimal Variable Choice

Cluster analysis is the set of techniques to group a set of objects in such a way that objects in the same group share similar properties; in general, the similarity is measured by some kind of distance. As expected there are several ways to do such task. In this work, the cluster analysis is performed to determine which are the best groups of study and which variables are more relevant to compute the efficiency ranking. The best method found, for this situation, was the Partition Around Medoid (PAM), which is, in fact, the most common realization of k-medoid clustering, i.e. the algorithm is related to the k-means algorithm and the medoid shift algorithm [4]. The distance is the GDM2, i.e. the GDM distance proposed by Walesiak [5].

2.3.2 Efficiency Scores and Relative Ranking

Players efficiency somehow means in a general sense using less resources (i.e. attempts, moves, games played) to provide the same result (e.g. number of wins), so we may translate it mathematically as minimizing the input resources when producing the same output levels [6]. The common way to do this is to solve a multi-objective optimization problem. The most common relative ranking algorithm is the so-called Data Envelopping Analysis (DEA), proposed by Charnes [7]. In contrast to the standard DEA, in the Multidirectional Efficiency Analysis (MEA), proposed by Bogetoft [8], the input reduction and output expansion are selected proportionally to the potential improvements in efficiency identified by considering the improvement potential separately in each input variable and output variable. MEA turn out to be, in some precise sense, a better relative ranking algorithm than DEA.

In this work, we use MEA for the analysis of the players efficiency, considering as inputs variables: NGP, NLeaved, and NimCtrl; and as output variables: NWon and the complement of NimLoss. The variable complement is defined as the maximum value of the output variable in all the database minus the value of the variable for the entity under consideration. Recall that the choice of variables where obtained by the clustering algorithm.
3 RESULTS

This is the first report about the analysis of the output data of our NIM game. Considering that we have now more than 150k games played under four rule variations, six levels, played by up to 3 players (i.e. humans or AI), we need to restrict the analysis to a particular subset. In this publication, our database of study is composed of games played with the rule “Classic”, when it is one human playing against the computer. Furthermore, only players with more than nine games played are considered. This subset still gives the impressive numbers of 6.514 games and 29.667 moves to be analysed.

3.1 General Characterization

In this section we give a brief characterization of the games played and their final status. The left image of Figure 2 shows the percentage of players distributed by the total number of games committed. An average user of the NIM game plays 20 games. It is interesting to see the steady decrease of player between 30 NGP and 60 NGP, being 60 and 100 the number of games played with fewer users. Nevertheless, let us point out that there are players with more than 500 games played. As expected, users mainly choose lower difficult levels (i.e. easy levels), where the computer plays some percentage of the time just randomly and, the other part of the time, takes in consideration the optimal strategy. For such reason, we have that 78% of the games where won, 19% were lost and, as little, as 3% were leaved games, see the right image of Figure 2. Note that it is a game in mobile platform, thus it is common to put away the mobile at any moment. Hence, in fact, the low rate of leaved games tell us that the users where quite committed with the game.

![Figure 2: Percentage of players versus number of games played (left); Percentage of wins/lost/leaved games (right).](image)

3.2 Players Efficiency

As described in subsection 2.3.2, we analyse the players’ performance by finding its efficiency in playing though the calculation of efficiency scores that produce a relative ranking. Recall that the percentage of winning games is quite high, almost 80%. However, a game can be won because the opponent was not clever enough to use the optimal strategy in order to get control of the game, so the player had many chances during the game to do the correct moves. For such reasons, it makes sense to see how efficient the players were in its actions and opportunities to win. A more detailed analysis will be made in the next subsection by studying NimCtrl and NimLoss.

To realize the distribution of the efficiency scores computed, we considered classes of NGP of size 10, and for each class we computed the mean value (MES) and the standard deviation (SES). Note that the last class just gather together every NGP above one hundred. Figure 3 presents in a solid blue line the MES and the dashed grey lines are MES-SES and MES+SES, respectively. In the range of less than 30 NGP, there is some average efficiency but may be related with the situation where players play very few games in easy levels and just give up to try further more and risk to loss. From 30 to 50 NGP the average efficiency rank is quite low, and increases substantially when approaching 60 NGP. It may tells us that above 50 NPG is the number needed of NGP for an average player, which tries until then heuristic moves, to produce a mental strategy near to the mathematically optimal one. The
results above 60 NGP tend to decrease because those players, that play more games, usually try
difficult levels which decrease highly their performance.

Figure 3: Efficiency ranking distribution by NGP.

To clarify further the above arguments, in Figure 4, we see the efficiency range for each group (i.e.
<50 NGP and >=50 NGP); divided in the top ones (i.e. efficiency score above 75%), middle range, and
bottom ones (i.e. efficiency score below 25%). We may reason that the inefficiency for players above
50 NGP, which together with the information of Figure 3, imply that there are players just playing for
the fun without much care or worry in winning.

Figure 4: Efficiency range for each group comparing the most efficient with the less efficient.

3.3 Moves and Game Control in “Classic” mode

The NIM game is a combinatory game, which has a mathematical optimal strategy for winning. Using
the optimal strategy allows a player to control the final game outcome. Contrary to other games with
these properties, the NIM optimal strategy is not too obvious for a player to realize it in a few moves.
Such turn out to be quite good in the learning point of view, because the brain will try to “learn” the
way to win, usually, by finding heuristic/golden rules which the player tests and adapt.

The optimal strategy of the NIM game is based on the so-called Nim-sum. For each configuration of
the board, where there is a certain number of pieces in each pile, the Nim-sum is compute as follows:
(a) we write the number of pieces in each pile as a binary number; (b) we apply the XOR sum to each
binary position of the numbers obtained in (a); and (c) the binary number obtained in (b) is converted
again to decimal. The decimal obtained is the Nim-sum of this board configuration. Then, the optimal
strategy follows from the following two mathematical results.

Lemma. If the Nim-sum is 0 after a player’s turn, then the next move must change it.

Lemma. The winning strategy, in “Classic”, is to finish every move with a Nim-sum of 0.
So a player, that receives a board configuration with a Nim-sum which is not zero, is able to do a move that will turn the Nim-sum to zero. The opponent will not have any other possibility then doing some move, which destroy the optimal Nim-sum zero. In this way, the player controls the game and it is sure that, if it keeps playing like this, will win the game. The key point is that the majority of the players are unaware about an optimal strategy. They tend to construct heuristics as “keep the piles with the same number of pieces”. Why? Because at some point they realize that the configuration (2, 2, 0, 0, …) is a winning configuration. However, such is not a general rule which works for any number of pieces and piles.

![Figure 5: Number of times the game was controlled by other player (NimCtrl) and number of times the player had the opportunity to take control of the game (NimLoss) versus the number of games played.](image)

In Figures 5 and 6, the number of times the game was controlled by other player (NimCtrl) and the number of times the player had the opportunity to take control of the game (NimLoss) versus the number of games played are shown in scattering graphs. The first observation is that again the behaviour of the players, that play less than 50 games, is quite different from the ones that play more games. The figures show that the players had more changes to control the game than the game was already in control by the opponent (i.e. the game is not to difficult). The linear regressions show the positive slope, i.e. an affine proportional relation, between NGP and NimCtrl/NimLoss.

![Figure 6: Number of times the game was controlled by other player (NimCtrl) and number of times the player had the opportunity to take control of the game (NimLoss) versus the number of games played.](image)
4 CONCLUSIONS

In this work, we reported the first results of analysing the gameplay data produced by a mobile application, which is a digital version of the ancient Nim game. We applied modern algorithms for measuring efficiency of players' performance, which, in a sense, measure how near their strategy of playing where of the optimal (mathematical) strategy. The evolution of efficiency along the games played may be correlated with the learning process in some extend. In this context, we dealt with 6,514 games and 29,667 moves for producing the stated results.

Besides the conclusions already mentioned before, we added some new ones. In general, users were quite committed to finish the game, even if it was in a mobile platform where perturbation and breaks are quite common to happen, e.g. see Section 3.1. The variation between NimCtrl and NimLoss shown in Figures 5 and 6, shows that the game may be ranked as a nice overall game experience, since the level of difficulty, the number of players opportunities to take control, and the percentage of won games seem well-balanced. Data seem to show that there are players that, when in the mood, play one or two games and, by this way, accumulate games played by using regularly the App for killing time and enjoying the moment. On the other hand, there are players quite committed to the game with a quite high efficiency score. Looking in more detail, we see that there exist players that, even being quite efficient and winning several times, just continue playing. That may point to the possibility that the game somehow turn out to be rewarding.

Although the obtained results are quite interesting, they are preliminary and show that a future and deep analysis of the data is needed, under different approaches, in order to understand more learning patterns of the persons solving combinatorial tasks.

ACKNOWLEDGEMENTS

This work was partially supported by Portuguese funds through CIDMA – Center for Research and Development in Mathematics and Applications, and the Portuguese Foundation for Science and Technology FCT – “Fundação para a Ciência e a Tecnologia”, within project UID/MAT/04106/2013.

REFERENCES