The Use of Inferential Information in Remembering Go Positions

Jay Burmeister and Janet Wiles
Departments of Computer Science and Psychology
The University of Queensland, QLD 4072, Australia

jay@cs.uq.edu.au
http://www.psy.uq.edu.au/~jay/

Abstract

Depending on the design of a memory experiment, it may be possible for subjects to make inferences from partly reconstructed responses which enable better performance than would be possible solely by memory. In explaining skilled chess memory performance, de Groot (1965; 1966) and Chase and Simon (1973a; 1973b) diminished the role played by inference, instead concentrating on perceptual and pattern recognition processes. In order to highlight the impact that inference may have on memory performance without the confound of the reliance on perceptual or pattern recognition processes, we tested our subjects episodic memory for sequences of Go moves and their ability to inferentially reconstruct the order of moves leading to a final Go position. Thus, we examined the use of information made available by inference (inferential information) and by episodic memory (episodic information) in remembering Go positions.

We report on two experiments in which the subjects had to indicate the order of play for the stones from the opening of Go games. In the first experiment, cued reconstruction was used whilst in the second experiment, unassisted reconstruction was used. There were two tasks in both experiments; in the episodic task the stones were presented sequentially every two seconds whilst in the inferential task no sequential information was provided. Thus, in the episodic tasks, both episodic and inferential information was available to the subjects whilst in the inferential tasks, inferential information was the main source of information available to the subjects.

We discuss our results which confirm earlier findings regarding expertise by de Groot (1966) and Chase and Simon (1973a; 1973b) and compute an upper limit for the proportion of performance which may be due to the use of inferential information. We conclude that the use of inferential information may have a significant impact on memory performance.

1 Introduction

Two of the main conclusions from de Groot's early work on chess players in the 30s and 40s was that the difference between master chess players and weaker players was in their perceptual abilities and that the basis of chess mastership was largely a matter of memory (de Groot 1965[1]). In his later work, de Groot performed a perception experiment on 5 master and 5 weak chess players which had two tasks: reproduction and blind-guessing (de Groot, 1966). In the reproduction task, the subjects were shown a chess position for 5 seconds and then given as long as they liked to concentrate on the position they had just seen. They were then given the chess pieces from the position and asked to reproduce it on a blank board. After their first attempt, the incorrect pieces were returned to them, and without again seeing the original position, they were asked to make a second attempt at reproducing the position by

[1] English translation of a Dutch manuscript published in 1946 which also reported new research.

placing the returned pieces on the board. This process was repeated until either all the pieces were correctly placed or 12 attempts had been made. In the blind-guessing task, the subjects were given the pieces from an unseen position and asked to place them where they thought they should be. The wrongly placed pieces were given back to them for replacement as in the reproduction task.

The weak players seemed to gain little advantage from the 5 second exposure to the position, performing at around the same level on both tasks. The master players performed as well at the blind-guessing task as the weak players did at the reproduction task. de Groot reached the conclusion that inference played an insignificant role in the masters' memory performance.

Building upon de Groot's work, Chase and Simon investigated chess skill and the chunking of chess information using a series of experiments (Chase and Simon, 1973a; 1973b). In their main experiment, there were two tasks: the memory task was similar to de Groot's reproduction task whilst in the perception task, subjects were asked to reconstruct a chess position which was in full view onto an empty chess board. Their subjects consisted of a chess master, a class A player, and a beginner and their results confirmed de Groot's earlier findings that the amount of information extracted from a chess board is related to chess skill. Their explanation for chess skill was that the labels of chess patterns held in long-term memory are placed in short-term memory by a salient pattern detector (i.e., skilled chess processing takes place at the perceptual level), and that this process also underlies chess memory performance (Chase & Simon, 1973b).

In examining chunk size and memory span, Chase and Simon expected that the number of chunks remembered by the subjects would not exceed the standard capacity of short-term memory and that the master's chunks would contain more pieces than the other subjects. Their results showed that the master's chunks were indeed larger than the other subjects and that the memory span of all subjects was within the standard capacity of short-term memory (7 +/- 2). However, they were surprised to find that the master also remembered more chunks than the other subjects. One of the reasons they advanced to explain this result was that the master (and the other subjects to a much lesser extent) sometimes inferred where the pieces should be from general chess knowledge rather than actually remembering where they should be.

To show how reconstructing a position from partial information is relatively easy, Chase and Simon referred to the pennies-guessing task in which all the pieces on a chess board are replaced with pennies and the subjects must substitute chess pieces for the pennies to 'reconstruct' the position (Chase & Simon, 1973b). Although it is unclear from the text[2] whose data they are citing or what the experimental design for the pennies-guessing experiment was, they claim that the master performed almost perfectly and the class A player performed at over 90%.

The impact of information made available by inference (inferential information) on a subject's memory performance seems to have been minimized by de Groot and by Chase and Simon. de Groot claimed that inference played an insignificant role in the master's memory performance claiming instead that chess skill makes information normally obtained by abstraction or inference available perceptually i.e., abstracted and inferred information is increasingly replaced by information gained perceptually as chess skill increases. Disregarding the master's performance on the pennies-guessing task, Chase and Simon dismissed the role that inference played in explaining the master's memory performance.

We believe that the impact of inferential information on a subject's memory performance has been neglected because of the attention paid to pattern recognition and perceptual ex-

[2]The text appears in a footnote on page 238 of Chase & Simon (1973b).

planations for memory performance. We believe that a complete explanation for memory performance should include both perceptual and inferential processes. In order to demonstrate that inferential information may have a significant impact on memory performance, we designed two experiments which limit the impact of perceptual processes (e.g., pattern recognition) on memory performance. By testing our subjects' episodic memory for sequences of moves in the fuseki (opening) of Go games rather than for static positions (like de Groot and Chase and Simon), we minimized the contribution made by perceptual or pattern recognition processes to their memory performance.

Before proceeding, we should first make clear our conception of inferential and episodic information. Inferential information is new information inferred from existing information (the premises). In our experiments, the premises consist of the subjects general Go knowledge and information gained by feedback during the reconstruction phase. Obviously experienced Go players will have access to more general Go knowledge and will also profit more from the feedback information than beginner players. Episodic information is information related to episodes (or events). Although the simplest form of episodic information is the placement of each individual stone, much more complicated episodes can be hierarchically constructed (e.g., joseki (opening) sequences), particularly by experienced players. In our experiments, episodic information comprised the main source of information available to the subjects in the episodic tasks although they could also make use of the inferential information available to them. Inferential information comprised the main source of information available to the subjects in the inferential tasks since no episodic information was provided.

2 Experiment 1: Cued Reconstruction

To demonstrate that inferential information may have a significant impact on memory performance, we used two memory paradigms: cued reconstruction, and unassisted reconstruction. By using cued reconstruction in the first experiment, the subjects' tasks were relatively easy since they were always selecting from a set of clearly defined options.

The first experiment was loosely based on the pennies-guessing task (Chase & Simon, 1973b). In the episodic task, stones were presented in order every two seconds; in the inferential task, an entire position was presented i.e., all the stones were visible. During the reconstruction phase in both tasks, all the stones were visible to the subjects for selection (similar to the pennies-guessing task).

The main source of information available in the episodic task was the subjects' episodic memory for the sequence and a secondary source was inferential information i.e., the subjects were able to infer from the stones already reconstructed which stone should be selected next. However, in the inferential task, there was no episodic memory cues and thus, the main source of information was inferential.

We hypothesized that the memory performance on both tasks would show an effect for Go skill (confirming the previous findings of both de Groot (1966) and of Chase and Simon (1973a; 1973b) on expertise) and that all subjects would perform better on the episodic task than on the inferential task. We also hypothesized that the level of the subjects' performance on the inferential task would approach a significant proportion of their performance on the episodic task.

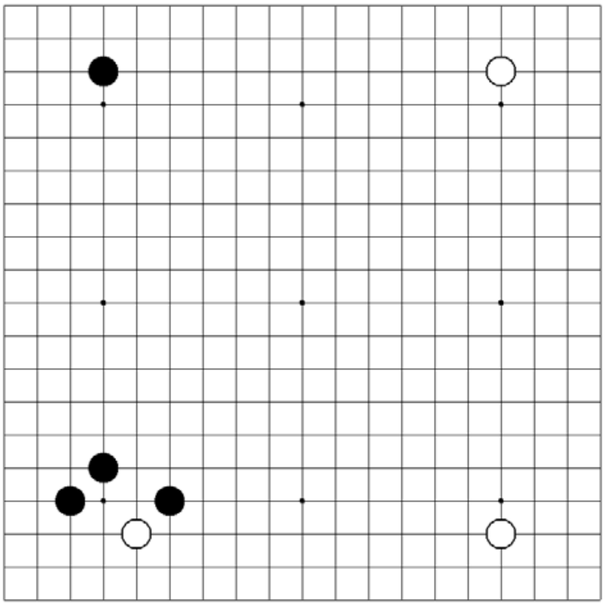

Figure 1: Sequential presentation in the episodic task. Stones were added every 2 seconds: the screen dump shows the board position 12 seconds after the first stone was presented.

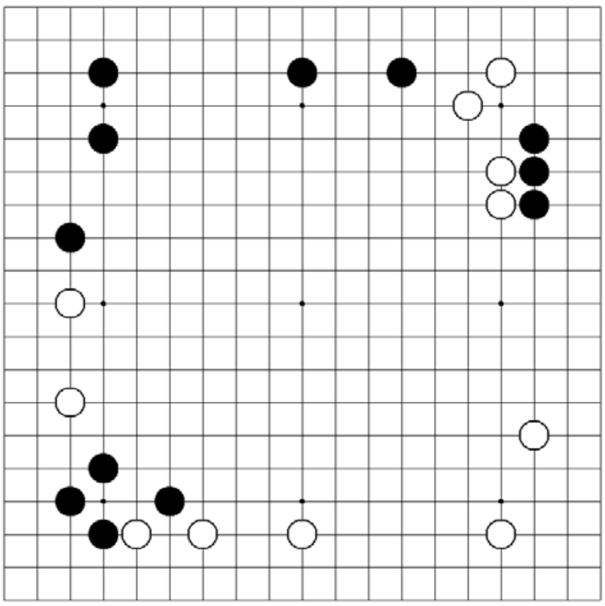

Figure 2: The entire position. At the beginning of the reconstruction phase of the episodic task and at the beginning of the inferential task itself, the entire position was presented.

2.1 Method

Subjects

The subjects were unpaid volunteer Go players who were divided into 2 groups: experienced and beginners. The 4 experienced[3] players ranged between approximately 10 and 1 kyu[4] and

[3]We use the term 'experienced' rather than 'expert' advisedly since our subjects can not be described as masters in the sense used by de Groot or by Chase and Simon.

[4]Ratings are relativistic in nature and these may not exactly correspond to Japanese ratings.

the 4 beginners ranged between approximately 25 and 15 kyu. All but one of the subjects had university level education. The ages of the subjects' were in the range of 25 to 50.

Materials

Seven board positions from the fuseki of Shusaku's[5] games were chosen (Power, 1982) in which there was a relatively even spread of stones around all the corners and edges and in which no stones had been captured. The positions contained between 17 and 35 stones (average 24.6). One board position was used as a practice board; the other six were randomly allocated into two groups of three (groups A and B).

Computer software specially written for the experiments was used to present the board positions to the subjects. Throughout the experiment, the subjects were seated in front of a monochrome X-window display. A mouse was used to select stones during the reconstruction phase.

Procedure

Subjects were asked to complete two tasks using the computer software described above. In the episodic task, the final position was cumulatively built by adding successive stones to the Go board on the computer monitor every 2 seconds in the order in which they had been played in the actual game (see Figure 1). The board grid and stones were cleared for 10 seconds and then the board was redisplayed with all the stones visible i.e., the entire final board position was visible (see Figure 2). Thus, during the reconstruction phase, the subjects were provided with the stones from the position as cues i.e., cued reconstruction was used.

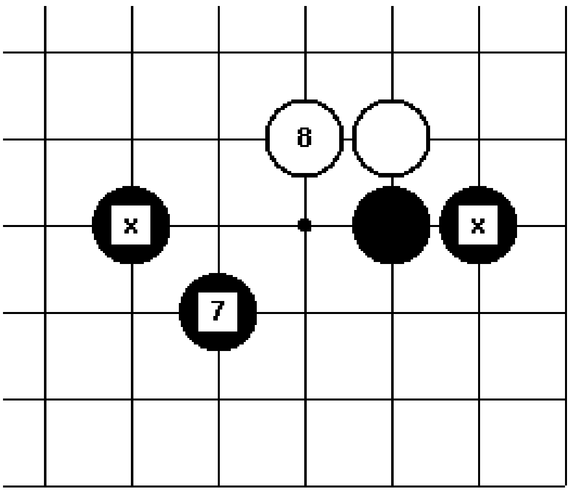

Figure 3: Feedback: wrong stone selected. The screen dump shows a portion of a board position during identification of the 9th move with 2 wrong stones selected (labelled with 'x').

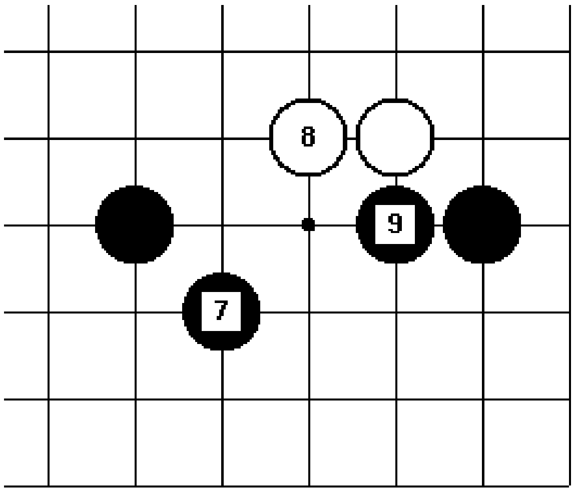

Figure 4: Feedback: correct stone selected. The screen dump shows the 'x' labels cleared when the correct stone is selected and the stone was labelled with the move number (e.g., '9').

In the reconstruction phase, the subjects were asked to indicate the order in which the stones had been played (i.e., added to the Go board on the monitor). When an incorrect stone was selected (i.e., the stone selected was out of sequence or the wrong colour), an 'x' was displayed in the middle of the stone (see Figure 3). When a correct stone was selected,

[5] Shusaku was a famous Japanese player of the mid 19th century.

the number of the move on which it was played was displayed in the middle of the stone and all stones with an 'x' in them were returned to solid colour (see Figure 4). If the subject could not correctly identify the next stone after 10 attempts, a dialogue box would appear and upon pushing the 'OK' button, the computer would identify the correct stone and clear the 'x' from the wrongly selected stones. The subject would then continue the reconstruction process by trying to indicate which stone was the next one played. At the beginning of the inferential task, the final position was presented to the subjects in its entirety. The subjects' task was the same as for the episodic task and the mechanics of stone selection and feedback was also the same.

To enable the subjects to become familiar with the use of the mouse and the feedback system employed by the software, they were presented with the practice board prior to being tested on the episodic task. An explicit practice for the inferential task was not necessary since the use of the software was the same as for the episodic task and the inferential task always followed the episodic task. The presentation of the positions from groups A and B were counterbalanced for subjects within the experienced and the beginner groups i.e., half the subjects in each group received the group A positions in the episodic task and the group B positions in the inferential task and the other half received the group A and B positions in the reverse order.

2.2 Results

The percentages of stones correctly identified on each attempt (i.e., on the first, second, third etc. attempt) were plotted as a cumulative total for the memory and inferential tasks for both the experienced and beginner subjects in Figure 5. Where subjects incorrectly selected stones of the wrong colour, the attempt was not counted i.e., the number of attempts to success was reduced by the number of wrong-colour selections. Since there was only one valid choice for the last two stones, they were not counted in the first attempt total and the percentages were therefore calculated from 2 less than the number of stones in the position.

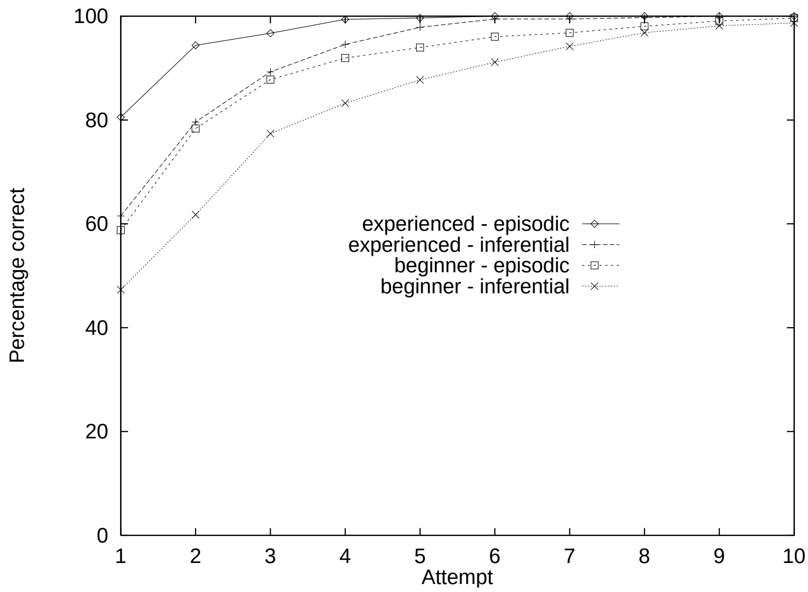

Figure 5: Experiment 1 results. The cumulative percentage of correctly identified stones for each attempt for experienced and beginner subjects on the episodic and inferential tasks.

It can be seen from Figure 5 that the subject's performance is related to Go skill and that performance on the episodic task is better than on the inferential task for all subjects. It can also be seen that the experienced subjects performed better on the inferential task than the beginner subjects did on the episodic task.

2.3 Discussion

It is possible to calculate an upper limit for the proportion of performance in the episodic task which may be due to inference (i.e., the impact of inferential information on memory performance) by calculating the ratio of the performance on the inferential task to the performance on the episodic task. The first attempt performance which may be due to the use of information made available by inference for the experienced subjects was 76.4% (61.5/80.5) and for the beginner subjects was 80.4% (47.3/58.8). These data are of course upper limits since it is not possible to accurately separate the performance in the episodic task into its episodic and inferential components. We do not suggest that the use of inferential information is necessarily as high as the upper limits suggest, rather that in the limit it it may reach those levels.

Since Chase and Simon's description of the results of the pennies-guessing task is somewhat vague as mentioned above, it is not possible to directly compare our results with theirs. However, we believe our results are qualitatively similar to Chase and Simon's with our experienced players performing at 89.3% from attempt 3 and perfectly by attempt 9. The initially low performance of 61.5% on the first attempt may have been due to our subjects not detecting tenuki moves (moves to another part of the board rather than a local response to the opponents last move) on the first attempt but quickly recovering on their second or third attempts.

3 Experiment 2: Unassisted Reconstruction

In the second experiment, we conducted a harder test of our hypothesis that inferential information may have a significant impact on memory performance by using unassisted reconstruction. During the reconstruction phase, the subjects had to recall or infer both sequential and spatial information.

The design of the second experiment was adapted from de Groot's experiment described above (de Groot, 1966). In the episodic task, the stones were presented in order every two seconds; in the inferential task, only the first stone was initially visible. During the reconstruction phase, the subjects had to select the point where they thought the next stone would be played since all the stones were not visible as in the first experiment.

As for the first experiment, the main source of information available in the episodic task was episodic with inferential information being a secondary source of information and in the inferential task, the main source of information was inferential. The tasks were anticipated to be much harder than the tasks in the first experiment and thus the performance to be lower. As for the first experiment, we hypothesized that there would be an effect for Go skill, that all subjects would perform better on the episodic task than on the inferential task, and that the subjects' performance on the inferential task would approach a significant proportion of their performance on the episodic task.

3.1 Method

Subjects

All of the 6 subjects who volunteered to participate had also participated in experiment 1. As in experiment 1, they were divided into 2 groups: experienced (10 - 1 kyu) and beginner (25 - 15 kyu); the ages of the subjects' were in the range of 25 to 50.

Materials

Eight board positions were selected from Shusaku's games (Power, 1982) using the same criteria as for the first experiment. The three positions used in the episodic task contained 25 stones and the three positions used in the inferential task contained 16 stones. Two practice board positions were used; one for each task.

Procedure

Using the same computer software used in experiment 1, subjects were asked to complete an episodic and an inferential task. In the episodic task, the board positions were presented in the same way as they were in the episodic task in the first experiment and the board grid and stones were also cleared for 10 seconds. A blank Go board was then displayed and the subjects were asked to indicate where each stone had been played in the order in which they had been added to the board. Thus, during the reconstruction phase, the subjects were provided with no cues i.e., unassisted reconstruction was used.

The subjects selected a point on the board via the mouse and a stone of the correct colour was placed at that point. The mechanics of stone selection and feedback were the same as for the first experiment including the computer placement of the correct stone after 10 unsuccessful attempts. The subjects practiced the episodic task before the reconstruction phase. In the inferential task, only the initial stone from the position was presented to the subject (see Figure 6). The subjects' task was to infer and/or guess where subsequent stones in the game had been played and they also practiced the inferential task before testing.

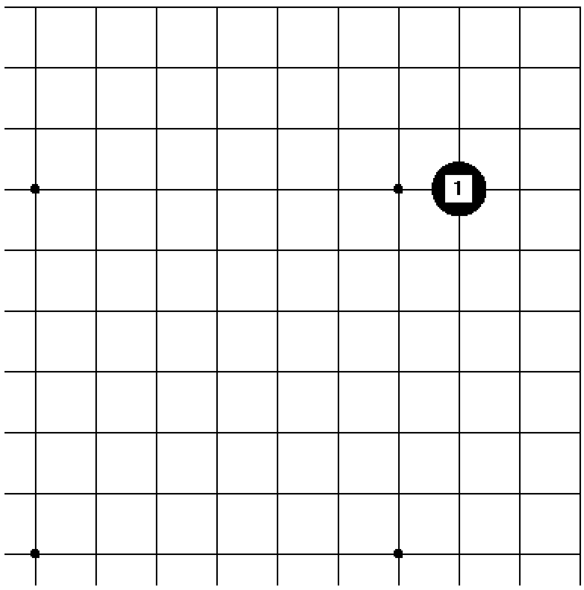

Figure 6: Initial stone presented in inferential task. Screen dump showing a portion of a board with only the first stone being visible with a '1' as a label.

3.2 Results

As for the first experiment, the percentages of stones correctly identified on each attempt were plotted as a cumulative total for the episodic and inferential tasks for both the experienced and beginner subjects in Figure 7. The percentages were calculated from 25 stones in the episodic task and 15 stones in the inferential task.

It can be seen from Figure 7 that, as in the first experiment, the subjects performance is related to Go skill and that performance on the episodic task is better than on the inferential task for all subjects. As in the first experiment, it can also be seen that the experienced subjects performed better on the inferential task than the beginner subjects did on the episodic task (which is a similar finding to de Groot's, as mentioned above).

3.3 Discussion

Once again, it is possible to calculate an upper limit for the proportion of performance in the episodic task which may be due to inference. The first attempt performance which may be due to the use of inferential information for the experienced subjects was 46.9% (23.7/50.5) and for the beginner subjects was 49.0% (9.6/19.6). These results are only indicative and further studies are required to substantiate them.

Performing the same computation on the first attempt data from de Groot's experiment returns 40.8% (37.3/91.4) for the master subjects and 83.7% (37.3/41.2) for the weak subjects. The disparity in absolute terms between our results and de Groot's for the weaker players may be due to two reasons. The first is the relative difference in the size of the search space (i.e., 8x8 as compared to 19x19) which meant that the task for our subjects was much more difficult than for de Groot's subjects. It was for this reason that we presented the first stone in the inferential task: the task would have been more difficult if the first stone also had to be identified. The second reason is that in de Groot's experiment, additional pieces could be added that would, on subsequent trials, provide inferential information to aid the replacement of a previously wrongly placed piece. However, in our experiment, each stone had to be identified in order which meant that less information was available for inferential use by the subjects. These difficulties were perhaps in some way alleviated by, but not completely offset by, the fact that there is only one type of piece in Go as opposed to 6 in chess and that our subjects had far longer to view the stimuli than the 5 seconds given to de Groot's subjects.

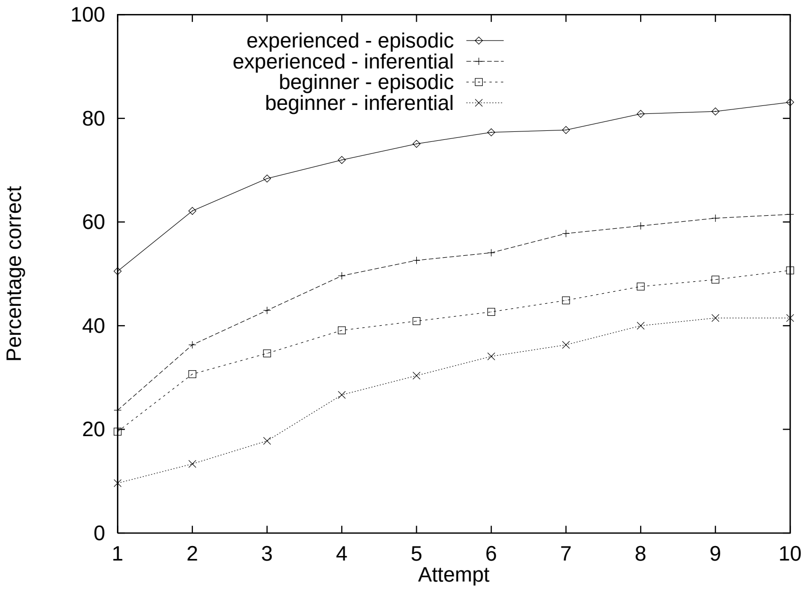

Figure 7: Experiment 2 results. The cumulative percentage of correctly identified stones for each attempt for experienced and beginner subjects on the episodic and inferential tasks.

4 General Discussion and Conclusions

Pattern recognition and perceptual level explanations for chess memory performance were advanced by de Groot and by Chase and Simon in which inference was denied a significant role. In order to limit (but not eliminate) the impact of perceptual and pattern recognition processes on memory performance, we tested episodic memory. Go is a good domain in which to test episodic memory since the final position results from a cumulative addition of stones and by careful selection of stimuli from the fuseki of professional games, it is possible to avoid captures and thus for each element in the stimuli (i.e., each stone), there is exactly one episode (i.e., event) to be remembered.

Our results are qualitatively similar to de Groot's (de Groot, 1965; 1966) and to Chase and Simon's (Chase & Simon, 1973a; 1973b) with regard to expertise i.e., our tasks have shown an effect for Go skill. We attribute this effect to a direct relationship between Go skill and the amount of information (including but not limited to episodic and inferential information) extracted from a Go position. The episodic information available to the subjects depended on their Go skill and ranged from the placement of each stone to hierarchical sequences (e.g., joseki sequences). The inferential information available to the subjects depended on the amount of general Go knowledge they possessed which also determined the amount of information they could extract from the feedback information provided to them.

We have demonstrated a method for computing the upper limit of the contribution of the use of inferential information to memory performance. Although the actual level of the contribution made by the use of inferential information to memory performance can not be determined from our method, we believe that from the upper limits calculated in the cued and unassisted reconstruction experiments, it can be seen that the use of inferential information may make a significant contribution to memory performance.

Our results are presented as a first study exploring the role of inference in memory performance. However, further studies are required to establish the extent of the role of inference in memory, and the context in which inference does and does not play a major role. Currently, in further studies, several other factors are being considered: we are using an alternative selection of games as stimuli (the Shusaku games used in this study are somewhat outdated in style which may have made it difficult for our subjects to make accurate inferences); and the range of subjects is being extended to include expert players.

Acknowledgements

We thank Andrew Hussey for the initial creation of the experimental software and for continuing on in a support role during its development. Thanks also to Yasuki Saito, Atsushi Yoshikawa, and Michael Humphreys for helpful discussions during this study.

References

Chase, W. G. and Simon, H. A. Perception in chess. *Cognitive Psychology*, **4**, pages 55-81, 1973.

Chase, W.G. and Simon, H. A. The mind's eye in chess. In W. G. Chase, editor, *Visual Information Processing*, pages 216-281. Academic Press, New York, 1973.

de Groot, A. *Thought and Choice in Chess*. Mouton, the Hague, 1965.

de Groot, A. Perception and memory versus thought: some old ideas and recent findings. In B. Kleinmuntz, editor, *Problem Solving: Research, Method, and Theory*, pages 19-50. John Wiley, New York, 1966.

Power, J. *Invincible: the Games of Shusaku*. Kiseido Publishing Company, Tokyo, 1982.