Smith, Ll., Gobet, F., & Lane, P. C. R. (2007). An investigation into the effect of ageing on expert memory with CHREST. *Proceedings of the United Kingdom Workshop On Computational Intelligence— UKCI07.*

An Investigation into the Effect of Ageing on Expert Memory with CHREST

Richard Ll. Smith
Brunel University
Uxbridge UB8 3PH UK
richard.smith@brunel.ac.uk

Fernand Gobet
Brunel University
Uxbridge UB8 3PH UK
fernand.gobet@brunel.ac.uk

Peter C. R. Lane
University of Hertfordshire
Hatfield AL10 9AB UK
peter.lane@bcs.org.uk

Abstract

CHREST is a cognitive architecture that models human perception, learning, memory, and problem solving, and which has successfully simulated numerous human experimental data on chess. In this paper, we describe an investigation into the effects of ageing on expert memory using CHREST. The results of the simulations are related to the literature on ageing. The study illustrates how Computational Intelligence can be used to understand complex phenomena that are affected by multiple variables dynamically evolving as a function of time and that have direct practical implications for human societies.

1 Introduction

The study of computational or human intelligence requires consideration of how experiences are stored in a memory. Human memory is still poorly understood, and the processes by which experiences are stored, retrieved, and compared are still part of ongoing study in diverse disciplines. In this paper, we describe a computational model of human memory, CHREST, and show how we can verify some basic properties of how experiences are used and stored by considering the effects of ageing on humans. CHREST stores its memories in a tree-like structure, known as a chunking network. Processes of perception and short-term memory (STM) enable CHREST to acquire data and form links between familiar patterns within its memory. By showing how CHREST captures the effects of ageing on humans, we provide new evidence that CHREST's form of memory structure captures the processes of human memory. We argue that the chunking network could form a good basis for storing data in systems for computational intelligence.

As a demonstration of CHREST's properties, we explore an important feature in the development of human intelligence: the process of ageing. In industrialised countries, the effects of ageing on expertise has recently become the focus of much interest; among other reasons, ageing may potentially have serious economic consequences, as it might erode the expertise associated with the workforce.

There is typically—although this varies considerably between individuals—a diminution of abilities such as vision, hearing, and memory as humans age [1]. A similar trend can be observed with intelligence, with the qualification that the ability to solve new problems, known as fluid intelligence, is more affected than the ability to use previously-acquired knowledge, known as crystallised intelligence.

An important question in current ageing research is the extent to which expertise might act as a moderator on the negative effects of ageing [2]. In particular, research into expertise has tried to identify more general compensatory mechanisms that might counterbalance the negative effects of ageing on cognition.

Charness [3-5] has carried out influential research on ageing in chess. In a memory task where positions were briefly presented, he found that, for the same skill level, younger players recalled chess positions better than older players. There was also an interaction between skill level and presentation time, in that the difference between younger and older players became greater as the presentation time was increased from 1 to 4 seconds. It is interesting to note that, in spite of producing worse performance than younger players of the same skill level in memory tasks, older players performed equally well in problem solving tasks where they had to choose the best move, and that they were also faster at choosing their move. However, according to [6], methodological issues limit the interpretation of these results: while the skill level was the same between the two age groups in the experiments carried out by Charness, it is likely that the older players had passed their peak and that their skill level was lower than was the case a few years previously; a consequence of this is that they had the (crystallised) knowledge of stronger players, which makes direct comparison with the younger players somewhat difficult.

So far, two computational models have been used to study the effects of ageing on chess expertise. Charness [7] developed a stochastic model of ageing inspired by Feigenbaum and

Simon's EPAM model [8] in order to investigate the assumption that, with increasing age, players' cognitive mechanisms get slower with the consequence that less exact information gets encoded per unit of time. In the model, expertise is implemented by varying the probability of detecting salient pieces on the chess board and finding information in long-term memory (LTM). To model ageing, Charness proposed the hypothesis that older players are 1.6 times slower than younger players at carrying out these cognitive operations [9]. The model's behaviour was compared to the data referred to above regarding the comparative ability of young and old players to memorise chess positions presented for 1, 2, or 4 seconds. The focus of the analysis was on the interaction between presentation time and skill. In general, the model does a reasonable job of simulating the human data. The interaction between presentation time and skill is explained by the fact that the salient piece detector is not often used with short presentation times, and the main skill difference in performance comes from the time required to find chunks. By contrast, with longer times, the salient piece detector is used frequently, and the ability of young players is boosted by the mechanisms conjointly. While interesting, Charness's simulations suffer from two limitations: first, the model cannot predict errors, and second, the model, being mathematical in nature, does not really carry out the task but merely makes predictions of behaviour as a function of the values of the independent variables.

More recently, Mireles and Charness [10] ran a series of simulations with neural networks in order to further explore the link between knowledge and ageing. They focused on the task of memorising sequences of moves from chess openings.

The neural network type used was a recurrent network with four layers. An input and an output layer were each fully connected to a hidden layer, which was linked to a context layer. The network was trained under supervised learning through backpropagation.

Ageing was modelled as modulations in the noise affecting the neural networks. The results of the simulations indicated that the effect of knowledge was to protect performance against the deteriorating effects of ageing. In line with the literature on ageing, the models simulated the fact that old players show a larger variability in performance than the young players. The simulations were not compared directly to human data; rather, the interest was in accounting for effects discussed in the general literature on ageing.

In this paper, we report further simulations of Charness's memory experiment, using as the subject the CHREST model of human cognition (as described below). We have also carried out a wider investigation into the extent to which CHREST predicts that the effect found by Charness with a specific level (club players) generalises to other levels of skill. Thus, as in Charness and Mireles' neural network experiment, we have also repeated the experiment on different simulated levels of chess skill.

2 The CHREST Cognitive Architecture

CHREST (Chunk Hierarchy and REtrieval STructures) is a cognitive architecture that models human perception, learning, memory, and problem solving [11, 12]. Influenced by the earlier EPAM model [8], it originated from modeling work on chess expertise [13, 14].

The model combines low-level aspects of cognition (e.g., mechanisms monitoring information in short-term memory) with high-level aspects of cognition (e.g., use of strategies). It consists of perception facilities for interacting with the external world, short-term memory stores (in particular, visual and verbal memory stores), a long-term memory store, and associated mechanisms for problem solving. Short-term memory in CHREST contains references to chunks held in long-term memory, which are recognised through the discrimination network from information acquired by the perception system. (See fig. 1 for an overview of the different parts of CHREST.)

Learning is seen as the acquisition of a network of nodes (chunks), which also become connected as a function of the similarity of their contents. Chunks can be seen as clusters of information that can be used as units of perception and meaning (the chunks in the simulations below will be fragments of chess positions). As in EPAM, long-term memory is represented as a discrimination network, which sorts and stores chunks.

Patterns that recur often in the environment make it possible for chunks to evolve into more complex data structures, known as templates [15]. Templates are schema-like structures that have slots allowing values to be encoded rapidly.

Simulations are carried out by allowing the model to acquire knowledge by receiving stimuli representative of the domain under study. For example, during the learning phase of the chess simulations, the program incrementally acquires chunks and templates by scanning a large database of positions taken from master-level games. This makes it possible to create networks of various sizes, and so to simulate the behaviour of players of different skill levels. Taken together with the presence of time and capacity parameters, this

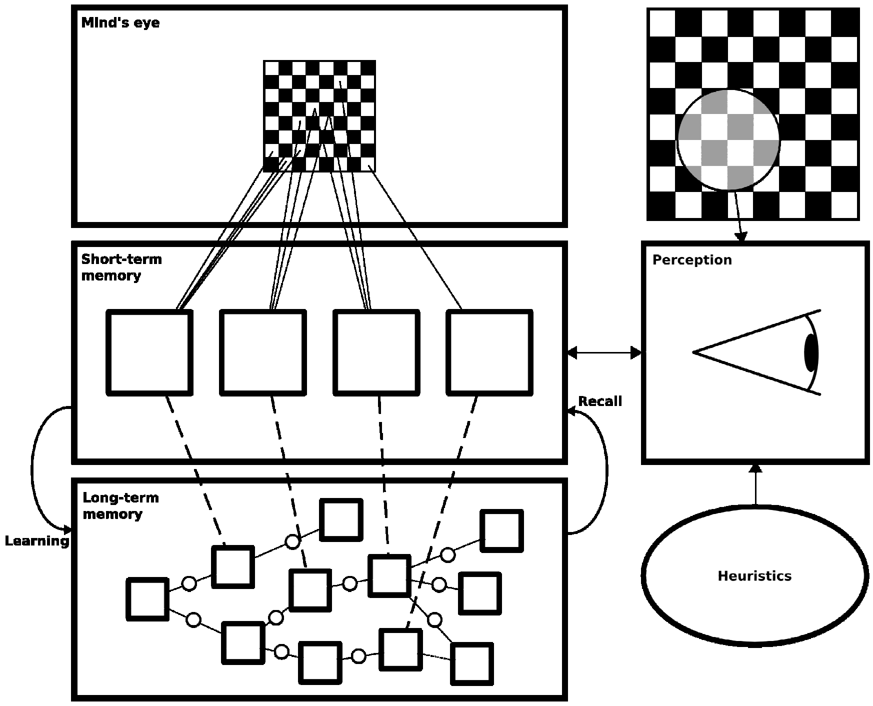

Fig 1. An overview of the components of CHREST. The environment is investigated through a perceptual process which is guided by the contents of STM and heuristics. STM may contain a limited number of references to chunks stored in LTM. LTM consists of a discrimination network of chunks (up to 300,000 in our simulations), connected by tests.

enables CHREST to make unambiguous and quantitative predictions.

A significant aspect of CHREST is the importance it places on the perception process. Rather than passively collecting information from the environment, the process of information gathering is directed by knowledge already learned; this, in turn, affects the knowledge that is captured, resulting in complex emergent behaviour. In the case of chess experiments, perception is equated with eye movements (approximately corresponding to attention), which are directed by chunks held in memory and heuristics.

An important requirement of any model that is claimed to simulate human cognition is that it not assume any abilities exceeding those of a human [16]. Thus, the parameters of the CHREST model are restricted to the human limits as understood by our current comprehension of human psychology. For example, by default the size of visual short-term memory is limited to four items, and the time required for moving the eye (known as a "saccade"), is set to 30 ms.

The majority of the variables in CHREST are time-related, so an internal clock is used to keep track of them. Each time an action is simulated by the system that is understood to take real time, such as mentally moving a piece, the clock is incremented by that period of time, as measured or estimated (note that this time representation is independent of the time taken to simulate the event; the actual processing time may be shorter or longer). Thus, time-restricted problems (such as the experiments we describe here) can be simulated.

The architecture has closely simulated phenomena in several domains, including chess expertise [13, 14, 17]; memory for computer programs [18], the use of multiple representations in physics [19], verbal learning [20], concept formation [21], children's acquisition of vocabulary [22] and children's acquisition of syntactic categories in four different languages [23].

A number of chess simulations are of interest in the context of the effects of age on expert performance. With perception, CHREST has simulated the eye movements during the first 5 seconds of the presentation of a position, as well as the rapid recognition of chunks and templates. With memory, CHREST accounts for the effect of various types of position modification and randomisation, the role of presentation time in memory, the type of errors made and the type of chunks replaced. It can also simulate the detail of how novices acquire chunks and templates. Here we use it to model how ageing affects the recall of chess positions.

3 Simulations of Ageing and Expertise

In previous simulations, CHREST's various parameters were kept constant, except for those parameters that are supposed to differ between experts and novices (e.g., the time required to move a piece in the mind's eye). CHREST can then be seen as implementing the typical configuration of a chess player. The key effects of the simulations were obtained by changes in the input (e.g., recall of typical game positions or meaningless random positions) or by changes in the number of chunks and templates acquired as a way to simulate the level of expertise.

Our goal in this paper is slightly different. We are interested in whether, and how, changes in some CHREST parameters can simulate ageing. Thus, our strategy will be to systematically vary some key parameters in order to establish whether their manipulation can reproduce a key result in the literature on ageing and expertise.

2.1 CHREST Configuration

The CHREST system was set up to model the expertise of chess players. To represent different levels of chess skill, discrimination networks of different sizes were grown (as described above) by allowing the system to learn from a set of 76,420 chess positions until the desired number of chunks were acquired.

Due to the impracticalities of manipulating a large number of different variables, the relevant parameters of CHREST were grouped according to similarity, producing six bundles as described below:

- *Capacity*, the size of STM
- *Learning*, the time taken to update information held in LTM, such as to add information to a chunk or create a new chunk
- *Discrimination*, the rate at which chunks are recognised from LTM
- *STM-Template*, the speed of storage and updating of chunks in STM
- *Eye Movement*, the time taken to "physically" move the eye
- *Mind's Eye*, the delay required to mentally focus on a square

These bundles were each allocated a meta-variable to act as a coefficient to each of the member variables of the set, thus allowing all of the values of a bundle to be modified simultaneously. For example, the parameters that make up the "Discrimination" bundle includes the time required to begin the discrimination process, set to 10 ms, and the time taken per node discriminated, also set to 10 ms. Setting the associated coefficient to 1.5 would adjust the value of both variables to 15 ms.

2.2 Simulation 1

In the experiment, CHREST was presented with a chess position for a fixed length of time. During this period, eye movements and the consequent assimilation of features into memory were modelled. An attempt was then made to reconstruct the position using the contents of CHREST's memory, which was assessed for accuracy (the model also makes predictions about number of chunks used, size of the largest chunks, number of errors by omission, and number of errors by commission, but these results are not reported here due to space limitations). In each case, the experiment was performed with 10 subjects (the same number selected in Charness's experiment), and each subject was assessed on 50 game positions independent of the training data.

We carried out the experiment under a number of conditions to simulate the effects of ageing. Using the system of bundles as described above, the meta-variables were systematically changed for each experiment. A coefficient of 1.0 was used to simulated the putative time parameters of a young chess player, whereas a value of 4.0 represented the slow down due to ageing (with respect to Capacity, 1.0 was used as the control value, and .25 the aged value). The aged value chosen is toward the upper limit of biological plausibility and was chosen to demonstrate the qualitative effect of the variable.

The variables were necessarily restricted to two values each due to a potential combinatorial explosion; the number of possible combinations is given by n^x, where n is the number of values and x is the number of variables. Our 6 meta-variables, with (for example) 7 values each, may be combined in 117,649 experiments; with 2 values this number is reduced to 64.

The experiment was repeated with each possible combination of values with exposure times of 1 second and 5 seconds.

The results of the experiment were collated in a matrix with each dimension corresponding to a variable: 6 meta-variables, presentation time, discrimination network size, and subject, for a total of 9 dimensions. In all, 7,680 experiments (each considering 50 positions) were performed.

Our interest in this particular simulation is the extent to which CHREST can replicate the key interaction between presentation time and age that was found in Charness's experiment. We chose the network of 1,000 chunks to approximate the skill of Charness's participants, who were Class C players (good club players, but 3 standard deviations below Master-level players).

In the following analyses we will investigate the

Table I
Results of the Analysis of Variance for Experiment I, Focussing on the Main Effects of Age

Proxy for Age	MSE	F Value	p value
Capacity	31955	7751.2	< .001
Learning	2.4	0.1	ns
Discrimination	36	0.9	ns
STM-Template	1656	44.5	< .001
Eye movements	202	5.3	<.05
Mind's eye	735	19.5	<.001

Notes: The degrees of freedom are 1 and 1,276 for all cases. MSE = Mean Square Error, ns = Not Significant.

Table II
Results of the Analysis of Variance for Experiment I, Focussing on the Interaction between Presentation Time and Age

Proxy for Age	MSE	F Value	p value
Capacity	11993	2909.1	< .001
Learning	5	0.1	ns
Discrimination	0	0	ns
STM-Template	27	0.7	ns
Eye movements	0.6	0	ns
Mind's eye	461	12.3	< .001

Notes: The degrees of freedom are 1 and 1,276 for all cases. MSE = Mean Square Error, ns = Not Significant.

main effects of age and the interaction between age and presentation time by assuming that age is mainly mediated by the effect of a single bundle of parameters we have describe above. In turn, we investigate each of the six bundles of parameters. Table I presents the results of the analysis of variance, focussing on the main effect of the age proxy, and Table II focuses on the key interaction between age proxy and presentation time. To simulate Charness's data, one needs both a main effect of the proxy of age and a statistically significant interaction between the proxy of age and presentation time, where the difference between the 'old' and 'young' models should be larger with the long presentation time (5 s). Figs. 2 to 7 illustrate how a given bundle and presentation time jointly affected recall performance.

Only two variables showed the required interaction (Capacity and Mind's Eye). Furthermore, as can be seen from Figs. 2 and 7, Capacity shows the correct interaction (the difference between the 'young' and 'old' models is larger with 5 seconds), while Mind's Eye shows the wrong pattern (the difference is actually smaller with 5 seconds). Note that, for all variables, there was a main effect of presentation time (not shown in the Tables).

Of the remaining variables, it was expected *a priori* that the learning bundle would have a negligible effect: minimal learning is expected to occur during a short experiment; however, it is useful to have experimental evidence of this.

To review, the outcome of this experiment is that Capacity was the only bundle of variables that was able to simulate the key results of Charness. In the next simulation, we investigate whether the same pattern holds across a wide range of skill levels, modelled by networks of 100 nodes to 300,000 nodes.

2.3 Simulation 2

The method was the same as for Experiment 1, with the difference that networks of varying sizes were used, and the analysis will focus on the Capacity variable.

As can be seen in Table III, all main effects were statistically significant. As Network Size gets larger, recall performance increases. As Presentation Time increases, so does recall performance. And as Capacity increases, so does recall performance (this main effect is crucial with respect to Charness's results). Similarly, all interactions were statistically significant. The interaction between Network Size and Presentation Time (illustrated in Fig. 8) indicates that the difference between the 1 second and 5 second presentations gets larger with larger networks. The interaction between Network Size and Capacity (see Fig. 9) is due to the fact that the difference between the .25 and 1 capacities increases from 7.6% to 13.2% from the 100-node network to the 10,000-node network, and then stays stable at about 13%. The interaction between Presentation Time and Capacity (crucial in simulating Charness's results) is due to the fact that the result we have reported with 1,000 nodes generalises to all network sizes: as the presentation time increases, so does the difference in recall between the .25 and 1 capacity levels.

Table III
Results of the Analysis of Variance for Experiment II

Source	df	Mean Square	F	Sig.
Corrected Model	23	73251.644	3052.813	.000
Intercept	1	3385083	141075.7	.000
NET	5	129777.7	5408.576	.000
PRESTIME	1	538965.6	22461.767	.000
CAPACITY	1	255898.2	10664.736	.000
NET * PRESTIME	5	30917.898	1288.525	.000
NET * CAPACITY	5	1643.098	68.477	.000
PRESTIME * CAPACITY	1	76938.226	3206.454	.000
NET * PRESTIME * CAPACITY	5	258.452	10.771	.000
Error	7856	23.995		
Total	7680			
Corrected Total	7679			

Dependent Variable: PERCENT

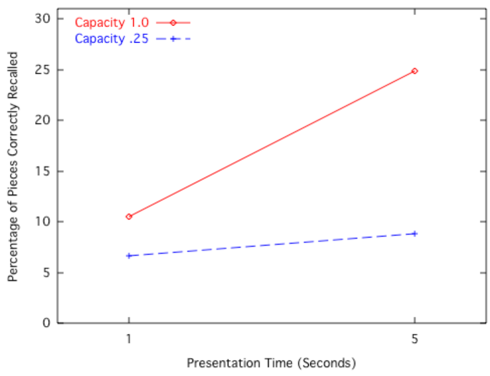

Fig. 2 The effect of Capacity and Presentation Time on recall ability.

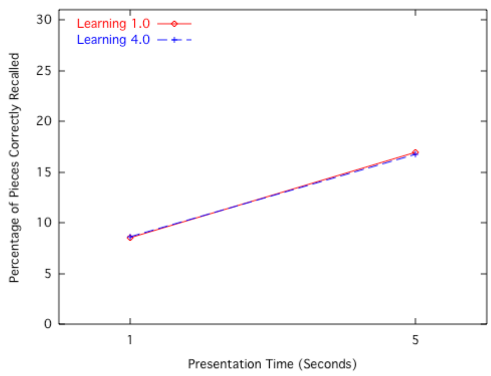

Fig. 3 The effect of Learning and Presentation Time on recall ability.

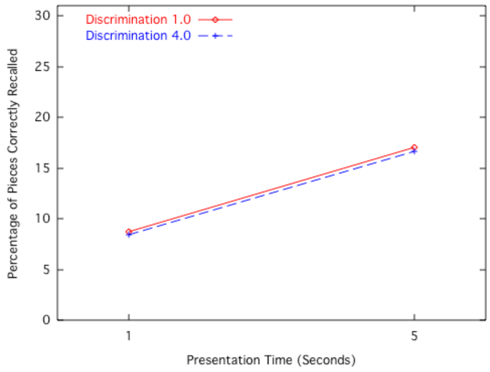

Fig. 4 The effect of Discrimination and Presentation Time on recall ability.

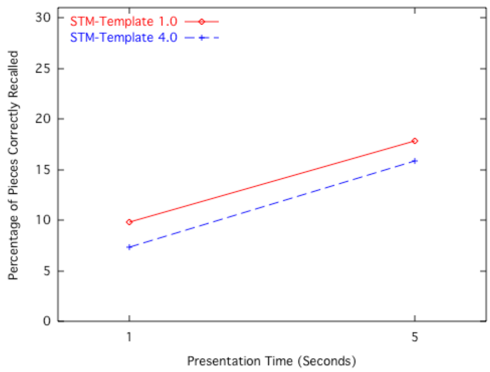

Fig. 5 The effect of STM-Template and Presentation Time on recall ability.

Finally, the 3-way interaction is due to the fact that, at a capacity level of .25, the small networks (below 100,000 nodes) barely take advantage of the longer presentation time, while the larger networks do; by contrast, at a capacity of 1, all networks obtain better recall performance with increased presentation time. Thus, the prediction of the model is that the effects observed by Charness with a sample of Class C players should generalise to all skill levels. We are not aware of data sets where this prediction has been tested.

4 Discussion and Further Work

Within the scope of the processes modelled by the system, and based on Charness's results, we have found evidence that the age-related degradation in the ability to recall chess positions is due primarily to decreased visual short-term memory capacity. We note that although the other variable bundles that were modelled affected the recall ability negatively when artificially aged, STM capacity was the only one that exhibited the required increased difference between age groups with a longer presentation time; we therefore conclude that reduced STM capacity is the variable with the most significant effect.

While Capacity, STM-Template, Eye movements and Mind's eye all showed a main effect of age, Capacity was the only bundle to exhibit the interaction found in Charness's study. Thus our conclusion is slightly different to that obtained by Charness [7], as in his model the slowing down parameter explaining the critical interaction had a general effect and in particular affected perception as well (what he called "salient piece detection"). The finding that several parameters used as a proxy for ageing affected memory in our simulations is certainly in line with what is known about ageing [1]. The critical role of memory capacity for explaining Charness's interaction is also consistent with other studies on ageing [24, 25], which show that age and visual short-term interact in complex ways with presentation time, as well as other variables such as stimulus complexity.

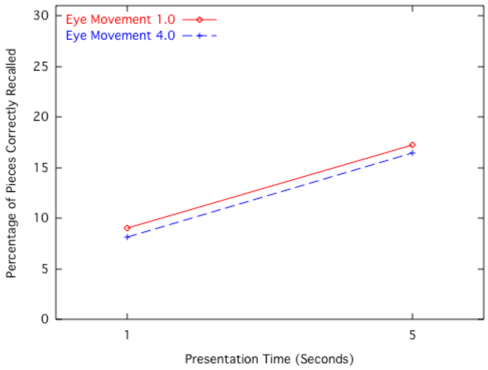

Fig. 6 The effect of Eye Movement and presentation time on recall ability.

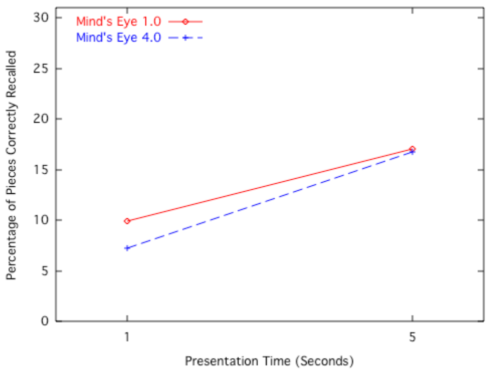

Fig. 7 The effect of Mind's Eye and Presentation Time on recall ability.

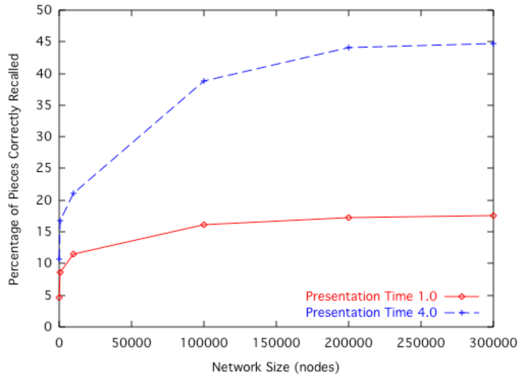

Fig. 8 The effect of Network Size and Presentation Time on recall ability.

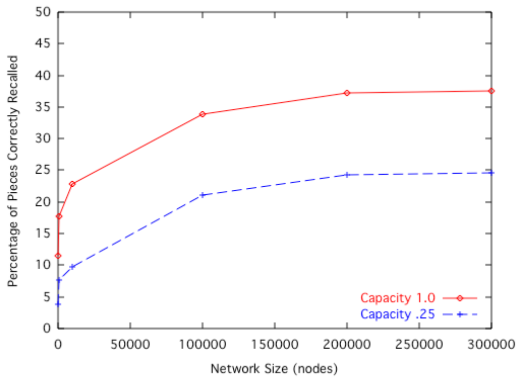

Fig. 9 The effect of Network Size and Capacity on recall ability.

We can further predict that players who show reduced performance in the capacities captured by the bundles STM-Template, Eye Movements, and Mind's Eye will perform less well in the recall test, though the effect of the mind's eye getting updated more slowly should be marginal at longer presentation times. Little or no difference will be seen in those players with slower Learning or Discrimination abilities. (Admittedly, it is not easy to experimentally verify these predictions with respect to some bundles.)

In view of potential future work, we have focused our analysis here on the overall effects of the variable bundles on recall abilities, but of potentially more interest is the extent of the interactions between the variable bundles; for example, to what extent variables can compensate for each other. Deeper analysis of our data, and possibly further experiments, would be needed to investigate this.

Having found indications of the importance of STM capacity, further experiments into the effects of this variable are needed. We note that because the model used to make these predictions is based on known processes and is using psychologically plausible human parameters, we could at this point make some tentative quantitative predictions as to the expected actual recall ability of a player under set conditions with a measured STM capacity. However, we believe we need to investigate the effects of a wider and more finely-grained range of values for the parameter in order to get an accurate estimate.

To conclude, we have shown how the CHREST computational model captures ageing effects through the modification of its capacity constraints. To show this, we used the model to simulate the recall of information acquired through experience. The CHREST model captures, in detail, the important processes of data acquisition, comparison and storage which occur in humans. In other work, CHREST has been shown to handle a range of memory and performance tasks. We suggest that CHREST's internal processes and chunking network provide a firm basis by which computational systems could exhibit some of the memory requirements of intelligent behaviour.

Acknowledgement

This research was funded by the Economic and Social Research Council under grant number RES-000-23-1601.

References

[1] R. Schulz and T. A. Salthouse, *Adult development and aging*, Third ed. Upper Saddle River, NJ: Prentice Hall, 1999.

[2] N. Charness and J. I. D. Campbell, "Acquiring skill at mental calculation in adulthood: A task decomposition.," *Journal of Experimental Psychology: General*, vol. 117, pp. 115-129, 1988.

[3] N. Charness, "Visual short-term memory and aging in chess players," *Journal of Gerontology*, vol. 36, pp. 615-619, 1981.

[4] N. Charness, "Aging and skilled problem solving," *Journal of Experimental Psychology: General*, vol. 110, pp. 21-38, 1981.

[5] N. Charness, "Search in chess: Age and skill differences," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 7, pp. 467-476, 1981.

[6] F. Gobet, A. J. de Voogt, and J. Retschitzki, *Moves in mind: The psychology of board games*. Hove, UK: Psychology Press, 2004.

[7] N. Charness, "The role of theories of cognitive aging: Comment on Salthouse," *Psychology and Aging*, vol. 3, pp. 17-21, 1988.

[8] E. A. Feigenbaum and H. A. Simon, "EPAM-like models of recognition and learning," *Cognitive Science*, vol. 8, pp. 305-336, 1984.

[9] J. Cerella, "Information-Processing Rates in the Elderly," *Psychological Bulletin*, vol. 98, pp. 67-83, 1985.

[10] D. E. Mireles and N. Charness, "Computational explorations of the influence of structured knowledge on age-related cognitive decline," *Psychology and Aging*, vol. 17, pp. 245-259, 2002.

[11] F. Gobet, P. C. R. Lane, S. Croker, P. C. H. Cheng, G. Jones, I. Oliver, and J. M. Pine, "Chunking mechanisms in human learning," *Trends in Cognitive Sciences*, vol. 5, pp. 236-243, 2001.

[12] F. Gobet and P. C. R. Lane, "The CHREST architecture of cognition: Listening to empirical data," in *Visions of mind: Architectures for cognition and affect*, D. Davis, Ed. Hershey, PA: IPS, 2005, pp. 204-224.

[13] A. D. De Groot and F. Gobet, *Perception and memory in chess*. Assen: Van Gorcum, 1996.

[14] F. Gobet and H. A. Simon, "Five seconds or sixty? Presentation time in expert memory," *Cognitive Science*, vol. 24, pp. 651-682, 2000.

[15] F. Gobet and H. A. Simon, "Templates in chess memory: A mechanism for recalling several boards," *Cognitive Psychology*, vol. 31, pp. 1-40, 1996.

[16] H. A. Simon, *The sciences of the artificial*. Cambridge, MA: MIT Press, 1969.

[17] F. Gobet and A. J. Waters, "The role of constraints in expert memory," *Journal of Experimental Psychology: Learning, Memory & Cognition*, vol. 29, pp. 1082-1094, 2003.

[18] F. Gobet and I. Oliver, "A simulation of memory for computer programs," Department of Psychology, ESRC Centre for Research in Development, Instruction and Training, University of Nottingham (UK), Technical report 74, 2002.

[19] P. C. R. Lane, P. C. H. Cheng, and F. Gobet, "CHREST+: Investigating how humans learn to solve problems using diagrams," *AISB Quarterly*, vol. 103, pp. 24-30, 2000.

[20] F. Gobet and P. C. R. Lane, "How do order effects arise in a cognitive model?," in *In order to learn: How ordering effects in machine learning illuminate human learning and vice versa*, F. E. Ritter, J. Nerb, T. O'Shea, and E. Lehtinen, Eds. New York, NY: Oxford University Press, in press.

[21] P. C. R. Lane and F. Gobet, "Discovering predictive variables when evolving cognitive models," in *Pattern Recognition and Data Mining, Pt 1, Proceedings*, vol. 3686, *Lecture Notes in Computer Science*, 2005, pp. 108-117.

[22] G. Jones, F. Gobet, and J. M. Pine, "Linking working memory and long-term memory: A computational model of the learning of new words," *Developmental Science*, in press.

[23] D. Freudenthal, J. Pine, and F. Gobet "Modelling the development of children's use of optional infinitives in English and Dutch using MOSAIC," *Cognitive Science*, vol. 30, pp. 277-310, 2006.

[24] J. K. Adamowicz, "Visual short-term memory and aging," *Journal of Gerontology*, vol. 31, pp. 39-46, 1976.

[25] J. K. Adamowicz and B. R. Hudson, "Visual short-term memory, response delay and age," *Perceptual and Motor Skills*, vol. 46, pp. 267-270, 1978.