Visualizing "Cognitive Fingerprints" from Simple Mobile Game Play

K. Gutenschwager[1], K. Leduc-McNiven[2], M. Aljumaili[2], R.D. McLeod[2], and M.R. Friesen[2]

[1] Institute of Information Engineering, Ostfalia University of Applied Sciences, Wolfenbüttel, Germany
[2]Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada

Abstract— **Serious Games and associated data analytics offer the potential of a complementary means of detecting early signs of mild cognitive impairment (MCI), which is often a precursor to more serious forms of dementias. As with all diseases and illnesses, the ability to mitigate the impact of the illness is directly correlated to early detection and intervention. In this work, a representative serious game is used to capture a "cognitive fingerprint" of a person's play, which is then used to analyze and visualize play. The long-term objective of the research is to demonstrate that data collected from serious games may be used to detect cognitive difficulties that may be pre-symptomatic, and outside the scope of normal age related cognitive decline. The present work assesses the viability of the platform for this purpose and opportunities in data visualization, but does not include clinical testing for MCI.**

*Keywords—**Serious games, mild cognitive impairment, machine learning, data visualization.***

I. INTRODUCTION

It is increasingly apparent that serious games are a becoming rich source of data associated with many aspects of mental health. Other researchers have surveyed current research and development of serious games for mental health data [1][2]. While there are other games on the market that claim to do 'brain training' to maintain cognitive function or potentially delay MCI, no game has been developed to *detect* MCI it – a subtle but important distinction. This work focuses on analyzing data from games toward MCI detection [3].

This paper briefly describes a serious game developed within the research group, denoted WarCAT, which illustrates a game's ability to capture a player's cognitive function related to strategy development, memory, recall, and other executive functions. We then briefly present an initial attempt to classify play using synthetic agents playing with various degrees of impairment. Limited success using various machine learning techniques prompted a more direct statistical approach of inferring cognitive ability by analyzing specific strategies a person may use or develop based on detailed moves made within the game. This process provided a data visualization of a person's "cognitive fingerprint" of play. The data visualization and rules of inference associated with play will be presented in detail with examples extracted from a small tournament.

II. WARCAT

WarCAT is a simple mobile game based on the familiar card game of WAR. Our variation is that a hand of 5 cards is dealt to the player, who can play them in any order they choose. The person is playing against a "bot" or machine that has also been dealt 5 cards (unseen by the player). The player and the bot both lay a card, and the higher card wins. The bot's strategy of play is to play its cards one at a time in descending order. There are other variations of bot strategy, but for the purposes here, this particular strategy is sufficiently representative. The player's challenge is to recognize that the bot is playing a strategy, and to consistently counter the strategy in one's own play.

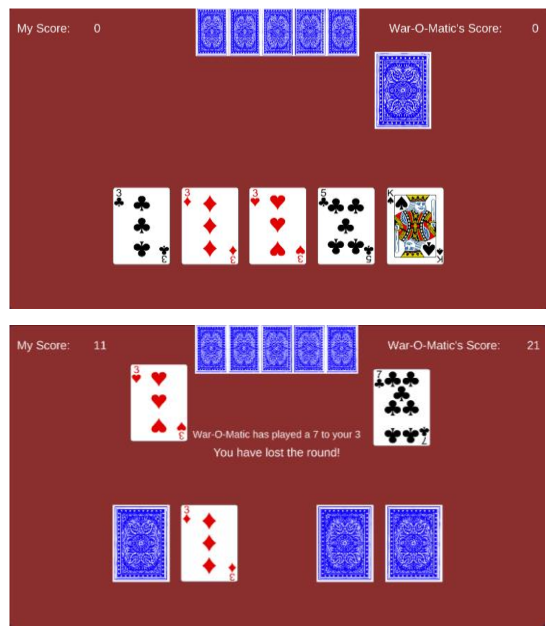

Fig. 1 Screenshot of one hand and instance of play.

Figure 1 illustrates a hand that was dealt and an instance of play. The player selected a 3 as the fourth card played against the bot's fourth card which was a 7. In this case, the bot won the hand. A game lasts for 50 (settable parameter) hands of play, at which time a player levels up. There are a programmable number of levels but for purposes of discussion, three will suffice. In the base version of the game, the bot does not change its strategy at subsequent levels. Other variations of the game have both changing strategy at subsequent levels, e.g. playing its cards in ascending order or other variations. At this time, as much player data as possible is collected, including the card played for each hand on both sides, whether the player won or lost the hand, the score, and timing information. From this, considerable information can be inferred as to whether (and how quickly) the player recognized and countered the bot's strategy, remembered the strategy to consistently beat the bot (when possible), or experienced points of confusion or momentary "loss of set" – i.e. forgetting the strategy, and then re-learning it (or not) at some later point in the game. As the cards are dealt stochastically, there is an inherent potential for confusion, for instance as a consequence of knowing that one played a winning strategy yet still lost the hand – which is a situation that happens approximately 25% of the time with the scoring system currently in use and playing a "good" strategy to beat the bot.

The scoring system scores a narrow margin of victory more significantly than by beating the bot by a wide margin. The score for each card played with a hand is:

$$score = 13 - (Winning\ card\ value - Losing\ card\ value)$$

unless there is a tie, in which case the score for that play is 0. This also supports strategy learning through reinforcement.

By contrast, a simple scoring system would award a point to the player or bot depending upon who simply had the higher card. The difficulty with simple scoring is that it is more difficult to play a winning strategy and more consistently win. With simple scoring, a person can expect to lose (as a consequence of the stochastic nature of the cards) approximately 35% of the time. It is easier to learn a winning strategy through reinforcement if one wins more often.

III. Machine Learning Forays

Initial attempts to use Machine Learning (ML) to classify play used synthetic input data generated by bot vs. bot play (computer agents), where one bot played at various degrees of impairment to mimic a human player with none-to-some MCI, while the other bot played their strategy consistently. Data were generated by bots that would play either a winning strategy or a random strategy governed by Bernoulli trials. For example, a bot would play at random with a probability p and a winning strategy with a probability of (1-p). The impairment model was improved to include impairment of a bursty nature using a process for modeling bursty channels [4].

The ML techniques investigated using synthetic data included dense neural networks (DNN) and convolutional neural networks (CNN) within the TensorFlow framework [5]. In both cases, the classification accuracy after training was in the mid-90%. Effectively, these results are a result of an accurate inference of whether the play was impaired or not. Having an accurate inference of whether the play was impaired of not would lead to similar (a few % less accurate) classification from simple statistics alone.

IV. Inferencing and Data Visualization

These observations lead to a more deliberate approach on attempting to differentiate impaired play from a player deliberately choosing an adaptive strategy with respect to the cards played by the bot. Even though WarCAT is simple and apparently straightforward, interpreting play (confused vs. adaptive) is difficult. This difficulty would be compounded if the game were any more complex.

A. Re-Interpreting Strategy

Much of the following is based on actual play and observations made during play, which has led to re-examining a person's optimal strategy. Knowing that the bot is always playing high to low cards, a reasonably good strategy which one could already consider as "perfect play" is to consistently play one's lowest card first followed by highest to lowest. In this strategy, the first card is sacrificed to see the bot's highest card. This strategy leads to wins approximately 75% of the time, but doesn't account for more subtle decision making during the hand. However, there are other strategies which may also lead to equal or better results. In order to evaluate the 120 possible basic strategies, defined as the possible permutations of the five cards, we performed a simulation experiment where 10M games were played against a bot that always plays its cards high to low. We further analyzed the results by storing information about the lowest and the highest card value, as more advanced players might use this information as input to a more adaptive strategy. Our results show that looking at all 78 possible combinations of lowest and highest card, 23 different permutations of cards lead to the best results at an average for at least one combination of lowest and highest card.

TABLE I: Winning strategies with respect to the lowest and highest card. Note that other strategies are minor variations.

ID	Sequence	Description
1	**1**-**2**-5-4-3	Sacrifice first two lowest cards at first and second position.
2	**1**-5-4-3-2	Sacrifice lowest card at the first position.
3	**1**-5-**2**-4-3	Sacrifice first two lowest cards, but the second best card at third position.
4	5-4-3-2-1	Play high to low
5	5-**1**-4-3-2	Sacrifice lowest card at the second pos.
6	**2**-**1**-5-4-3	Sacrifice first two cards (second best first)
7	5-4-3-**1**-2	Sacrifice lowest card at the forth pos.
8	5-4-**1**-3-2	Sacrifice the lowest card at the third pos.

If a player always plays an adaptive strategy, then the respective sequence to play their cards based on the present combination of the lowest and highest card (in their hand) leads to the percentage of wins increasing to about 83.4%. The main eight of the 23 strategies are given in Table I.

If we reduce the number of different strategies to four (IDs 1, 2, 4, 5), we obtain a rather manageable set of strategies and combinations of lowest and highest cards and when to use them. The percentage of games won is still very high with a value of about 82.7%.

Note that if a player also takes into account the first card played by the bot and then adapts their strategy based on that card, the chance of winning further increases by about 1%.

Looking at the limited number of games a player completes in one round, the difference between these two approaches is hardly recognizable and statistically insignificant. We therefore differentiate a few main strategies that actually lead to statistically significantly different results which a player may actually recognize playing only a few rounds.

Table II estimates the expected points awarded based on the strategy played. For example, if a player played at random, the average number of points expected would be 0. If they developed a strategy of consistently sacrificing their first card followed by playing their cards from high to low, their expected points per hand would be 16.6. Within one level, a person would play 50 hands, and if they proceeded to play that strategy, they would end up with approximately 50*16.1 points.

TABLE II: Expected points per hand, based on strategy played

ID	Strategy	Avg. points
1	Random	0
2	High to low	0
3	Low to high	0
4	Sacrifice first card, rest high to low	16.6
5	Sacrifice first two lowest cards, rest high to low (lowest at pos. 1)	16.1
6	Adaptive (1): If the value of second lowest card < 8 then sacrifice two cards, else one	18.1

Overall, we can vary the strategy a player follows as the first step, and then analyze if the player gets distracted from that strategy over time. However, there are a number of complexities that need to be taken into account. A player may tend to deliberately change strategy over time to improve their overall score. These variations make it more difficult to infer moments of confusion or loss of set.

The first three strategies can be differentiated by the sequence played, even though the average score does not differ. One can use a simple distance measure for comparing two sequences played, denoted as P1 and P2, where pos_j states the position of the j-lowest card in the sequence:

$$dist = \sum_{j=1}^{5} abs(pos_j^{P1} - pos_j^{P2}) \quad (1)$$

Computing the average distance for the last n games to the (best) sequence of the respective main strategy (2-7), one can analyze the current distraction as a deviation from the optimal sequence of the strategy the player follows momentarily (except for random play). Taking into account the average score obtained using a specific strategy, one can further compare the average score the player achieved with the respective expected value. This information might give further insights into why a player changes strategy, i.e., if the actual score is lower than the expected value (computed by our simulation experiments), there might be a higher probability that a player got distracted from the chosen strategy.

The classification on whether and how significantly a player's behavior reflects a cognitive impairment needs to be corrected by this influence. Furthermore, one needs to take into account a learning curve, such that we compare a player's current play to the best strategy learned so far. This means that the "perfect play", as introduced above, is actually to be considered as dynamic if one wants to judge the influence of cognitive decline.

B. Visualization of a Cognitive Fingerprint

In order to better understand the patterns of play, a technique was developed to visualize an inferred strategy, duration of play at that level, and overall score achieved during play (see Fig. 4 for further explanation)..

Fig. 4: Visualization of winning play

To illustrate this with real data, a small tournament was organized early December 2018 to play for fun with an incentive of the high score receiving a $100 prize. This was an informal tournament that allowed the collection of a small player pool of anonymous data. The objective was simply to explore the visualization of data extracted from play. The tournament consisted of games of 3 levels of play where the bot would play its cards from high to low at each level. Each level consisted of 50 hands. Players were allowed to play multiple times. University research ethics approval was obtained to assess the game hedonics, software verification, and to illustrate the interpretation and visualization of game-play. The tournament was not carried out to assess MCI.

Based on the inferences of Section IV(*A*) the data was displayed on a polar plot with coordinates governed by the time of the hand and an inference of the strategy of play. Figure 4 illustrates play by the winner of the tournament. In this case, most of the play was inferred to be Advanced, with only a small portion inferred as Random or Basic.

Another example of play is visualized in Fig. 5. The player was an adolescent who had never seen the game previously. Various strategies of play can be seen as the person quickly developed more advanced strategies.

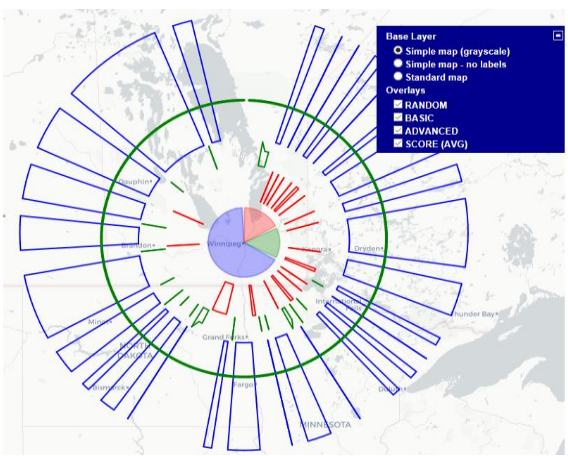

Fig. 5: Visualization of novice play

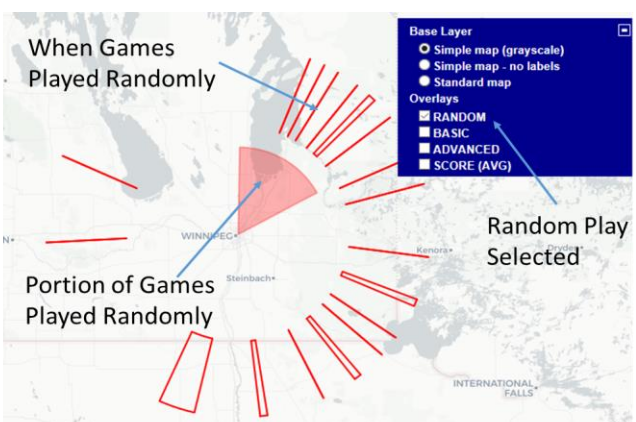

Fig. 6: Visualization of random play during game

Figure 6 illustrates a selection of inferences made of Random or Basic strategies of play, and these features that are conjectured to be useful as ML inputs. As with many ML applications, data preprocessing is essential in obtaining useful and accurate predictions. Figure 7 illustrates a visualization of the best play seen to date.

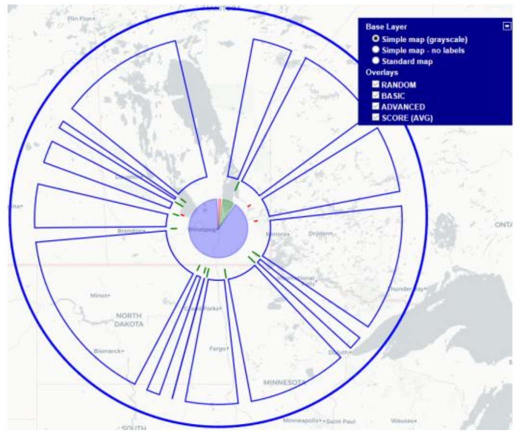

Fig. 7 Visualization of best play to date

V. Conclusions

Building serious games for MCI detection and attempting to extract players' patterns of play to assess cognitive approaches during play is very difficult. Even at the level of our very simple game, analyzing data is complex. These complexities will be compounded by the subtleties of MCI and the real (future) challenges of labeled data. Future work will be in developing agents or bots that will learn to beat the game consistently, for the purposes of emulating human play. Depending upon the degree of agent training, it may be possible to label various levels of play or impairment. These type of synthetic data will be visualized with the methods developed here as well as with CNNs for more accurate predictions of cognitive impairment distinct from normal cognitive decline.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

1. McCallum S. Gamification and serious games for personalized health. Stud Health Technol Inform. 2012 Jan;177(2012):85-96.
2. Muscio C, Tiraboschi P, Guerra UP, Defanti CA, Frisoni GB. Clinical trial design of serious gaming in mild cognitive impairment. Frontiers in aging neuroscience. 2015 Mar 11;7:26.
3. Petersen RC, Smith GE, Waring SC, Ivnik RJ, Tangalos EG, Kokmen E. Mild cognitive impairment: clinical characterization and outcome. Archives of neurology. 1999 Mar 1;56(3):303-8..
4. Elliott EO. Estimates of error rates for codes on burst-noise channels. The Bell System Technical Journal. 1963 Sep;42(5):1977-97.
5. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.