Proceedings of the Workshop on The Cognitive Science of Games and Gaming July 26, 2006
Vancouver, British Columbia, Canada

Embedded Assessment of Cognitive Performance with Elders' Use of Computer Games in a Residential Environment

Holly Jimison (jimisonh@ohsu.edu)
Medical Informatics
Oregon Health and Science University,
3181 SW Sam Jackson Park Road Portland, OR 97201
Spry Learning Company
1325 NW Flanders Street , Portland, OR 97209, USA

Misha Pavel (pavel@bme.ogi.edu)
Biomedical Engineering
Oregon Health and Science University,
20000 NW Walker Rd, Beaverton, OR 97239, USA
Spry Learning Company
1325 NW Flanders Street , Portland, OR 97209, USA

Katherine Wild (wildk@ohsu.edu)
Neurology
Oregon Health and Science University,
3181 SW Sam Jackson Park Road Portland, OR 97201
Portland OR 97239

James McKanna (zephy@sprylearning.com)
Spry Learning Company
1325 NW Flanders Street
Portland, OR 97209, USA

Payton Bissel (payton@sprylearning.com)
Spry Learning Company
1325 NW Flanders Street
Portland, OR 97209, USA

Devin Williams (devin@sprylearning.com)
Spry Learning Company
1325 NW Flanders Street
Portland, OR 97209, USA

Abstract

Elders are the fastest growing demographic of new computer users, and those over the age of 75 are at risk for medically related cognitive decline and confusion. The early detection of cognitive problems enables earlier treatment that may be much more effective. To address this issue, we have developed a method for embedding cognitive assessment algorithms within computer games that are enjoyable for elders to play on a routine basis. The cognitive assessment algorithms also serve as input to tailored automated hints and help functions for users of various cognitive abilities. In this paper we describe a software architecture and methodology for monitoring cognitive performance using data from a suite of computer games designed to assess multiple dimensions of cognitive performance.

Introduction

Cognitive performance is a key health concern of elders in the United States. In fact, maintaining cognitive health is often the most important factor in being able to age in place. Nearly 50% of all people over the age of 85 are found to have a measurable decline in cognitive function (Callahan, 1995). However, common clinical practice does not offer methods for detecting cognitive decline at an early stage, when therapies may be more effective. Recent research has demonstrated the importance of detecting cognitive decline in an early stage (Chen, 2000). Some cognitive issues have immediately treatable causes, such cognitive disturbances due to medication interactions or short-term medical conditions. However, even with long-term conditions, such as dementia, there are many new therapies that researchers presume would have improved efficacy with earlier detection. In this paper we describe a framework for using unobtrusive computer interaction data to infer cognitive changes on the part of computer users. Frequent assessments allow us to detect relevant changes in various aspects of performance that can be used to adapt the user interface in real time and also provide a mechanism of early detection of cognitive problems.

.

Growing Use of Computers by Elders

Elders are the fastest growing demographic of new computer users in the United States. In a recent survey conducted by the Pew Internet and American Life Project (Pew Internet Project, 2004), they found that 22% of American adults over the age of 65 use the Internet. Interestingly, elders in this group are even more likely than other Internet users to go online and check email each day. In addition, nearly 35% of elders who use a computer have played a game online, comparable to 39%, the average rate of computer game play for other age groups. Given this rapid growth of computer use by users at risk for cognitive problems, as well as the current large use of computers by the advancing wave of baby boomers, we have an important opportunity to collect and interpret naturalistic computer interaction data for diagnostic purposes. In our project on cognitive monitoring with computer interaction data, we have focused on keyboard and mouse data from standard word processing and Web browsing applications, as well as more focused data interpretation of interactions in computer games that we have specifically designed to probe cognitive performance.

Current Methods of Cognitive Assessment

In standard clinical practice, cognitive screenings are usually performed only at advanced age or if there are already patient or family concerns about cognitive dysfunction. These screening tests, such as the Mini-Mental State Exam, the Kokmen Short Test of Mental Status, and the Memory Impairment Screen, can be performed in a physician's office, but are fairly course and not particularly useful for the early detection of problems (Callahan, 1995). More complete neuropsychological batteries can be performed to obtain more sensitive diagnostic information. These normally include measures of short-term and working memory, divided attention, motor speed, planning, and general executive function (Howieson, 2003). Typical tests include:

Verbal Fluency - This test is focused on semantic processing and recall from long term memory. The test procedure requires the participants to recall as many words as possible given a specific semantic category or one or more phonemic constraints.

Word-List Acquisition - This test is focused on learning and recall from short term memory. The test procedure requires the participants to learn and recall a list of words.

Word list Recognition - This is a test of the ability to recognize words previously presented during the Word-List Acquisition test. The participant is asked to discriminate between the words that were presented in the list from distracter words. Together with the Word-List Acquisition test, the recall test can distinguish whether the "forgotten" items were truly lost or the memory trace was just too weak to support reliable recall.

Constructional Praxis - This test is focused on the ability to integrate visual and motor processing. The participants are asked to copy several black-and-white drawings of simple forms such as circle, diamond, etc. In addition, this test is used to assess the participants' visual memory.

Trail-Making Test - This test is focused on complex visual scanning, mental tracking and mental flexibility. The participants are asked to trace a sequence of digits or interposed digits and letters.

Symbol Digit Modalities Test - This test is used to assess the ability to sustain attention and to perform coding task. The participant is given a table associating a simple but novel symbol with each digit and then asked to assign a number to each of a long list of these symbols.

Letter-Number Sequencing - The focus of this test is working memory and focused attention. The participant hears a list of letters and digits, presented in random order. The task is to repeat the presented items, digit first in the numerical order and then letters in the alphabetic order.

Finger tap test - Although this test is focused on the speed of motor control, there is increasing evidence in the literature that this type of test is useful to predict future decline in cognitive abilities. The participant in this test is asked to push a switch as many times as he or she can within a ten second interval. One feature of this test is that the results of the performance are insensitive to educational level and other demographic variables.

Advantages of Frequent Home-Based Monitoring

The conventional tests described above are usually performed by trained psychologists and usually done no more frequently than once per year. One of the hallmarks of cognitive impairment is the increasing variability in performance. Infrequent assessments do not offer a mechanism to pick this up. In fact, the sensitivity of standard cognitive measures is clouded by a need to reference the performance metrics directly to population norms. Many cognitive tests are highly affected by differences in educational level, language abilities, etc.

In our work with monitoring computer interactions to infer cognitive performance, we attempt to incorporate these conventional metrics of verbal fluency, short-term and working memory, planning abilities, and divided attention into computer activities that are enjoyable for elders to play on a routine basis. With this method we are able to make frequent assessments using each elder participant as their own control. Although our computer assessments are less direct and potentially more noisy on an individual trial, we have the benefit of multiple nearly continuous measures to filter and / or average, and in addition, this technique allows us to analyze within subject trends. Comparing an individual's current performance to their own baseline substantially reduces unwanted confounding effects due to education, language abilities, and culture. In addition, we are able to characterize variability in performance over time, which in itself is a powerful indicator of cognitive function.

Describing Elders' Preferences for Computer Activities

In our project on monitoring elders' computer interactions, we first performed a needs assessment to define elders' preferences for computer applications, games, and potential barriers to computer use. We used focus groups and surveys to help us define a set of features for an elder Web portal that we could use as a research environment to collect real-time interaction data. We also defined a set of enjoyable computer games that could be adapted for cognitive monitoring. To select the games for further development, we observed which features were most enjoyable and easily understood by elders and then also did a cognitive task analysis on each of the games to characterize its appropriateness for providing information on one of the cognitive dimensions described in the previous section on standard cognitive tests.

Measuring Cognitive Dimensions within Computer Game Play

We currently monitor all keyboard and mouse interactions, both within game play, and in conventional computer applications. As one measure of motor speed, we monitor typing speed on the computer keyboard. Although we monitor general typing speed in word processing or game applications, our measures of average login speed provide us with the most robust measure of general motor speed. It is less likely to be influenced by learning effects and other confounding factors. Our repeated measures of login typing speed is a useful proxy for the Finger Tap Test described in the previous section. This is a simple test of motor speed that is highly predictive of cognitive decline. Similarly, we use mouse trajectories (speed and smoothness) within repeated and consistent conditions to provide another measure of motor speed.

In addition to motor speed, we also measure word complexity in word processing and game applications. Our complexity measures include average word length and word frequency in the English language (greater rare word usage corresponding to higher cognitive function). In our home monitoring research, we then compare these results to standard tests of verbal fluency. We also use simple computer word games, as shown in Figures 1 and 2, to provide us with additional assessments of language performance. In these games, the user's speed of word discovery and creation of longer and more sophisticated words (against time and difficulty of available letters) we rate as having higher verbal fluency. We concentrate on monitoring relative performance (with respect to the user's baseline) to look for differences. This is likely to be a more sensitive measure that is less influenced by education and language abilities, and more influenced by cognitive changes.

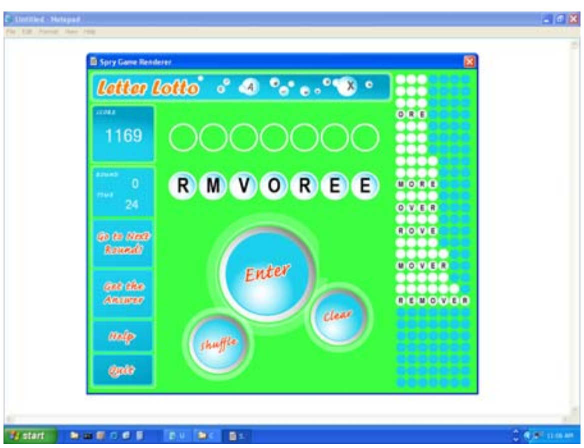

Figure 1: An example of a word game interface (word jumble).

Figure 2: Example of an interface for a word game where users connect adjacent letters to form words.

Standard play in most computer games offer at least an indirect measure of memory. However, to obtain a more direct measure of short term and working memory, we adapted the standard Concentration card game, as shown in Figure 3. Users must remember the location of various cards they select (turn over to view the face of the card) and then match pairs. Game difficulty is adapted based on number of cards and the cognitive difficulty of the matches. These range from simple shape and color matches to cognitively more difficult matches, such as matching a digital clock time with the analogue picture equivalent.

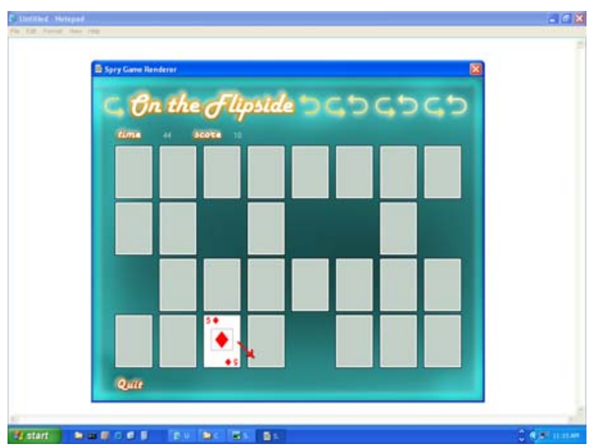

Figure 3: Example of a memory computer game

We have also designed other computer games to specifically test additional dimensions of cognition. Figure 4 shows a shape and color matching game that provides us with measures of planning (inferring the number of steps ahead a user would have to be able to plan in order to be successful). In this game we can also manipulate difficulty and provide added features to test memory and divided attention.

Figure 4: Color and shape matching game that tests planning ability, memory and attention.

Evaluating Cognitive Monitoring with Computer Games in a Home Environment

Most of our experience and testing of computer games for cognitive monitoring has come from our work with an implementation of the popular Solitaire game of FreeCell, as shown in Figure 5. We found that this game was by far the favorite with the elders that we interviewed and in addition, it is a game that requires a significant degree of planning to complete the more difficult layouts of cards.

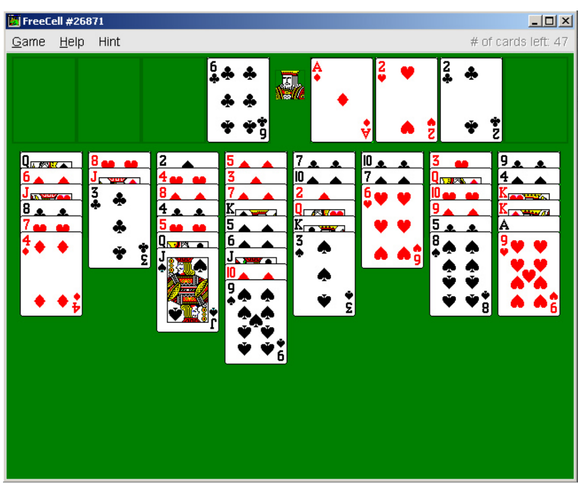

Figure 5: Sample game of FreeCell (Solitaire game requiring significant planning).

In our research version, to measure cognitive performance, we compare user performance to our computer solver. The computer solver provides us with a difficulty metric for any initial and mid-game layout of cards by calculating the minimal number of moves to complete the game from that layout. Figure 6 shows a plot comparing the move-by-move difficulty for a sample game of FreeCell. In this case, the game difficulty starts at 82 moves to optimal solution. The lower line shows the computer solver's direct path to solution, and the upper line shows the subject's moves going toward and away from best solution, with the new difficulty calculated whenever the user changes the board layout. We use the slope of the subjects performance as a measure of efficiency of play. In our early pilot work comparing FreeCell performance of cognitively healthy elders to those with diagnosed mild cognitive impairment, we were able to use the efficiency metric to distinguish the two groups (Jimison, 2004).

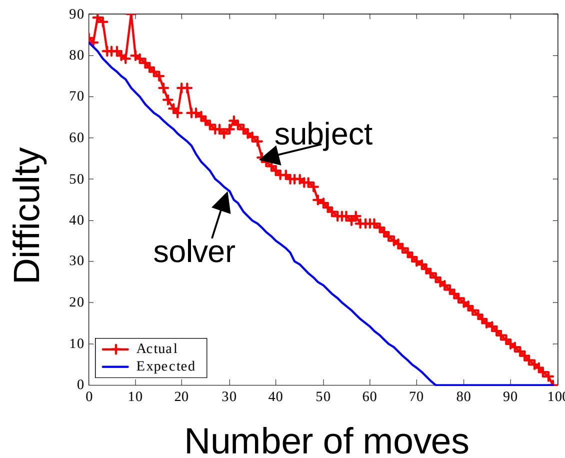

Figure 6: Diagram showing subject performance versus computer solver on the research version of FreeCell.

Table 1 shows the results of our early pilot tests to show the feasibility of monitoring computer interactions in the home. We monitored 12 elders in a local senior residential facility for a period of 3 weeks. Using conventional neuropsychological tests described earlier, we found that 3 of the elderly subjects (mean age 80.2 +/- 8.0) had mild cognitive impairment. Using only data from their FreeCell performance we were able to distinguish cognitively healthy subjects from those with mild cognitive impairment. Interestingly, the variability of the measures over time was in itself a useful feature in classifying cognitive impairment (Jimison, 2004).

Table 1: Classification Performance of FreeCell Metrics

	Ave of Subjects' Ave Efficiency	SD of Subjects' Ave Efficiency	Average of Subjects' SD Efficiency
Normals	0.58	0.12	0.38
MCI	0.27	0.72	0.55

Software Architecture for Cognitive Monitoring

We have developed a rich set of tools for assessing cognitive performance based on the unobtrusive collection of computer interaction data. Our measures are based on keyboard and mouse interactions for both cognitive computer games and conventional applications. The measures include metrics of verbal fluency (word processing and word games), motor speed (login typing, game speed), memory, attention, planning and general executive function. Figure 7 shows our general software architecture for collecting and analyzing the monitoring data.

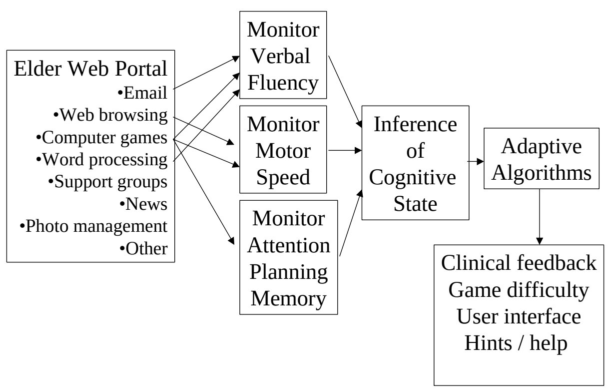

Figure 7: Overview of software architecture for cognitive monitoring.

Real-time analysis of game data takes place on the elder's local machine. This data is used to adapt the difficulty of the ongoing computer game (in cases where appropriate) and also used to adapt the level of difficulty for a user's upcoming games. We attempt to ensure a win rate of between 50% and 80%. Our goal is to keep users engaged by having the activities be challenging but not overly frustrating. In addition, win rates in this region provide us with more sensitive cognitive monitoring data, ensuring that we avoid "ceiling" or "basement" effects sometimes seen on conventional tests that are either too hard or too easy for a patient. We also use real-time analysis and feedback to tailor hints and help messages as part of the user interface. If we realize that a user is having memory problems or divided attention problems, we are then able to immediately adapt our user interface.

Most importantly though, our work on cognitive monitoring is designed to provide clinical feedback to the elder. Based on the elder's preferences, he or she may choose to share this information with caregivers and clinicians.

Conclusion

We have demonstrated proof of concept for a software architecture for real-time unobtrusive monitoring of computer interactions for the purpose of inferring cognitive performance. This approach offers substantial benefits in being able to measure within subject changes over time in a natural setting. Our ability to detect trends in cognitive performance offers the possibility of detecting cognitive decline earlier than conventional methods. We plan to test the effectiveness of this approach in a large prospective long-term trial in elders' residences. Our hope is that this monitoring information may be an inexpensive way of facilitating cognitive health management for elders, helping them maintain their quality of life and independence.

Acknowledgments

This work was supported by the National Institute on Standards and Technology's Advanced Technology Projects, by the Intel Corporation, and by ORCATECH, Oregon's Roybal Center for Aging & Technology with a grant from NIA (Grant NIA P30AG024978).

References

Callahan, C. H. Hendrie, and W. Tierney, "Documentation and evaluation of cognitive impairment in elderly primary care patients," *American College of Physicians*, vol. 122, pp. 422-429, 1995.

Chen, P., Ratcliff, G., Phil, D., Belle, S., Cauley, J., De Kosky, S., and Ganguli, M ."Cognitive tests that best discriminate between presymptomatic AD and those who remain nondemented," *Neurology*, vol. 55, pp. 1847-1853, 2000.

Garber, A. M. 1989. "Long-Term Care, Wealth, and Health of the Disabled Elderly Living in the Community." In David A. Wise (Ed.), *The Economics of Aging*. Chicago: the University of Chicago Press

Geldmacher, D. S., Hills, E. C., "Effect of stimulus number, target-to-distractor ratio and motor speed on visuospatial search quality following traumatic brain injury." Brain Injury 1997; 11(1):59-66

Hofer, S., Sliwinski, M., & Flaherty, B., (2002) Understanding Ageing: Further Commentary on the limitations of cross-sectional designs for aging research, Gerontology. 48, 22-29.

Howieson, D., Bowman, S., Coehlo, D., Lim, S. and Kaye, J. A. "Natural history of cognitive decline in the old old," *Neurology*, vol. 60, pp. 1489-1494, 2003.

Jimison, H.B., Pavel, M., McKanna, J., Pavel, J. Unobtrusive Monitoring of Computer Interactions to Detect Cognitive Status in Elders, *IEEE Transactions on Information Technology in Biomedicine*, Vol. 8, No. 3, September 2004, pp. 248-252.

Kiernan, R., Meuller, J., Langston, J., and Van Dyke, C. "The Neurobehavioral Cognitive Status Examination: A brief but differentiated approach to cognitive assessment," *Annals of Internal Medicine*, vol. 107, pp. 734-739, 1987.

Petersen R.C., Stevens .JC., Ganguli M, Tangalos EG, Cummings JL, DeKosky ST. Practice parameter: Early detection of dementia: Mild cognitive impairment (an evidence-based review): Report of the Quality Standards Subcommittee of the American Academy of Neurology. Neurology 2001 May 8;56(9):1133-42.

Pew Internet Project. Older Americans and the Internet. Pew Internet and American Life Project, www.pewinternet.org . March 2004.

Social Security Online, http://www.ssa.gov/OACT/TR/TR04/ index.html, last update March 23, 2004.

US Census Bureau, www.census.gov/population/ socdemo/age.html, last revised: November 04, 2003.