Model-based Evaluation of Playing Strategies in a Memo Game for Elderly Users

Felix Putze, Tanja Schultz
Karlsruhe Institute of Technology
Cognitive Systems Lab
Karlsruhe, Germany
felix.putze@kit.edu

Sonja Ehret, Heike Miller-Teynor, Andreas Kruse
University of Heidelberg
Institute for Gerontology
Heidelberg, Germany
sonja.ehret@gero.uni-heidelberg.de

***Abstract*—In this paper, we analyze game protocols for a Memo game for elderly users. We show how we can use generative statistical models to automatically reveal different playing strategies. We present a quantitative and qualitative evaluation of the approach on simulated and real data. We show that we can reliably detect different strategies and that we can use those strategy profiles to uncover relevant information on the players beyond pure performance measures.**

***Index Terms*—mental model; memory; cognitive and social activation; elderly users; dementia**

I. Introduction

In the aging society, the number of people with dementia is steadily growing, being estimated at 35.6 million people today and projected to double by 2030 [1]. Among other cognitive deficits, people with dementia suffer from impaired memory, which affects their ability to maintain an independent life. It is long known that cognitive and social activation (e.g. reminiscence therapy, game playing, etc.) of people with dementia has a positive effect both on their mental health and their quality of life [2]. The digital implementation of activation paradigms brings with it a number of benefits, for example multimedia presentation of dynamic content and the potential for adaptation to individual needs. Another major advantage of digital activation applications is the ability of collecting detailed personalized information on the progress of an activity in automatically generated protocols. The general positive effect of such activation is known among caregivers and practitioners and has been shown and quantified in many lab studies [3]. However, there still is a lack of individual assessment of the outcome of an activation paradigm. One tool which will help to close this gap is the analysis of automatically generated activation protocols.

There exists a number of research projects which aim at the cognitive or social activation of elderly people with and without dementia. For example Smith et al. [4] described the development of a reminescence tool in form of individual multimodal life stories on DVDs to induce positive emotions. Webster et al. [5] presented the "Portrait" system as a tool for evoking emotional reactions and triggering episodic memory. The designated purpose for this system is the bonding process between caregivers and elderly people by presenting biographic content. The CIRCA system ("Computer Interactive Reminiscence and Conversation Aid") by Gowans et al. concentrates on positive psycho-social effects of computer-based reminiscence therapy in group sessions [6]. It was refined by Botella et al. [7] described the improvement of the CIRCA system following a user-centered design approach. Alm et al. [8] could show that the CIRCA system evoked positive effects which were absent for a traditional reminiscence therapy. Other research concentrates more on the quantifiable training of cognitive capabilities. Buschkühl et al. [9] showed that a intensive memory training can yield generalizable improvements in cognitive ability for people even at high age. Basak et al. [10] showed similar results when training with a complex computer game.

While this selection of related work shows that computer-aided cognitive and social activation is beneficial to elderly people in multiple ways, we found that those works concentrate on one of two aspects in their analysis: Either, they perform a purely high-level observational analysis or they focus purely on the data-driven, quantitative evaluation of performance metrics, e.g. to show an improvement in measurements of cognitive ability. In this paper, we aim for a combination of both perspectives: We will look at quantitative data which is automatically collected during the execution of an activation, but we will use this data to look beyond pure performance metrics. Instead, our goal is to uncover information about the mental model a participant has of a certain activity.

In this paper, we therefore present a novel statistical approach to analyze the automatically generated protocols of computer-based activation paradigms. Specifically, we look at data generated from a multimedia Memo game, in which players take turns to reveal cards from a tableau to find card pairs with matching content. A Memo game is a both, a social activity as well as a cognitive stimulation. When used with carefully selected prototypical or individual cards, a Memo game can stimulate long-term memory or evoke emotional responses. Stimulating long-term memory of a person, e.g. during reminiscence therapy, has a longstanding and successful tradition in working with elderly people with and without dementia [11].

Using this Memo game, our main focus is not to document improvements in cognitive abilities by training, but to use automatic data analysis to arrive at a better understanding of the mental model that individual users have of activation paradigm and how this model drives their behavior during the activity. Such an understanding would allow caregivers to draw conclusions on strong changes in the cognitive representation and understanding of the environment (e.g. when the mental model changes) and to identify optimal activation paradigms for individuals (e.g. when a person is not interested in highly structured games but still enjoys certain aspects of those activities).

II. The Memo game

For cognitive and social activation of elderly people with and without dementia, we developed a Memo game based on a tablet computer. Tablet computers with touch input have already been shown to work well with elderly users [12]. The Memo game can be played with one or two players, variable numbers of cards and allows the inclusion of multimedia content on the cards, e.g. pictures, sound files or videos. It is also possible to include individualized content, e.g. personal photos. Sounds and personalized speech prompts were used to guide the players during independent play, e.g. to notify about success, errors or to indicate the currently active player. Similar approaches for tablet-based Memo games (for example MemoCare [13]) exist, but developing our own Memo application gave us the opportunity to generate detailed protocols of all game turns for the corresponding player profiles. Figure 1 shows a screen shot of an ongoing game.

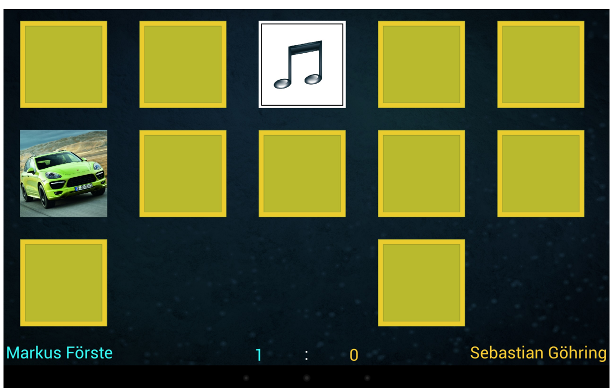

Fig. 1. Screenshot of the Memo game user interface.

For the analysis of played games, one might at first be temped to use easily accessible game metrics to draw conclusions about a player's performance or progress during a game: Number of discovered pairs, response time or total playing time can be directly calculated from game protocols and provide a superficial impression of cognitive performance within the game. However, those metrics can easily be misleading: For player with dementia, winning the game might not always be in the focus of the activity. Instead, they might enjoy looking at the cards, or engage in the social aspect of the game by talking about the cards or the game progress. As a consequence, there are players which do not follow the "intended" playing strategy of the game. This does not imply that those players do not act according to a plan or a model of the game. On the other hand, there may be players which do follow the intended playing strategy but suffer from impaired short-term memory. In both cases, using the mentioned game metrics would classify such players as "low performers" without further differentiation.

The goal of this work is to provide a statistical framework to uncover such different playing strategies. For this purpose, we use a number generative statistical models of different game playing strategies. Those strategies are evaluated against recorded game and help to identify the most likely strategy played in a certain game.

III. Model-based Strategy Profiling

As stated in the previous section, we expect not all players to actually follow a strategy which aims at the most efficient way to clear the card tableau by removing pairs. In this work, we differentiate between four distinct strategies which we try to identify from recorded game protocols.

Those four strategies were identified by gerontology experts from theory and from observations of live plays and video recordings: For a majority of players, we observed compliance to the declared and intended rules of the game. This behavior was driven by group conformity, a desire to win or the wish to demonstrate that the cognitive challenges offered by the game can still be tackled by the player. However, other types of playing strategies could be observed frequently: Some players follow a mirror strategy, copying the actions of the other player or repeating their own previous actions (imitation behavior [14]). Such behavior may have its origin in insecurity of the player or in the desire to reveal certain content (e.g. with a connection to the personal background or interests). Another distinct type of playing strategies defines individual ordering principles according to which cards are revealed, i.e. following certain spatial patterns. This is in accordance with the theory of Goldstein [15] who claims that ordered behavior can often be observed as compensation for cognitive deficits. Finally, there were players which did not seem to reveal cards following a systematic strategy but selected cards randomly. In the following, we will formalize models to automatically identify by which strategy a player chose the cards to reveal in a given game.

To automatically identify the strategy a player is following, we use a model-driven approach using the data from a game protocol (which contains the sequence of game turns for each player). For every strategy, we define a generative model which assigns a *reveal probability* to each available card in a tableau. To calculate the probability of a strategy to have generated a given game turn sequence, we iterate over all turns of this game for one player. For each turn, the reveal probability $P(C|H)$ for the card C picked by a player in this turn (given the game history H) is calculated[1]. Those probabilities are accumulated for all game turns. To explicitly assign one strategy to a game, we can select the maximum likelihood strategy. However, in many cases it is useful to look at the complete *strategy*

[1]While we calculate the probabilities separately for each player, the history contains game turns of both players as cards are revealed publicly

profile, i.e. the normalized probability distribution assuming equal priors for the strategies.

The Win strategy is the one which follows the intended goal to win the game by revealing matching cards. The Win strategy will try to reveal known pairs (*exploit*) or – if no pairs are known – to gather more information about unknown cards (*explore*). However, the memory of a player is not perfect, i.e. players do not remember every pair which is potentially available. This fact is especially true when modeling players with memory impairment. To represent this imperfect memory within the strategy, we implement the formula for *base level activation* from the cognitive architecture ACT-R [16], see Equation 1. Base level activation of a memory item (a card identity in our application) depends on the frequency and recency of learning events related to a card. As a learning event, we treat each game turn (by any of the players) which revealed that card. In Equation 1, t_i is the age of learning event i and d is the decay parameter. From base level activation, we can calculate the *retrieval probability* of an item, see Equation 2: The higher the activation, the more likely the item is to be retrieved from memory (depending on the threshold parameter τ and the noise parameter s).

$$A(C|H) = \log \sum_{i=1}^{n} t_i^{-d} \tag{1}$$

$$P_{BL}(C|H) = \frac{1}{1 + e^{-(A(C|H)-\tau)/s}} \tag{2}$$

Using this retrieval probability, we can define the Win strategy as follows (see Equation 3 for the reveal probability of a card): A player can either exploit available knowledge on pairs of cards or explore new cards. If exploiting, the strategy prefers cards for which the partner card is also known (see Equation 4). If exploring, the strategy prefers cards of which the identity is not yet (or not anymore) known (see Equation 5). The weighting between exploration and exploitation is performed by Equations 6 and 7: Exploration is only likely if no pairs can be retrieved with high probability.

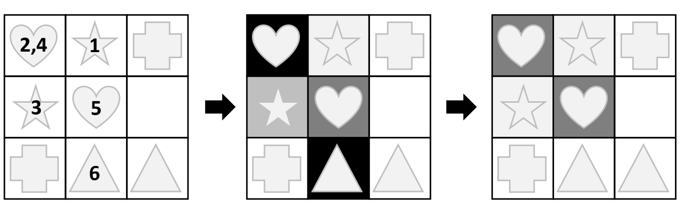

Fig. 2. Win strategy: Sequence of card selections (left) results in a certain retrieval probability (center) which results in a certain reveal probability (right) for each card. Darker shading indicates higher probability.

$$\begin{aligned} P(C|H) = &P(exploit|H) \cdot P_{exploit}(C|H) \\ &+ P(explore|H) \cdot P_{explore}(C|H) \end{aligned} \tag{3}$$

$$P_{exploit}(C|H) = \frac{\min(P_{BL}(C|H), P_{BL}(Partner(C)|H))}{\sum_{i=1}^{n} \min(P_{BL}(C_i|H), P_{BL}(Partner(C_i)|H))} \tag{4}$$

$$P_{explore}(C|H) = \frac{1 - P_{BL}(C|H)}{\sum_{i=1}^{n} 1 - P_{BL}(C_i|H)} \tag{5}$$

$$P(explore|H) = \prod_{i=1}^{n} 1 - P_{exploit}(C_i|H) \tag{6}$$

$$P(exploit|H) = 1 - P(explore|H) \tag{7}$$

In contrast to the Win strategy, the High-Activation strategy does not follow the intended goal of the game to quickly identify pairs of matching cards. Instead, it is based on the principle that individual cards which are familiar to the player are flipped, regardless of the knowledge of a matching partner card. While being different from the Win strategy, High-Activation shares the base level activation mechanism with it. Instead of deciding between exploration and exploitation, the High-Activation simply prefers cards with high retrieval probability, i.e. the reveal probability is identical to P_{BL} in Equation 2. The difference can be seen when comparing Figures 2 and 3, which show the different probability patterns resulting from an identical game history.

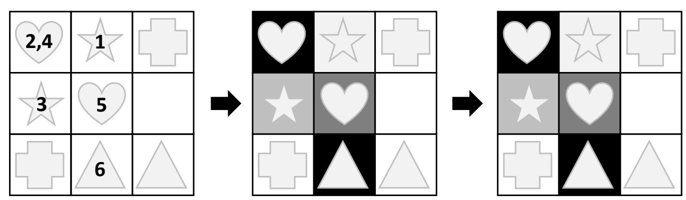

Fig. 3. High-Activation strategy: Sequence of card selections (left) results in a certain retrieval probability (center) which results in a certain reveal probability (right) for each card.

The Neighbor strategy does not make use of base level activation values. Instead, it implements the behavior of choosing a card which is spatially close to the last revealed card. This strategy represents any behavior which follows spatial patterns to iterate the cards (e.g. in row-order) but is also flexible enough to allow for more irregular spatial patterns or changes within one pattern over time. Figure 4 shows show a history of game turns translates to assigned probabilities.

Finally, there is the Random strategy which acts as a baseline strategy. Random disregards the history of game turns and assigns uniform probability to all remaining cards on the tableau, see Figure 5.

The integration of base level activation already accounts for errors made by the players due to limited short-time memory. However, this mechanism does not account for all kinds of errors. For example, a player might remember the previous occurrence of two matching cards, but be unsure about their

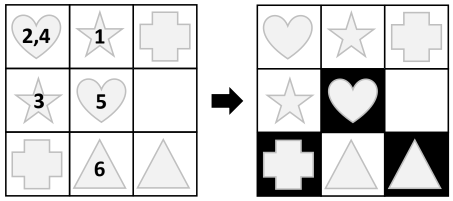

Fig. 4. NEIGHBOR strategy: Sequence of card selections (left) results in a certain probability (right) for each card.

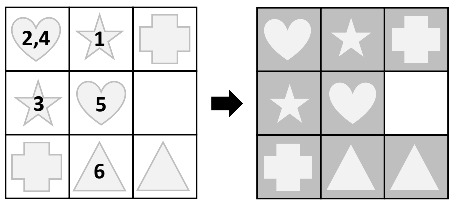

Fig. 5. RANDOM strategy: Sequence of card selections (left) results in a certain probability (right) for each card.

exact location. In that case, the player will not guess with a uniform distribution between all cards. Instead, the player (and consequently, our model) *spreads* the probability of a card to its neighboring cards. Formally, a fixed fraction of the reveal probability of each card is distributed equally among all (vertical or horizontal) neighbors and is added to their final reveal probability, see Figure 6. This mechanism is designed in analogy to the spreading activation mechanism of the ACT-R memory model [16].

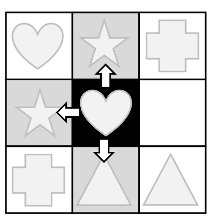

Fig. 6. Illustration of the probability spreading mechanism: A fixed fraction of the activation of the central card is distributed across its neighbors.

IV. EVALUATION

To evaluate the proposed approach, we perform two different steps: First, we perform a quantitative evaluation on simulated data. This allows us to show that the approach is actually able to robustly discriminate different playing strategies. Second, we perform a quantitative evaluation on real games with elderly players and players with dementia.

A. Evaluation on Simulated Data

For the quantitative evaluation, we simulated 100 games for each strategy with varying number of initial cards, with varying number of players and varying player performance (probability for errors) for each strategy. Each simulated game resulted in one game protocol which was analyzed using the model-based approach. For evaluation, we look at the averaged strategy profile and the resulting maximum likelihood strategies and compare those results to the expected outcome.

We applied the strategy profiling to the resulting data and evaluated, whether selection extracted the correct strategy reliably. The games were not generated by the strategy models in the profiling component. Instead, we implemented a separate game generator to produce typical game turn sequences for each strategy. To increase the variability of the generated data, the simulated games contained a number of "untypical actions" (e.g. random actions for non-random strategies) or different skill levels (e.g. different memory horizons for the WIN strategy). This was done to avoid an overly optimistic result due to overfitting.

Table I shows the resulting probabilities for the different strategies for the different configurations. We see that the correct strategy is usually identified correctly with a large margin compared to the second best estimate. We observe (for example when examining the probabilities for the WIN strategy) that for games with more cards (i.e. longer games), the margin increases compared to games with fewer cards. This can be explained by the fact that in longer games, there occur more opportunities to identify the characteristic game actions for each strategy. For example, during the first turns of a game, WIN will resemble the RANDOM strategy as it depends on the knowledge of complete pairs.

The RANDOM strategy is not identified as well as the others when following a maximum likelihood approach, i.e. the difference between the probability for the correct and the second-best strategy is much smaller than for the other three generated strategies. We explain this by the observation that the model for the RANDOM strategy predicts a completely uniform probability distribution and therefore cannot be expected to as selective as the other models. Instead, robust identification of a random play can be performed by calculating the entropy of the strategy probability distribution: Random play can then be identified by a value close to $\log(\#strategies)$.

B. Evaluation on Real Data

In this subsection, we present results from the analysis of game protocols collected in two centers for elderly day care in the region of Heidelberg. At both locations, a tablet with the Memo game was present for 12 weeks. In total, we had 12 participants (6 male, 6 female) with a mean age of 83.8. The participants played a total of 189 games. The evaluation process was accompanied but not guided or controlled by two researchers.

For such realistic but uncontrolled data, a systematic quantitative evaluation of the strategy profiling is difficult, as we have no ground truth on the chosen strategies. Furthermore, the participants in the study were allowed to chose card sets, number of cards and number of players according to their preferences. This increased the ecological validity of the data, but also makes quantitative comparisons between different participants difficult. Therefore, we pursue a more qualitative perspective in which we look at a number of case studies which focus on different players or game configurations. While the

Simulated Strategy	# cards	P(WIN)	P(HIGH-ACTIVATION)	P(NEIGHBOR)	P(RANDOM)
WIN	6	**0.47**	0.09	0.25	0.19
WIN	12	**0.6**	0.1	0.18	0.11
WIN	24	**0.65**	0.12	0.13	0.09
HIGH-ACTIVATION	6	0.2	**0.37**	0.21	0.21
HIGH-ACTIVATION	12	0.21	**0.4**	0.19	0.19
HIGH-ACTIVATION	24	0.23	**0.41**	0.18	0.19
NEIGHBOR	6	0.15	0.19	**0.44**	0.21
NEIGHBOR	12	0.1	0.17	**0.6**	0.12
NEIGHBOR	24	0.09	0.16	**0.68**	0.08
RANDOM	6	0.28	0.21	0.22	**0.28**
RANDOM	12	0.25	0.24	0.24	**0.27**
RANDOM	24	0.25	0.24	0.24	**0.27**

TABLE I
AVERAGED STRATEGY PROFILES FOR DIFFERENT TYPES OF SIMULATED GAMES. THE MAXIMUM LIKELIHOOD STRATEGY IS PRINTED IN BOLD FONT.

evaluation on the simulated data showed the reliability of the strategy profiling, the evaluation on the real data aims at the information we can uncover using the strategy profiles.

Case Study 1: In this case study, we evaluated the data of participant A with mild cognitive impairments. In her youth, she regularly and successfully participated in gymnastics competitions. The caregivers described her as competitive and as having the tendency to compare herself to others. A's game performance was not good: with 0.24 retrieved pairs per turn on average, A showed the worst performance of all players for medium card count. Still, the strategy profile classified her games consistently as following the WIN strategy, see Figure 7. Of all players, A exhibits the second-largest difference between the maximum likelihood strategy and the second-most likely strategy. This result shows that A was able to consistently conform to the rules of the game, despite deficits in short-term memory which prevented her from achieving a higher game score. We conclude that looking only at performance may not always yield a full picture of the players' mental model.

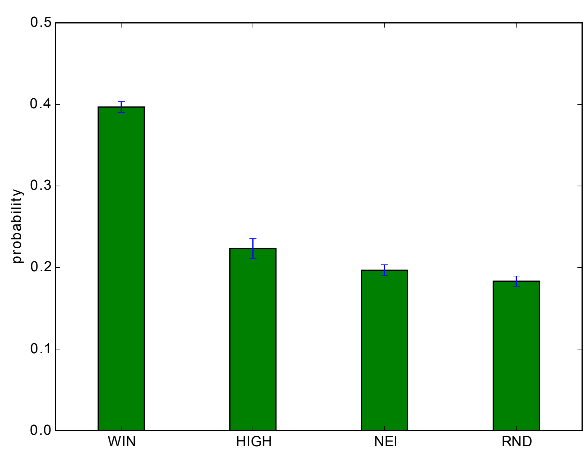

Fig. 7. Strategy profile of participant A. Whiskers indicate standard deviation.

Case Study 2: In this case study, we evaluated the data of a participant B, diagnosed with dementia. During the repeated game sessions, the caregivers explained the game's principle to B. Despite those efforts, B lacked the ability or the interest to follow those instructions. The probabilities assigned by the strategy recognition reveal this behavior: In contrast to the diagram for A, we see that the NEIGHBOR strategy is selected as the maximum likelihood result of the strategy profile. This result can be interpreted that B does not follow the intended game principle, but was still able to exhibit a consistent, self-defined behavior. An analysis purely based on performance metrics would have simply identified B as a bad player of the (implicitly assumed) WIN strategy.

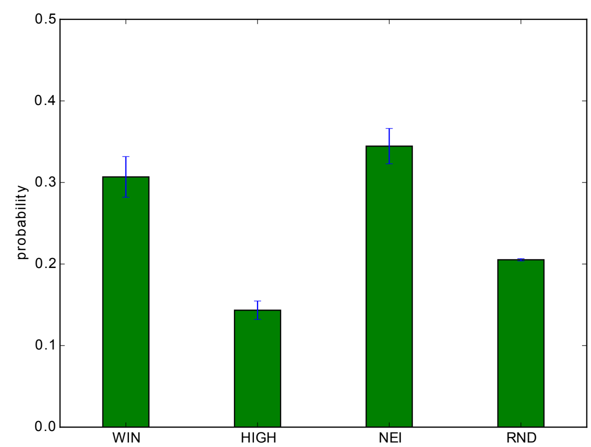

Fig. 8. Strategy profile of participant B. Whiskers indicate standard deviation.

Case Study 3: In this case study, we evaluated the data of participant C, diagnosed with dementia. When looking at C's strategy profile in Figure 9, we observe a probability distribution which is similar to the patterns in I. The probability of the RANDOM strategy is second-highest. Among all participants, C's strategy profile exhibits the highest entropy. This shows that C did not follow the intended rules of play, but in contrast to participant B, did also not follow a different consistent pattern.

Case Study 4: While the first three case studies concentrated on individual players, we now compare game protocols across all participants of the study. For this purpose, we compare games which were played with standard card sets with games which were played with individualized card sets.

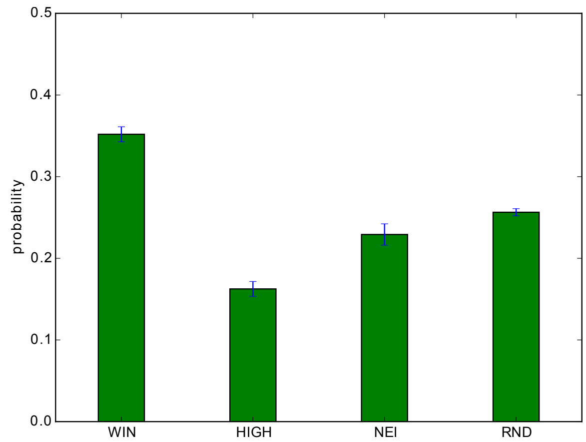

Fig. 9. Strategy profile of participant C. Whiskers indicate standard deviation.

Figure 10 shows the probability of the different strategies to occur as the maximum likelihood result for games with standardized card sets (red) and individual card sets (blue). Most prominently, we see that while the most frequently selected strategy is still WIN, the relative frequency for the HIGH-ACTIVATION and the RANDOM strategy are more than doubled for the individualized card sets compared to the standard card sets. We take this observation as an indication that for individualized cards, the wish to reveal, experience and share personal content increased. Simultaneously, the urge to conform to the intended rules of the game decreased. This is also in accordance with our observations during the performed games: during games with individual cards, the amount of biography-related discussion among players was substantially higher than for games with standard cards.

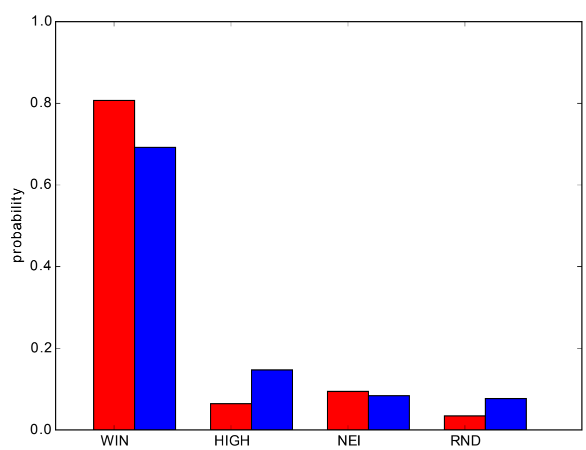

Fig. 10. Relative frequency of strategies to occur as maximum likelihood result for standard card sets (red) and individualized card sets (blue).

V. Conclusion

In this paper, we presented a novel approach to analyze game protocols by extracting the mental model of the players in form of their playing strategy. This statistical generative approach is able to discriminate between different playing strategies, even if metrics on game performance are similar. On simulated data, we showed that strategies can be recovered reliably. On the real data, we used the method to identify and discuss different strategy profiles. The presented approach can be extended in several different ways. First, we plan to extend the approach by looking at a model which does not treat all cards of a game identical. This would enable us to discriminate between different card types within one game (e.g. sounds vs. pictures) or to discriminate between familiar and unknown content. Second, we can analyze the model parameters for individual players as interpretable descriptors of cognitive functions and combine this information in a multiple-criteria approach with observational data and game statistics.

References

[1] W. H. Organization and others, *Dementia: a public health priority*. World Health Organization, 2012.

[2] M. Vernooij-Dassen, E. Vasse, S. Zuidema, J. Cohen-Mansfield, and W. Moyle, "Psychosocial interventions for dementia patients in long-term care," *International Psychogeriatrics*, vol. 22, no. 07, pp. 1121–1128, 2010.

[3] H. Brodaty and C. Arasaratnam, "Meta-analysis of nonpharmacological interventions for neuropsychiatric symptoms of dementia," *American Journal of Psychiatry*, vol. 169, no. 9, pp. 946–953, 2012.

[4] K. L. Smith, M. Crete-Nishihata, T. Damianakis, R. M. Baecker, and E. Marziali, "Multimedia Biographies: A Reminiscence and Social Stimulus Tool for Persons with Cognitive Impairment," *Journal of Technology in Human Services*, vol. 27, no. 4, pp. 287–306, Nov. 2009.

[5] G. Webster, D. I. Fels, G. Gowans, and V. L. Hanson, "Portraits of Individuals with Dementia: Views of Care Managers," in *Proceedings of the 25th BCS Conference on Human-Computer Interaction*, ser. BCS-HCI '11. Swinton, UK: British Computer Society, 2011, pp. 331–340.

[6] G. Gowans, J. Campbell, N. Alm, R. Dye, A. Astell, and M. Ellis, "Designing a Multimedia Conversation Aid for Reminiscence Therapy in Dementia Care Environments," in *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, ser. CHI EA '04. New York, NY, USA: ACM, 2004, pp. 825–836.

[7] C. Botella, E. Etchemendy, D. Castilla, R. M. Baos, A. Garca-Palacios, S. Quero, M. Alcaiz, and J. A. Lozano, "An e-Health System for the Elderly (Butler Project): A Pilot Study on Acceptance and Satisfaction," *CyberPsychology & Behavior*, vol. 12, no. 3, pp. 255–262, Jun. 2009.

[8] N. Alm, R. Dye, G. Gowans, J. Campbell, A. Astell, and M. Ellis, "A Communication Support System for Older People with Dementia," *Computer*, vol. 40, no. 5, pp. 35–41, May 2007.

[9] M. Buschkuehl, S. M. Jaeggi, S. Hutchison, P. Perrig-Chiello, C. Dpp, M. Mller, F. Breil, H. Hoppeler, and W. J. Perrig, "Impact of working memory training on memory performance in old-old adults," *Psychology and Aging*, vol. 23, no. 4, pp. 743–753, 2008.

[10] C. Basak, W. R. Boot, M. W. Voss, and A. F. Kramer, "Can training in a real-time strategy video game attenuate cognitive decline in older adults?" *Psychology and Aging*, vol. 23, no. 4, pp. 765–777, 2008.

[11] B. Woods, A. E. Spector, C. A. Jones, M. Orrell, and S. P. Davies, "Reminiscence therapy for dementia," in *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd, 1996.

[12] M. Kobayashi, A. Hiyama, T. Miura, C. Asakawa, M. Hirose, and T. Ifukube, "Elderly User Evaluation of Mobile Touchscreen Interactions," in *Human-Computer Interaction INTERACT 2011*, ser. Lecture Notes in Computer Science, P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, and M. Winckler, Eds. Springer Berlin Heidelberg, Jan. 2011, no. 6946, pp. 83–99.

[13] N. Neuroscience, *Ratgeber Demenz fr Betreuer und Angehrige*. Plejaden Communication GmbH & Co KG, 2006.

[14] B. Dubois, A. Slachevsky, I. Litvan, and B. Pillon, "The FAB A frontal assessment battery at bedside," *Neurology*, vol. 55, no. 11, pp. 1621–1626, 2000.

[15] K. Goldstein, *Der Aufbau des organismus*. M. Nijhoff, 1934.

[16] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, "An Integrated Theory of the Mind," *Psychological Review*, vol. 111, no. 4, pp. 1036–1060, 2004.