The Neurological Scaling of Human Expertise

T. Bossomaier, A. Delaney, J. Crane
Centre for Research in Complex Systems
Charles Sturt University
Bathurst, Australia
Email:tbossomaier@csu.edu.au

F. Gobet
Department of Psychology
University of Liverpool
Liverpool, UK
Email:Fernand.Gobet@brunel.ac.uk

M. Harré
Complex Systems Group
Fac. of Eng. and Info. Technologies
The University of Sydney
Sydney, Australia
Email: mike.harre@gmail.com

***Abstract*—Although chip clock rates seem to have plateaued, the inexorable rise of computing power in accordance with Moore's law continues. We can easily measure the increase in performance using a portfolio of metrics or a Pareto surface across them, including clock rate, memory latency, bus speeds and so on. In this paper, we address two questions. The first of these is what it would mean to scale a human brain, in the way that the primate brain has been getting steadily bigger and more powerful in the lead up to *homo sapiens.* The second is whether, if we could scale the human brain at the same rate as computer power, human algorithms and computational processes would continue to dominate in the domains where humans still reign supreme. To consider these questions we will phrase much of our practical considerations in terms of board games, particularly the games of Go and Chess.**

Keywords*—*neural energy use; scaling; expertise; patterns.

I. Introduction

The last couple of decades have seen a steady erosion of the superiority of human cognition over that of computers. A decade ago, humans were vastly superior at face recognition. Now computer face recognition is a routine technology, albeit with people still better in poor light, or with distorted or corrupted images. CAPTCHAs have to be made increasingly difficult for people to avoid their interpretation by web bots.

The key questions of this paper surround whether the *computational strategies of the human brain would scale if the brain were to become more powerful.* Put another way, were we able to build computers of the power of the human brain, to what extent would we want to adopt the brain's strategies for these powerful computers? To provide some definite focus, we consider expertise in games, particularly Go, where humans still reign supreme.

A key dimension of such questions, particularly with respect to problem solving in the broad sense (finding food, finding a mate, playing the best move in a Go position) is that of pattern recognition and search [1], [2]. Pattern recognition makes it possible to solve problems by quickly recognising important and often frequently occurring pieces of information that elicit good-enough actions. Search is the systematic exploration of the space of possible solutions, possibly those that have been prompted by pattern recognition [3].

The human cognitive architecture excels at pattern recognition. This makes sense from an evolutionary point of view: it is more important to react quickly to a threat with an action, one that might not be optimal, but is good enough, than it is to find the optimal solution. By contrast, the architecture of most computers (Von Neumann serial architecture) offers an excellent support for search.

Although computers are catching up, they frequently are doing so by using different routes to humans, typically involving much more search. Take the case of Deep Blue, the IBM Chess computer which beat world champion Gary Kasparov in 1997. Deep Blue did incorporate some heuristics derived from human play, but its fundamental approach to the game was different – pruning and *search of the game tree,* carried out far, far faster than any human could ever achieve [4]. There is no reason why advanced computation should track the strategies of the human brain. Aircraft were inspired by birds, but they are now much bigger, faster and don't flap their wings.

A successor to Deep Blue attacks a different game, this time with a focus on natural language processing and understanding. IBM's Watson [5] played the TV game show "Jeopardy!" against two of the most successful players of all time and defeated them quite readily. The importance in this outcome lies in the nature of the game Jeopardy! Contestants are provided the answer to a question and they have to provide the question. The answers are notoriously ambiguous and typically require very subtle contextualisation of the clues in order to play successfully. Watson consists of 90 substantial computer servers with 15 terabytes of RAM and 2,880 processor cores. It operates through massively parallel search of documents without any understanding in a human sense.

We show that increases in the power of the brain will enhance pattern recognition, but not necessarily search. Evolution has produced a human brain that is massively parallel at the neural level and has only limited serial capacity at the cognitive level. The latter can be seen in the limited capacity

 ISBN: 978-1-61208-273-8

of short-term memory and the narrow focus of attention. Most people can recall only about 7±2 digits when they are rapidly dictated [6], and people are surprisingly bad at identifying the difference between two alternating pictures that are nearly the same – a phenomenon called change blindness [7]. The bottleneck of attention also affects learning, including a large part of implicit learning, as it is unlikely that unattended stimuli will be learnt.

Will then humans be increasingly good at pattern recognition, without improving significantly their search behaviour? Compared to computers, will they continue to excel at Go but be weak in games where pattern recognition is hard? More generally, the human brain has evolved a specific path in the space of possible cognitive architectures. This means that, if one does not modify the basic architecture of the human brain, there are things that are very difficult for the human brain to do. Search behaviour is one example.

Watson fills a big room and uses 200KW, 10,000 times the brains of the human contestants at a mere 20W. We argue that some aspects of human expertise are extremely energy efficient and some recent brain imaging results reinforce this point. Considerations such as energy consumption as well neural architecture, conductance speed and total numbers of neurons will be considered in terms of what is currently understood regarding human expertise.

II. Complex Games and go

Go, one of the most popular and well studied recreational board games in the world, originated in China sometime before 400 BC. Since that time, it has spread throughout Asia and the rest of the world.

Perhaps the most striking aspect of Go is that it combines three characteristics that are not commonly found in a single game: relatively simple rules; combinatorically large search space for moves; and gameplay that is incredibly engaging for human players. Games with simple rules and large search spaces are almost trivial to generate, but they are very rarely interesting enough for people to want to play them at length, and certainly it is rare for any game to survive for 2,500 years.

Current research interest in Go comes from two particular communities: artificial intelligence (AI) and cognitive psychology. The psychological community's interest in Go has been similar in nature to that of Chess in that these games are a fascinating and contrasting source of data on both the extent and the limits of trained and untrained human abilities. On the other hand, for the AI community Go has replaced Chess as a grand challenge since Deep Blue's success against Kasparov.

A. Comparing Go with Chess

Serious Go players can rapidly acquire some competency within a year but understanding the game to any significant depth can take a lifetime. In technical terms, the search space of moves is considerably more vast than Chess, as is the number of possible games that can be played.

The success of Deep Blue in Chess has stimulated recent attempts to achieve the same result for Go. Deep Blue relied extensively on brute force search of many different lines of play and a relatively weak evaluation function of 8,000 features [4], compared to the tens of thousands of chunks in human expertise [8]. The weak evaluation function was compensated for by the extremely large size of the move tree Deep Blue was able to search.

In the case of Go, there is no accurate, explicit evaluation function for intermediate positions, i.e., positions that are not at the end of the game. Instead, a new and very efficient tree search algorithm, called UCT Monte Carlo, has recently been introduced. It has been very successful on small boards and is the leading contender for algorithms in current Go AI Research [9]). The *pattern recognition strategies of human players do not dominate the current best AIs.*

B. The Nature of Human Game Expertise

A popular explanation for expertise in general and in board games in particular is that players acquire a large number of *chunks,* patterns that become increasingly large with practice and that not only encode perceptual information, such as the location of pieces on a Go board, but also provide information about what kind of actions could be profitably carried out given the presence of a given pattern [8] Pattern recognition works in most board games because features tend to be correlated. For example, the pawn structure in Chess provides a considerable amount of information about the likely placement of pieces.

Simon, Gobet and others [10], [8] studied how skilled practitioners in games such as Chess are able to overcome limitations such as working memory. Subsequently these results were incorporated and extensively discussed in Chess simulation software.

High-levels of expertise implicate a number of brain areas, well illustrated by board games, since a Go or Chess player has to recognise patterns, look-ahead for tactical opportunities, plan ahead at a higher level of abstraction, remember standard tactics and strategies, to mention just some of the cognitive processes involved [8].

Several studies have attempted to identify the location of chunks in the brain. For example, comparing the memory for Chess positions having actually occurred in masters' game and random positions, as well as memory for visual scenes unrelated to Chess, Campitelli et al. [11] showed that the fusiform gyrus and parahippocampal gyrus, both located in the temporal lobe regions were engaged. Bilalić et al. [12] found that Chess experts used temporal and parietal object-recognition areas bilaterally in an identification task with Chess stimuli, but only in the left hemisphere in a control task involving geometrical shapes. Wan et al. [13], using shogi (Japanese Chess), have identified a neural circuit that implements the idea that experts recognize patterns giving access to information on the type of action to take. Pattern recognition would be in the precuneus (part of the parietal lobe), while information on possible action would be stored in a more central and older part of the brain, the caudate nucleus, one of the basal ganglia.

However, some of the patterns of Go expertise are low-level perceptual templates [14], akin to the eyes, nose, mouth features which make up a face. Such patterns are likely to occur lower down, in visual areas or infero-temporal cortex.

 ISBN: 978-1-61208-273-8

1) Patterns and their Frequency of Occurrence in Go: Like Chess, Go throws up a huge range of positional patterns but these are not uniformly likely. The distribution of contextual patterns follows in such games a Zipf-like power law [15]. A useful form of Zipf's law for the k^{th} most frequently occurring element from a set of elements is given by: $freq(k^{\text{th}}$ ranked item$) = c(k+b)^{-\rho}$ where c, b and ρ are constants to be estimated from the data. These distributions are typically associated with 'complex systems' where there are many strongly interacting components giving rise to surprising patterns and dynamics.

Such a power law appears in Go. Liu et al. [15] extracted patterns from 9,447 professional games, featuring over 2 million moves. In terms of the above equation expressed in logs, $\log(freq) = c' - \rho\log(k+b)$ we have $c' = \log(50,000)$, $k \in \{1, 2, \ldots, 20\}$, $\rho = 0.91$ and $b = 0.5$.

Given these frequencies, it is interesting to note some of the previous results of pattern chunking. Simon and collaborators have estimated the number of such patterns needed to be learned to be in the order of 50,000 to 100,000 in order to attain the level of 'master' or above (see Figure 2). Using the equation for the frequency of pattern occurrences and the parameter estimates, we can solve for the frequency of occurrence of patterns of the 50,000 most frequently occurring pattern. All patterns that occur more frequently than the fifty thousandth are presumably more easily learned as they occurred more often. The equation we are left to solve is: $\log(freq) = \log(50,000) - 0.91 \times \log(50,000 + 0.5)$, i.e., $freq = 2.65$. This is significant in that if the frequency were less than 1 then the fifty thousandth pattern is unlikely to be seen during the play of 10,000 games. Note that the $\log(50,000)$ term comes from the parameter estimation whereas the 50,000 in this term: $0.91 \times \log(50,000 + 0.5)$ comes from the need to find the frequency of the fifty thousandth pattern.

This is an important result combining the empirical distribution of pattern frequencies and the number of pattern necessary to be learned in order for a level of expertise to be attained: to become an expert you will have seen the fifty thousandth chunk in your repertoire somewhere between 2 and 3 times. At an estimate of one game per day, this would take over a decade, consistent with the ball park of the 10,000 hours or 10 years required to become an expert that has been suggested in the literature on expertise [16].

If a player has seen significantly less than 10,000 games, then the frequency of patterns seen will drop and a player will not observe a pattern often enough in order to have placed it in memory. For example, if a player has only seen the most frequently occurring pattern in the game 10,000 times (i.e., $k = 1$, all other parameters remain the same), then the pattern that has been observed on average about 2.5 times is the pattern $k = 9,084$, i.e., if 2.5 observations is the frequency required to learn a pattern, then if you reduce the number of games played by a factor of 5 (the $k = 1$ frequency drops from 50,000 to 10,000) then the highest ranked pattern that you can expect to observe 2.5 times is about the nine thousandth most frequently occurring pattern. 9,084 patterns is a significant reduction in the numberof patterns that a player can have acquired given their reduced experience of the game. Considered in the light of Figure 2, it can be seen that such a reduction in the count of acquired patterns, such a player could only expect to achieve the rank of 'expert' rather than 'master'. Figure 1 shows four different instances of fat-tailed curves (straight lines on a log-log plot) for the frequency of contextual patterns in Go [15]. The CT line is the comprehension threshold, so called because it represents the minimum number of observations required before a player is able to remember a given pattern (2.5 observations based on the calculations above). In terms of Gobet et al's work [8], this is the threshold regular patterns of play have been learned sufficiently to be understood as unitary chunks of information. We assume this threshold is fixed, although in practice it is likely influenced by multiple factors. The $\Delta_{i,i+1}$ terms represent the count of extra patterns that have been observed often enough to breach the CT barrier between ranks i and $i + 1$. Note that for linear increases in rank, i.e., skill level s_{i+1} is *reasonably approximated* as a fixed multiple of skill level s_i, $\Delta_{i,i+1}$ is exponential in i due to the logarithmic scale of the x-axis. This approximation is accurate if skill level s_i corresponds to player rank and patterns *above* the comprehension threshold correspond to chunks as in Figure 2.

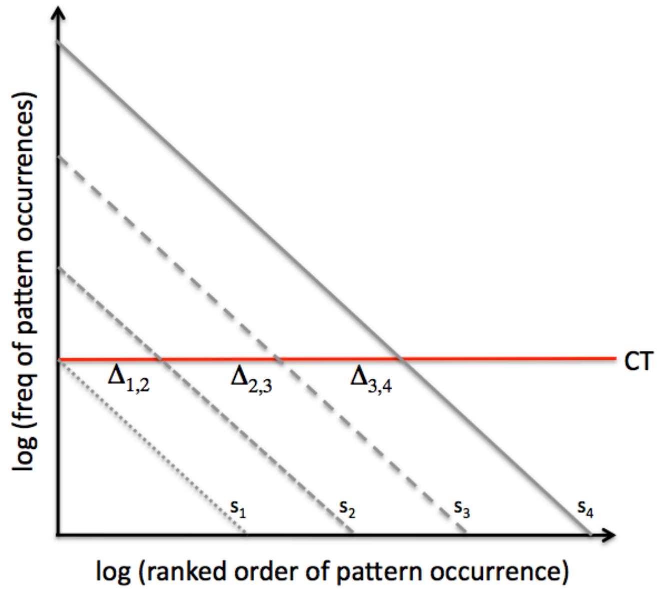

Fig. 1. A stylised log transformed plot of the ranked order of pattern occurrence against the frequency with which each pattern occurs, adapted after [15].

Recent work by Gobet and Lane has shown that the number of chunks increases exponentially with rank (Figure 2). As players increase in skill they master more and more patterns, but each new pattern is likely to be less common than its predecessors. So in a game or tournament, which is a roughly constant number of patterns, then to have a good chance of being aware of a pattern unseen by the opponent, requires an exponential increase in known patterns.

Humans have essentially an infinite capacity for remembering patterns or chunks of data [17]. The decreasing frequency with which rarer patterns are observed puts an upper bound on how much can be stored in practice. To see this, note that the number of patterns that needs to be observed frequently enough to be turned into chunks (i.e., above the comprehension threshold in Figure 1) increases exponentially with increasing

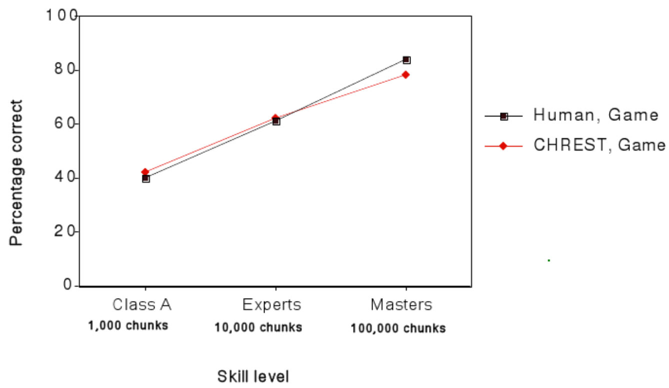

Fig. 2. Graph showing the number of chunks needed with expertise and corresponding model predictions. (courtesy of Gobet and Lane)

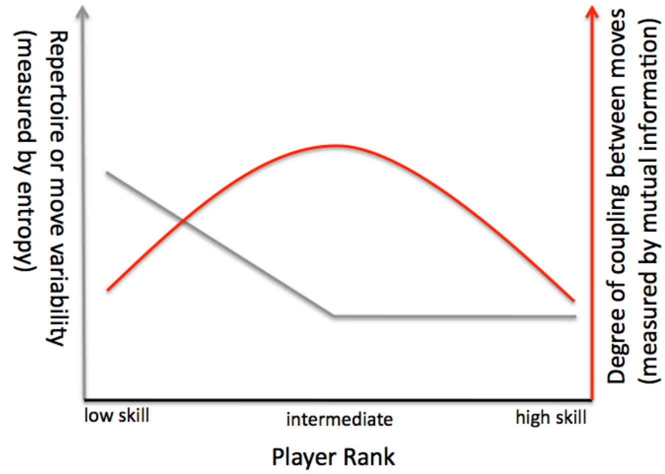

Fig. 3. A stylised plot of the two principal results in [18]. The entropy linearly decreases with rank indicating that certain moves are preferentially chosen more often as skill increases (grey curve). The red curve shows the mutual information between position and move as a function of skill

rank, i.e., the $\Delta_{i,i+1}$ terms in Figure 1 increase *exponentially*. Thus, players need to improve in some other aspect of play and we argue that this improvement occurs in refining the relationships between the patterns learned rather than learning more patterns.

But technology has led to players being able to gain experience faster. In 2009, Joe Cada set the record for the youngest ever Poker World Series Champion at just 21. It is now possible to play many more hands in a given time online than at the table. In Chess, the career peak was around 35 years of age. Recently, this peak went down and it is more like 25 years of age, although Anand, the current world champion, is 40. Thus, the limit of expertise might be set by age. In the absence of dementia or trauma, is it brain size or the slow loss of neurons with age which is limiting?

2) Information Content in Human Decisions: As skill levels increase in Go, the entropy, a measure of the diversity of strategies used, *decreases* to a point where it plateaus. This is where the players have been able to acquire as much information as is extractable from the local configuration of the board, (grey curve in Figure 3).

The difference between strong players and weak players lies in the degree to which the *local* information on the board interacts with the *global* information, i.e., expertise past a certain point is mediated not by a repertoire of patterns, but by the relationships (correlations) between those patterns (see figure 3). We conjecture that this implies *simultaneous grasp* of these global properties, or in psychology terminology, *holding them in working memory* (§II-B3).

Figure 3 also shows a phase transition in expertise, measured by a peak in mutual information [18]. Here, there is some reorganisation of patterns and relationships, necessary before advancing in expertise.

3) Working Memory in Games: At the time of writing (late 2012), computer Go programs have reached professional dan levels on 19x19 boards [19]. However as the handicaps (in terms of speed and board positioning) that favour the computer decreases, human players rapidly gain dominance. In other words, humans have a much better grasp of the interaction between local and global positional elements. In most positions, a maximum of around five live masses can exist on the Go equivalent of 19x19 board. Given that there may be more candidates for ultimate survival, this number is about the order of working memory.

Consistent with a possible relationship with working memory is that the board size for almost all (human) competition play evolved to 19x19 and locked in at this figure. Now, with any game, there is always a lock-in factor for tournaments and rankings to remain coherent. Nevertheless, moving to a 21x21 board would create a greater number of global chunks to be manipulated, making an increase in working memory necessary. On the other hand, tree search scales exponentially with board size.

III. Scaling the Brain's Performance

Despite its impressive performance, the human brain is far from perfect. Gary Marcus in his book [20] catalogues numerous ways in which the brain is biased, gets confused and is prone to all sorts of extraneous influences. He argues that the brain is non-optimal because evolution is a tinkerer, only able to use what is at hand, even though better designs are possible. Thus, some improvements in our cognitive abilities are beyond the space of possible human brain architectures.

An important aspect of expertise in board games is that pattern recognition and search are interleaved [8], [21]. When exploring the space of possible moves, players generate a search tree, albeit much smaller than computers. When trying to find a move in a given position, possibly after having already generated several moves, players tend to use pattern recognition to generate moves automatically [8], [21]. When evaluating a position at the end of branch in their search tree, players also tend to use pattern recognition, rather than a systematic and conscious combination of features of the position on the board. A consequence of this interleaving between pattern recognition and search is that it is likely that the effect of increasing neural efficiency, which is likely to benefit pattern recognition, will be mitigated because of the relative slow speed with which search can be carried out with a human brain.

 ISBN: 978-1-61208-273-8

A. Evolution of Brain Architecture

Brain architecture has been relatively constant since the appearance of mammals, but the neocortex has got bigger. Obviously there might be new transitions in brain evolution, which will create entirely new capabilities, but for the purpose of this article, the primary concern is further increases along alreeady established directions, such as that of size (§III-A1).

The one aspect of architecture that could change is the overarching connectivity. The brain has small world network characteristics and this seems to be important on energetic grounds [22]. But increasing the number of connections per neuron will have an energy penalty, but could increase working memory (§III-B). Could higher levels of connectivity make it any better? But a big new wave is building of sophisticated computation at the sub-neuron level, computation on the dendritic tree (§III-C). The potential here is enormous, but it is intricately tied up with synaptic noise.

Where the human brain is orders of magnitude ahead of its silicon counterpart in terms of energy use, some of the possible wetware enhancements, such as increase in connections per neuron, come at an energy cost (§III-D).

1) The Neocortex and its Recent Evolutionary Enlargement: There are a great many different explanations as to why the human neocortex is as relatively large as it is, but perhaps the most influential has been the demands of living in large, complex groups [23]. Such drivers would also have contributed to good play in complex games.

The relative size of major brain components scales with total brain size across a large number of species [24]. The relative size of the neocortex to whole brain volume scales with the size of a species' typical social group size. It is the neocortex size, which is of interest here, since overall brain size is strongly correlated with body size.

2) Impulse Conduction Speed: If we discount the need to respond in real-time, then conduction speed and age are to first order reciprocally related. However, there could be other considerations. Maintaining synchronised activity across the brain has been advocated for solving the binding problem in perception through to the emergence of consciousness itself. Thus, increasing conduction speed within a given brain, may allow a greater level of synchonisation, since impulses may then arrive before small synaptic activities have decayed away.

B. Connections per Neuron

Increasing the number of dendrites, synapses and links to other neurons offers several advantages. It will likely to increase the, already huge, number of patterns that the brain can discriminate [17]. The number of connections also constrains *working memory*. Roudi et al. [25], using a detailed biologically realistic neural network simulation, show that working memory depends strongly on the average number of connections made per neuron and, at most weakly, on the number of brains cells. The number of connections imposes an energy penalty, however (§III-D). The energy penalty arises because the Excitatory Post Synaptic Potentials (EPSPs), where the spike is converted to activity in the neurons to which it is connected, require slightly more energy than spike generation in human cortex, so increasing connectivity will increase energy requirements [26], [27].

C. Computation on the Dendritic Tree

Over two decades of research in artificial neural networks and a great deal of computational neuroscience has assumed that the inputs to a neuron from other neurons arrive at a single point (effectively the soma). The dendritic tree has often been considered as an anatomical detail, of little computational importance.

But no longer are dendrites viewed as passive players in neural networks. Both the passive and active properties of dendrites endow them with a range of computational abilities [28], such as OR and AND-NOT logic operations (passive), and powerful mechanisms for temporal integration and co-incidence detection (active) [29]. Indeed, it has now been suggested that individual dendritic branchlets should be considered computational units in their own right.

Enhancements in dendritic processing might arise through increased complexity in the morphology of the dendritic tree; more complex patterning of synaptic inputs; alterations in the amount and pattern of expression of active conductances; alterations in the biochemical signalling within dendrites [30]; and increases in hippocampal dependent learning and memory in mice [31].

The computational potential is huge and so far not clearly understood, or at least, its real-world significance has not yet been demonstrated. Furthermore, using the models of neuronal energy use discussed in section III-D, since the dendritic computation is neither part of the axonal or EPSP costs of the cell, it would seem to be exceptionally efficient compared to computation at the network level.

D. Energy Issues

Laughlin and Ruderman found that transferring one bit of information at a synapse needs about 10^4 ATP molecules, the energy transfer mechanism used throughout the animal kingdom [22]. So the brain is operating around 10^5 times above the absolute theoretical limit. Current computers, such as Watson, are at least 10^4 worse.

A later study by Lennie [27] revealed that not only is the human brain relatively efficient compared to computers, its rate of glucose metabolism, the brain's only energy source, is three times lower than in rat and 1.5 times lower than in monkey. He concludes that far fewer neurons are active in human cortex. This would seem to fit the pattern model of expertise rather well. The human brain is much more diverse than rat, storing, and being able to do, many more things. Since these are not all happening at the same time, its *average* energy use is lower.

The wave of connectionist thinking, which began in the mid 80s, emphasised distributed representations. Recent evidence suggests, however, that individual neurons may be extremely specific, such as having a response just to the *concept* of Bill Clinton, regardless of whether this is a picture, his voice, or some ideas or events with which he is strongly associated [32].Clearly such representations engender very low average cortical activity. Thus, if increasing expertise involves laying down more and more increasingly rare patterns, then energy costs may not go up. Where the human brain might seem to score would be in retrieval of a rare, little used pattern with very little additional latency over something used everyday.

Recalling that the biggest consumer of the energy in the neural activity in the brain is in the delivery of the spikes (III-B), the EPSPs. The more synapses a neuron makes (i.e., the higher the number of connections), the greater the energy cost. Thus working memory, which depends on number of connections according to simulations carried out by Roudi et al [25] will demand an energy increase approximately linear in the number of connections.

IV. Conclusions

Human expertise relies strongly on pattern recognition and the building of very large pattern libraries over time. Since patterns themselves do not occur with equal frequency, and often follow a power law in probability of occurrence, this strategy has excellent scaling properties for energy consumption.

The use of the patterns requires some sort of addressing mechanism and this probably occurs through working memory. The best simulations to date suggest that working memory scales linearly with the average number of connections made by a neuron, and, in turn, energy consumption is approximately linear in number of synapses. Thus, although we do not at present know how to calculate the address space (i.e., total possible number of patterns one can store/access), it seems likely that the energy penalty will be logarithmic in number of patterns.

Thus, our view would be that the pattern recognition strategies of human expertise would scale exceptionally well for feasible enhancements in neural architecture. However, it would be significantly harder to improve search. We are left to conclude that neural enhancements to the brain would likely enhance that which it is already good at.

References

[1] H. Berliner, "Search vs. knowledge: An analysis from the domain of games," in *Artificial and Human Intelligence*, A. Elithorn and R. Banerji, Eds. New York, NY: Elsevier, 1984, pp. 105–117.

[2] A. Newell, *Unified theories of cognition*. Cambridge, MA: Harvard University Press, 1990.

[3] F. Gobet and H. Simon, "Pattern recognition makes search possible: Comments on holding (1992)," Psychological Research, vol. 61, 1998, pp. 204–208.

[4] M. Campbell, A. Hoane, and F. Hsu, "Deep blue," Artificial Intelligence, vol. 134, no. 1-2, 2002, pp. 57–83.

[5] D. Ferrucci, "Build Watson: an overview of Deep QA for the Jeopardy! challenge," in *Proceedings of the 19th international conference on Parallel architectures and compilation techniques*. ACM, 2010, pp. 1–2.

[6] G. Miller, "The magical number 7, plus or minus 2: Some limits on our capacity for processing information," Psychological Review, vol. 61, 1956, pp. 81–97.

[7] R. Rensink, K. O'Regan, and J. Clark, "To see or not to see: The need for attention to perceive changes in scenes," Psychological Science, vol. 8, 1997, pp. 368–373.

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

[9] S. Gelly and D. Silver, "Achieving master level play in 9× 9 computer go," in *Proceedings of AAAI*, 2008, pp. 1537–1540.

[10] H. Simon and W. Chase, "Skill in chess," American Scientist, vol. 61, 1973, pp. 393–403.

[11] G. Campitelli, F. Gobet, K. Head, M. Buckley, and A. Parker, "Brain localisation of memory chunks in chessplayers," International journal of Neuroscience, vol. 117, 2007, pp. 1641–1659.

[12] M. Bilalić, A. Kiesel, C. Pohl, M. Erb, and W. Grodd, "It takes two: Skilled recognition of objects engages lateral areas in both hemispheres," PLoS ONE, vol. 6, 2011, p. e16202.

[13] X. H. Wan, H. Nakatani, K. Ueno, T. Asamizuya, K. Cheng, and K. Tanaka, "The neural basis of intuitive best next-move generation in board game experts," Science, vol. 331, no. 6015, 2011, pp. 341–346.

[14] M. Harré, T. Bossomaier, and A. Snyder, "The perceptual cues that reshape expert reasoning," Nature Scientific Reports, vol. 2, no. 502, 2012.

[15] Z. Liu, Q. Dou, and B. Lu, "Frequency Distribution of Contextual Patterns in the Game of Go," Computers and Games, 2008, pp. 125–134.

[16] K. Ericsson and A. Lehmann, "Expert and exceptional performance: Evidence of maximal adaptation to task constraints," Annual Review of Psychology, vol. 47, 1996, pp. 273–305.

[17] T. Brady, T. Konkle, G. Alvarez, and A. Oliva, "Visual long-term memory has a massive storage capacity for object details," Proceedings of the National Academy of Sciences, vol. 105, no. 38, 2008, pp. 14 325–14 329.

[18] M. Harré, T. Bossomaier, A. Gillett, and A. Snyder, "The aggregate complexity of decisions in the game of Go," European Physical Journal B, vol. 80, 2011, pp. 555–563.

[19] D. Ormerod, "Zen computer Go program beats Takemiya Masaki with just 4 stones!" Accessed March 31,2013, 2012. [Online]. Available: http://gogameguru.com/zen-computer-go-program-beats-takemiya-masaki-4-stones

[20] G. Marcus, *Kluge: The haphazard construction of the human mind*. Houghton Mifflin Harcourt, 2008.

[21] F. Gobet, "A pattern-recognition theory of search in expert problem solving," Thinking and Reasoning, vol. 3, 1997, pp. 291–313.

[22] S. Laughlin, R. d. Ruyter van Steveninck, and J. Anderson, "The metabolic cost of neural computation," Nature Neuroscience, vol. 1, no. 1, 1998, pp. 36–41.

[23] R. Dunbar and S. Shultz, "Evolution in the social brain," Science, vol. 317, no. 5843, 2007, pp. 1344–1347.

[24] K. Zhang and T. Sejnowski, "A universal scaling law between gray matter and white matter of cerebral cortex," Proceedings of the National Academy of Sciences of the United States of America, vol. 97, no. 10, 2000, pp. 5621–5626.

[25] Y. Roudi and P. Latham, "A balanced memory network," PLoS Computational Biology, vol. 3, no. 9, 2007, p. e141.

[26] D. Attwell and S. B. Laughlin, "An energy budget for signaling in the grey matter of the brain." J. Cereb. Blood Flow Metab., vol. 21, 2001, pp. 1133–1145.

[27] P. Lennie, "The cost of cortical computation," Current Biol., vol. 13, no. 493–497, 2003.

[28] K. Sidiropoulou, E. Pissadaki, and P. Poirazi, "Inside the brain of a neuron," EMBO reports, vol. 7, no. 9, 2006, pp. 886–892.

[29] M. London and M. Häusser, "Dendritic computation," Neuroscience, vol. 28, no. 1, 2005, p. 503.

[30] T. Branco and M. Häusser, "The single dendritic branch as a fundamental functional unit in the nervous system," Current Opinion in Neurobiology, vol. 20, 2010, pp. 494–502.

[31] M. Nolan, G. Malleret, J. Dudman, D. Buhl, B. Santoro, E. Gibbs, S. Vronskaya, G. Buzsáki, S. Siegelbaum, E. Kandel *et al.*, "A Behavioral Role for Dendritic Integration: HCN1 Channels Constrain Spatial Memory and Plasticity at Inputs to Distal Dendrites of CA1 Pyramidal Neurons," Cell, vol. 119, no. 5, 2004, pp. 719–732.

[32] R. Quiroga, L. Reddy, G. Kreiman, C. Koch, and L. Fried, "Invariant visual representation by single neurons in the human brain," Nature, vol. 435, 2005, pp. 1102–1107.