New Ideas in Psychology 31 (2013) 108–121

Contents lists available at SciVerse ScienceDirect

New Ideas in Psychology

journal homepage: www.elsevier.com/locate/newideapsych

What is the nature of the mind's pattern-recognition process?☆

Alexandre Linhares*, Daniel M. Chada

Getulio Vargas Foundation/EBAPE, Praia de Botafogo 190 office #509, Rio de Janeiro 22250-900, Brazil

Keywords:
Cognitive models
Perception
Memory
Cognitive psychology
Experimental psychology
Context

A B S T R A C T

If we look at the human mind as a pattern-recognition device, what is the nature of its pattern-recognizing? And how does it differ from the majority of pattern-recognition methods we have collectively devised over the decades? These broad philosophical questions emerge from the studies of chess thought, and we propose that a major task of the mind is to engage in "experience recognition" (Linhares & Freitas, 2010). One of the basic tenets of that proposal is that pattern recognition, in cognitive science and related disciplines, does not accurately reflect human psychology. As an example, the well-known article by Chase and Simon, "perception in chess", and the benchmark cognitive computational models of chess, by Gobet et al. were criticized. Lane and Gobet (2011) provide serious skepticism concerning some of those arguments, and here we take the opportunity to respond and expand the theoretical constructs of "experience recognition". We postulate that the mind's pattern-recognizing process holds the following properties: it is a highly path-dependent process; it prioritizes internal encodings; it is a self-organizing process in constant change; and it constructs its future information-processing pathways by continuously recognizing the possibilities that lie within the adjacent possible.

> "Thus intuition (or recognition [of patterns]) provides a ready explanation of some of the apparently extraordinary memory feats of which experts are capable in their domain of expertise" (Simon, 1986, p. 243)
>
> "It has been proposed that intuition may be largely explained by pattern recognition (Simon, 1986). This is the route followed by our models." (de Groot & Gobet, 1996, p. 247)

1. Introduction: what is the nature of the mind's pattern-recognition process?

If we look at the human mind as a pattern-recognition device, what is the nature of its pattern-recognizing? And how does it differ from the majority of pattern-recognition methods we have collectively devised over the decades? While these are broad philosophical questions, we will start from studies of chess thought, and propose that a major task of the mind is to engage in what we term "experience recognition" (Linhares & Freitas, 2010).

Linhares and Freitas (2010) argued that cognitive scientists place too much emphasis on "pattern recognition"—and scarce emphasis on "experience recognition". The study of (static) pattern recognition generally holds a database of known patterns, and a system (or theoretical model), given a *new* pattern, is faced with the task of classifying it against the database store. While the field of pattern recognition has

☆ This work has been supported by grant E-26/110.790/2009 from the Fundação de Apoio à Pesquisa do Estado do Rio de Janeiro (FAPERJ-www.faperj.br), grants 470341/2009-2 and 301207/2009-7, grants E-26/110.540/2012 and E-26/111.846/2011 from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq-www.cnpq.br) of the Brazilian Ministry of Science & Technology, the Programa de Apoio à Pesquisa (PROPESQUISA) of the Getulio Vargas Foundation (www.fgv.br).

* Corresponding author.
E-mail address: linhares@clubofrome.org.br (A. Linhares).

0732-118X/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.newideapsych.2012.08.001

brought deeply impactful contributions toward solving some central problems of modern science, to the goal of modeling human cognition we argue in favor of a similar, but in many ways distinct approach. We believe that pattern recognition as it is currently defined may not be the most accurate reflection of human psychology, as it focuses on (generally) static, (usually) visible, entities, rather than the silent, invisible, cognitive *process* through which we experience our surroundings. The focus of "experience recognition", on the other hand, is a process in which the progression of the sensory influx of information is, at any point in time, mapped against a large store of experiences (*trajectories of information processing*).

Traditional (static) pattern recognition generally consists of training upon a data set which is presumed to be representative of all the patterns which will be seen and subsequent classification of every encountered pattern based on what was 'learned'. Dynamic pattern recognition, in turn, achieves 'online' learning, that is, a system remains able to add new patterns to its roster of recognizable elements. This amounts to a constant re-training of the model at every 'time iteration' adding the pattern seen at time t to the training set of $t+1$. This is a more flexible but yet lacking format, when the goal is a faithful model of human decision-making.

It is important to note that pattern recognition, as a field, is orthogonal to the pursuit of cognitive science: it is an immensely successful field with countless applied results, and does not hold the modeling of cognition as a principal (or even secondary, some might say) goal. We reinforce this notion by adding that the vast majority of pattern-recognition models do neither address fundamental issues of human decision-making, nor should they. Furthermore, we invoke pattern recognition here in order to address the fact that extant literature (Gobet, 1998; de Groot & Gobet, 1996; Gobet & Jackson, 2002; Gobet & Simon, 2000) employs models and arguments that display characteristics and techniques of static pattern recognition.

Computational models based on (static or dynamic) pattern recognition usually are i) context-free, ii) temporal-sequence-independent, and iii) culture-free. The current pattern being processed is (typically) not affected by the previous one(s) seen; the temporal sequence of patterns the system has acquired is (usually) irrelevant; and there are no developed biases: one pattern is as good as any other. Past trajectories of information processing do not seem to affect future processing to a large extent (other than, perhaps, incremental adjustments in pre-selected internal parameters). These characteristics seem to hold for a large number of methods appearing, for example, in journals such as *IEEE Transactions on Pattern Analysis and Machine Intelligence*. We postulate that the mind's pattern-recognizing has enormous *path-dependence*.

Additionally, priming studies and studies from the heuristics and biases school show the enormous extent to which we are *context-bound* (Bargh & Chartrand, 1999). In here, the concept of 'context' is fundamentally different from that explored in standard vector-space models such as Latent Semantic Analysis (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990). The mind's information-processing is highly path-dependent: someone who has experienced trauma is intrinsically different from someone who has never experienced it, yet is about to. And of course we are also strongly context-dependent: bound by taboos and customs, social norms, political institutions, formal and informal power (and status) structures, etc. (Tomasello, Carpenter, Call, Behne, & Moll, 2005). We postulate that the mind's pattern-recognizing is self-organizing, constantly creating and releasing constraints for further exploration: a process capable of molding itself.

Linhares and Freitas (2010) analyzed cognitive models of chess, influenced by Chase and Simon's (1973) well-known study, "perception in chess". The computational cognitive models in that analysis concentrated in the work of Gobet (1998), de Groot & Gobet, 1996, Gobet and Jackson (2002), Gobet and Simon (2000), which are quite simply the benchmark in chess cognition models. Lane and Gobet (2011) provide commentary and criticisms of Linhares and Freitas (2010). Here we respond to those criticisms, concentrating on the critique of Chase and Simon (1973), which is the most influential work in question, with over 800 citations in the ISI database,[1] and further expand on the aforementioned questions.

Lane and Gobet claim that the views expressed by Linhares and Freitas were "revolutionary" (using the term under quotes), yet no grandiose words like "revolution" or "paradigm" are to be found therein. In fact, the term experience recognition seems a natural convergence of a number of previous ideas, including: Hofstadter's (2001) 'analogy at the core of cognition', Klein's (1999) 'recognition-primed decision' framework, Hawkins' (2005) 'hierarchical temporal memory', and, to a lesser extent, Saariluoma and Kalakoski's (1998) view of 'apperception'.

In the following section we comment on Chase and Simon (1973) and address the concerns raised by Lane and Gobet (2011). In Section 3 we address and expand on necessary traits of experience recognition models.

2. Chess and chunks

In this section we address technical concerns over chess cognition raised in Lane and Gobet (2011) and expand on the concepts initiated in Linhares and Freitas (2010). We shall initiate our discussion by revisiting Chase and Simon (1973).

2.1. A flaw in delineating chunks: further analysis of Chase and Simon (1973)

Chase and Simon (1973) proposed a method for "isolating and studying the perceptual structures that

[1] For reasons of space, we cannot review the entire literature of chess cognition pointed out by Lane and Gobet; we refer the reader to Linhares (2005), Linhares and Brum (2007), and Linhares et al. (2011) for our basic position concerning cognitive computational architectures regarding the chess game. There is one point that bears responding, though: Figs. 1 and 2 of Lane and Gobet's article seem to imply that Linhares and Freitas claimed that "all positions can be entirely distorted", which is not true, obviously (the claim was that some positions can be entirely distorted, which shows that specific location coding cannot explain much). Finally, we of course do not expect to see rapid advances in context-dependence, path-dependence and cultural bindings. It is much easier to point out limitations in computational models than to develop the immense breakthroughs the field aspires toward. It is with respect and a constructive mindset that we criticize the work of others.

players perceive" (p. 55). They set out to find out the nature, size, number, and content of chess chunks. Consider their problem: How should one proceed in order to uncover the information encoded in chess chunks? The reader might find a two step solution to this problem:

1. Clearly delineate the object of study: clearly separate each chess chunk from others;
2. Analyze the information contained in the collected chunks.

This is standard practice, and these were the steps taken by Chase and Simon. Regrettably, they have committed a logic error in step (1) which invalidates the analysis in step (2).

Chase and Simon needed initially to delineate chess chunks: a task that would separate the information contained in a chunk. Theory states beginner has few chunks, while a master has a large array of chunks; so this task would need to provide *striking differences* in the acquired data. The arrow of causality would have the master's chunks generating different data from that generated by class A players, or by beginners.[2]

As their subjects reconstructed the chess positions, they measured the time taken. More specifically, the "within-glance" and "between-glance" timings involved. The within-glance timing was measured in interpiece-interval seconds taken *without glancing* at the original position; the between-glance timing measured seconds taken between pieces placed, *with a glance* taken at the original position. They hypothesized that: "pieces placed on the board by the subject in the perception task after a single glance correspond to a single chunk" (p. 64).

The timing data showed that most within-glance intervals were around 2–2.5 s. This would then provide their boundary for delineating chunks.

However, this timing data was the same, regardless of the skill of the player: There was no discernible difference between a master and a beginner when it came to the within-glance or between-glance data. The data implies that skill and experience do not affect the timings. The unanswered question here is: what *caused* the beginner's behavior? If not a vast repertoire of chunks—something beginners lack—what, then, made their timings indistinguishable from the advanced player, or the master player? Given that expertise entails a vast array of chunks and ignorance entails their absence, one is led to question the choice of timing by Chase and Simon as a delimiter of chunks. It is clear that chunks do not affect the timings, and a task that is not affected by chunks cannot be used to delineate chunks.

Consider Fig. 1, which presents the within-glance timing data of the perception experiment. That is, as subjects reconstructed the board, the figure plots the (relative) time spent on combinations of features (A: attack; D: defense; C: same color; S: same piece; P: proximity relations). The collected data does not clearly differentiate between the master and beginner (and class A subject), which seem to interchange positions quite randomly. The largest gap obtained is in the PS combination (Spatial Proximity; Same Piece), showing that the master subject took more time placing nearby pieces of same type (mostly pawns) than the Class A player, or the beginner. This could be important information concerning the chess chunk; however, an alternative explanation must include acquired motor behavior: since masters are more experienced at placing pieces at the board (given years or decades of play, setting up boards, and setting up chess problems), they may be able to reconstruct pawn chains with greater dexterity. Because the rest of the graph has these three players of immensely different skill levels randomly swapping positions, it should bring us pause. Do beginners switch positions with masters to this extent? There is no clear differentiation between skill levels.

Or consider Fig. 2, which plots the *probabilities* obtained by Chase and Simon. At least three serious issues stand out:

i) Positions continue to be interchanged between the master, the class A subject, and the beginner, with no discernible pattern;
ii) The "pattern" (PS) seen on the previous figure has all but vanished into one of the closest clusters of the data set, weakening the view that information was meaningful; and
iii) The intermediate player (Class A subject) is, quite incredibly, a strong outlier, presenting 7 extreme data points of 11, instead of intermediate data points.

Observing Figs. 1 and 2, one is led to ask: Where are the gaping skill differentials between players? How can the intermediate subject be this large an outlier? In fact, how could the intermediate player be an outlier at all? Can such data illuminate the content of the chunks that masters have accumulated but beginners lack? If so, how can we explain the seemingly random exchanges of (relative) positions between the skilled master, the intermediate class A player, and the naïve beginner?

Chase and Simon's task does not produce data in which the master's skill clearly stands apart from others. Is this really a sound data set in order to make inferences, as they have done, concerning acquired chunks, or skill? Chase and Simon certainly treat this data as if it were meaningful—they even extend Table 1 to compute averages and standard deviations, assuming a normal distribution given three observations.

Note that we are not pre-selecting a subset of data to confirm our views: these questions apply to Chase and Simon's presented data in their Figs. 3 and 4, Table 1–4[3] (please see next section). Hence our skepticism: if the data does not clearly differentiate the skilled from the incompetent, the data cannot be suitable for drawing

[2] There were two experiments: a *perception* experiment, and a *memory* experiment. Because the same error of logic was committed in both, I will refer to the perception experiment for simplicity; similar arguments hold for the memory experiment, as it also failed to produce disparities in the data.

[3] Figs. 1 and 2 concern the first part of Chase and Simon, which we believe is sound, and a remarkable achievement.

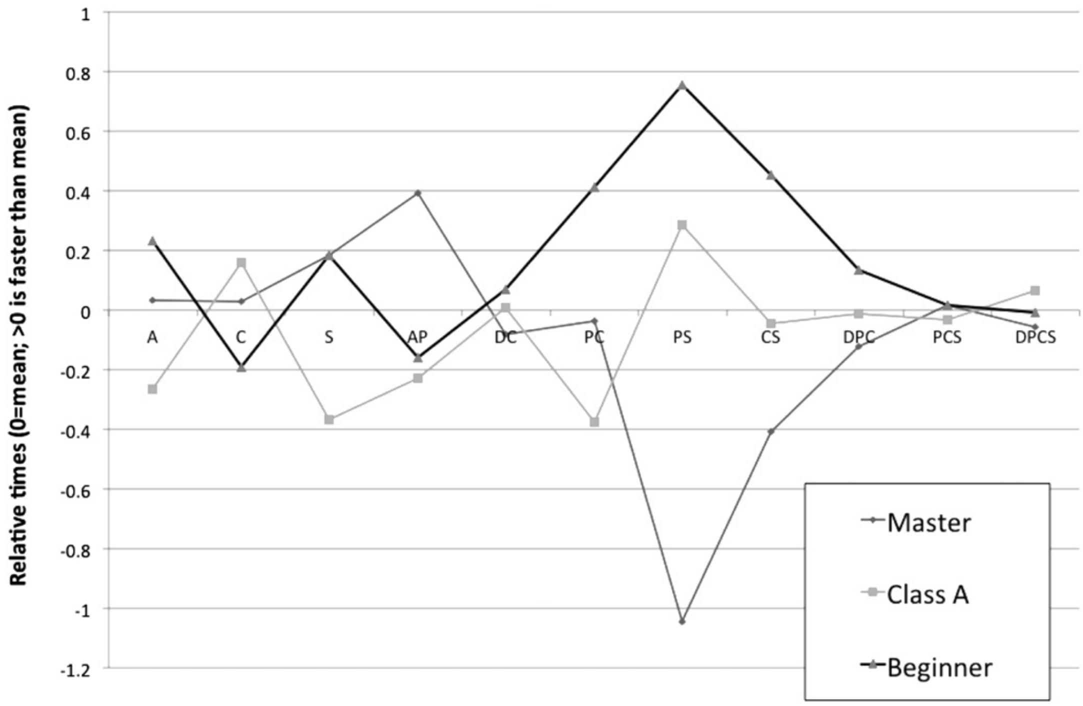


Fig. 1. Relative within-glance latencies obtained by Chase and Simon's perception task (data extracted from Chase and Simon's Table 1, relative to the mean, taken from the combinations in which data for the three subjects is available for comparison). Negative values imply that a subject took more time than the mean time (positive values imply less time taken). *A: chess pieces under attack; C: same color; S: same* piece type; D: defense relationships; P: proximity; AP, DC, PC, PS, DPC, PCS, DPCS are combinations of the underlying relationships.

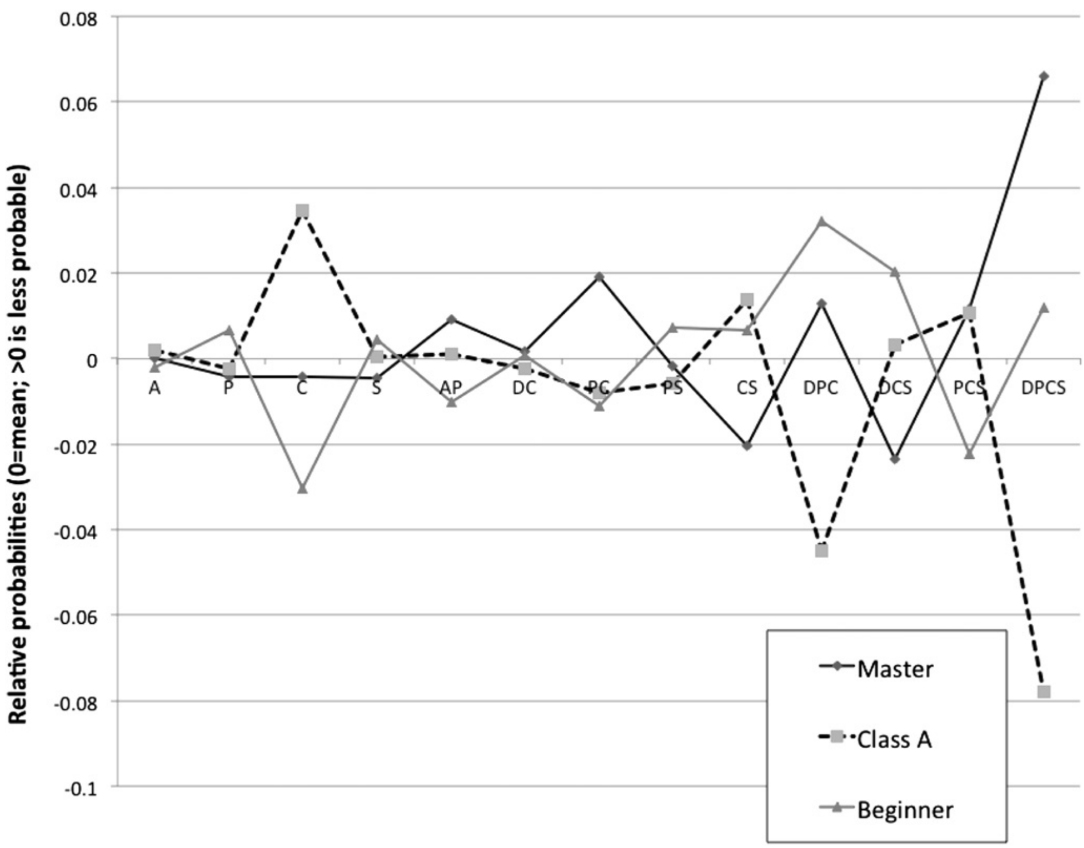


Fig. 2. The corresponding figure for the probability data for each subject (distance from the mean, data extracted from Chase and Simon's Table 1).

inferences concerning skill—which is what the study of chunks is, in our case.

2.2. Response to the arguments set forth in Lane and Gobet (2011)

Lane and Gobet state that the above claims are false, and provide some arguments, prior to re-stating Chase and Simon's claimed results:

1. "*Incorrect claim*". Lane and Gobet point out that the statement, that "the data was exactly the same for masters and beginners" (Linhares & Freitas, 2010, p.71) is incorrect. They point out that there was an instance in which small timing differences were found by Chase and Simon. Linhares and Freitas did indeed ignore that instance, for Chase and Simon (1973) themselves start the paragraph in question with the following: "The results are straightforward and roughly the same, with one exception, for all three subjects" (p. 62). Chase and Simon noted the *exception*, but basically ignored it, and moved on to other subjects. It would indeed have been precise to write "roughly the same" instead of "exactly the same"; yet, our claim does not alter any of the above arguments. There is no hard evidence that chunks do have causality over the timings. Numerous confounding factors could have produced those timings.
2. "*Quoted passage literally out of context*". Lane and Gobet claim that Linhares and Freitas have misled readers, through a quote out of context. This, of course, deserves clarification: In the second part of Chase and Simon, the collected data for the master player, the class A player, and the beginner, is "roughly the same" in Fig. 3, in Fig. 4, in Table 1 (correlation between *master* and *beginner* is an astounding .95, p. 65), in Table 2 (correlation between *master* and *beginner* is .89, p. 65), in Table 3 (the correlations are not informed, but "the pattern of probabilities was still the same between subjects", p. 70), and in Table 4 (correlation between *master* and *beginner* is again .95, p.70). There is such an overwhelming amount of "roughly similar" data between masters and beginners that, at one point, the word *data* was used in an encompassing form, including that from Figs. 3 and 4 and in the subsequent phrase, Chase and Simon's statements concerning the data from Tables 1 and 2. Note that this encompassing use of the term "data" (which include Chase and Simon's Figs. 3 and 4, and Tables 1–4) does not alter the argument—it only reinforces the fact that the task does not delineate chunks, or skill differences for that matter. With no causation between chunks and the test's results, the data cannot provide insight concerning chunks. The task generates noise, not signal.
3. "*Difference between within-glance and between-glance*": Lane and Gobet point out Chase and Simon's argument that there were indeed differences between the within-glance and between-glance data. Such differences would supposedly demonstrate that the information gathered came from chunks. One should be skeptical about this argument, because *this difference can be obtained by the beginner's data alone*. The within-glance-between-glance disparity can be found in the beginner's data, a subject who has no or few chunks. Anything that could be obtained by a beginner's data alone cannot teach us about expertise, by definition.

In Chase and Simon's concluding sections, they do arrive at some data that distinguishes masters from beginners. Here one finds, for example, that, in middle-game positions, the average size of chunks per trial is estimated to be, respectively, masters ∼7.7 pieces, class A ∼5.7 pieces and beginner ∼5.3 pieces. Chase and Simon accept these results (after discussing how peculiar they are, p. 77). Lane and Gobet seem to accept them, mentioning also that they reside within Miller's magic number (see Linhares, Chada, & Aranha, 2011 for criticisms of the computational modeling of short-term memory limits). But perhaps the conclusion that a *beginner* starts with a 5.3-sized chunk and might, with luck, have a gain of 38% (reaching ∼7.7) after a whole decade of hard work might be better seen as a *reductio ad absurdum* of the method employed.

Studying Chase and Simon led Linhares and Freitas to conclude that, in order to understand the content of chess chunks, one needs to look at *performance differentials*.

2.3. Performance differentials: an alternative empirical approach

Though we remain skeptical that any (current) system is able to precisely delineate chunk boundaries, we do believe that it is possible to study the *content* of chess chunks, by contrasting the (many) errors beginners commit against the (few) errors experts commit. Linhares, Freitas, Mendes, and Silva (2011) conducted a chess position reconstruction experiment, and compared the (numerous) *errors* committed by beginners with the (few) errors committed by experts. The results are different from Chase and Simon's, as the data shows an enormous gap between strategically perfect reconstructions and poor reconstructions. The reader might want to compare Fig. 3, displaying some of these performance differentials, to Figs. 1 and 2, produced with Chase and Simon's data.

Finally, in the following section we propose a "dense" chunks hypothesis. We may not know the precise size of chess chunks, and we may not know their precise number. There is a possibility that chunks are "dense", that they fluidly re-transform during one's processing of a scenario, and that a proper delineation might turn out to be an impossible task.

2.4. The density of chunks hypothesis

Alfred Binet is said to have remarked: "Could we look into the head of a chess player, we should see there a whole world of feelings, images, ideas, emotion and passion". It is possible that chess chunks encode much more than the visible, superficial, piece-to-square bindings. Consider Fig. 4, which presents one of the most interesting reproductions obtained in Linhares et al. (2011). If one looks at piece-square bindings, the reproduction seems to have lost most of the original information. As much as 52% of the

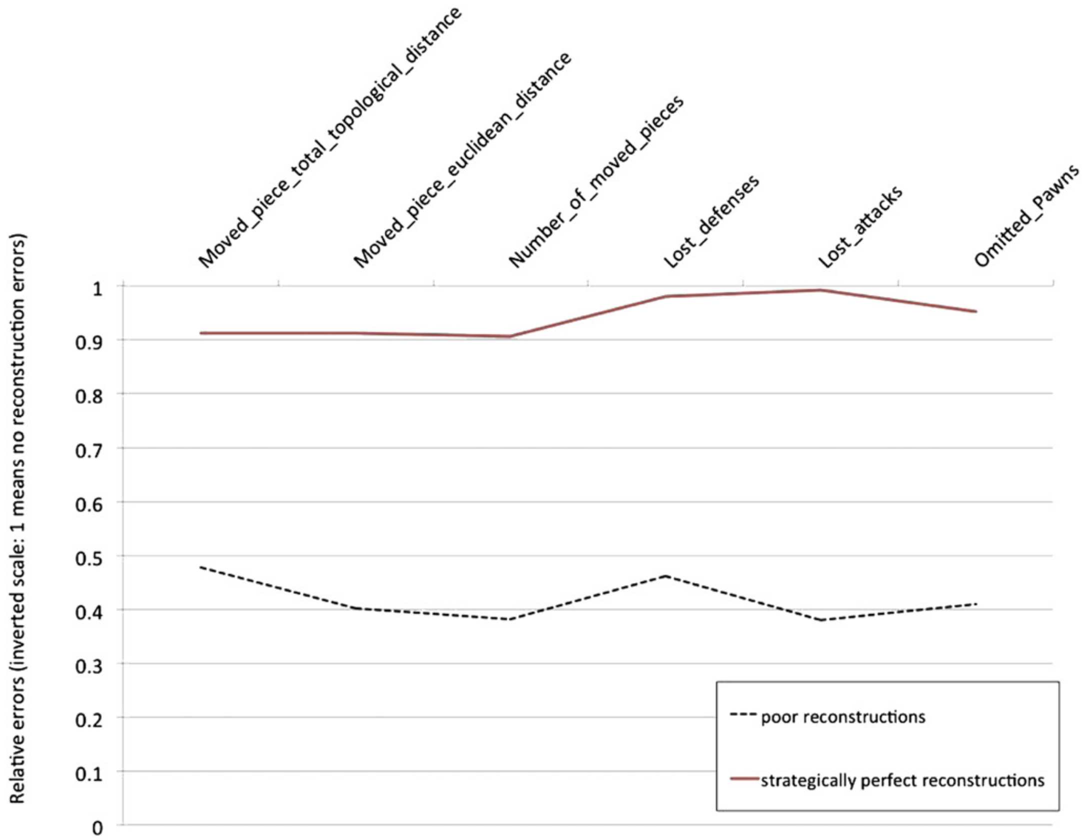


Fig. 3. Data collected in Linhares et al. (2011) compares the (many) reproduction errors conducted by beginners with the (few) errors conducted by grandmasters (though the main focus of the study was on the reproductions rather than the experts per se). This data clearly demarcates different performance levels (higher is better; 1 represents a perfect recollection of the feature in question). The strongest the *quality difference*, the higher the probability that a given feature is encoded in chess chunks. For example, grandmasters rarely neglect to reproduce an attack between two pieces (Lost_attacks feature), while reproductions by less skilled players usually do. This seems to imply that grandmasters have learned to rapidly encode such feature, while less skilled players are, at best, yet to.

pieces are either misplaced or omitted. Yet, the strategic situation remains "unchanged": in the original position, white's mate can only be delayed by blocking the rooks. In the reconstruction, the (misplaced) white queen assumes the same role as the original rook. Because chess is a combinatorial game, a single misplaced piece can, in most cases, alter the strategic situation. Our subject was able to preserve the higher qualities of the strategic scenario, while losing most of the superficial, visible, bindings. The probability that this is a 'false positive' is vanishingly low.

Recently, the field of "graph redescriptions" in theoretical computer science has shown how all computation can

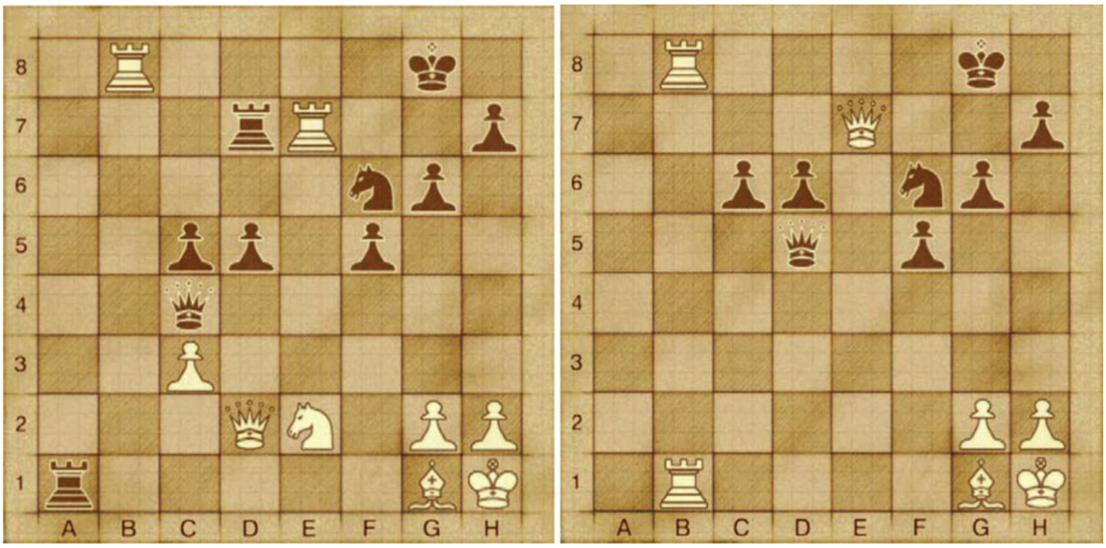


Fig. 4. Left: Original position. Right: a player's reconstruction preserves the essence of the situation while committing a large number of errors of various sorts.

be seen as transformations in graphs. Perhaps chess chunks, if viewed in graph theoretical terms, are *dense*, with multiple hierarchical levels of description, and increasingly difficult to cleanly tease apart from one another. Perhaps chunks have considerable overlap; encoding much more than the visible superficial information (piece type, square, color, etc.). Linhares (2005, submitted for publication), and Linhares and Brum (2007) have proposed, for instance, that chunks encode abstract roles perceived to be played by pieces (or even by empty space), and that chunks also encode a player's intention. If chunks were sparse or easily clustered, like prisoners self-organized into different gangs or students into cliques, it might be easier to separate them for analysis. If they are dense, however, then attempts at a clean demarcation may be counterproductive.

The idea that chunks may contain highly abstract features and have a dense topology stems from the computational school in cognitive science that—as opposed to Fodor's claims of "modularity"—postulates that there can be no "perception" module. This school proposes that there is no vision, or any other sense for that matter, without interpretation. And the interpretation process is both culturally and contextually dependent, indeed intrinsically so (Chalmers, French, & Hofstadter, 1992; French, 1995, 1997, 2008; Hofstadter & FARG, 1995; Linhares, 2000, 2008).

With this we conclude our response to Lane and Gobet's concerns. From here to the end of this text we aim to define some of the necessary traits of an experience recognition model. Moreover, we argue their relevance toward any computational model of cognition.

3. Experience recognition: Simagin's blunder and information-processing trajectories

If the human mind is to be seen as a pattern-recognition device, what kind of pattern-recognition task is it accomplishing? What insights for future computational models can we gather from an 'experience recognition' viewpoint? We believe we should strive for pattern-recognition that is i) highly path-dependent, ii) prioritizes internal encoding over external data, and iii) anticipates the adjacent possible.

Concerning computational models, two additional ideas seem valuable: firstly, we should strive for *theories of process*. Secondly, we should focus analysis and comparisons on the *information-processing pathways* (not on final outputs that match/surpass human subjects). We expand on these concepts in this and the following sections.

If one's objective is to create a world chess champion computer—as it had been for the AI community for 50 years—, how could we evaluate progress toward that goal? How could we know whether a new proposal was a scientific advance? How much can be gained by specific hardware, by a new pruning strategy, and so on? The historical answer has been to look at *results* obtained, in terms of quality of play. Hence, most research has been *results-oriented*. Yet, the criteria for measuring advances in psychologically plausible, *cognitive*, chess models should be *process-oriented*, as opposed to *output-oriented*.

Consider the notion of an "information-processing cone", defined by *all the possible future information-processing pathways of an agent at any given point in time*. Such a cone expresses all possible courses of interpretation and subsequent action by a decision-maker, and thus the entire space of causal reaction by that decision-maker.

Let us define 'understanding' as *relevance within an information-processing cone*: a system with few branch-points is *focused*, while a system with a wide trajectory cone is able to explore numerous pathways, that is, to consider a wider range of choice in reaction to a certain snapshot of stimuli. The range of possibilities reflects the *temperature* of the system at a given point in time (Hofstadter & FARG, 1995); a system with high temperature has a wider range of possibilities than a system with a narrow range. This space is defined by what Kaufmann (2000) calls 'the adjacent possible': the continuous cognitive process of anticipating multiple subconscious counterfactuals—and being surprised when the expectations are broken by an unanticipated state. Given a representational state at any point in time, the relevance of the set of possible future states of a cognitive system define its understanding, its level of *meaningful anticipation* (this notion is close to the 'memory-prediction framework' sketched by Hawkins, 2005). Understanding is not a property of a particular state, of a particular representation, or of a particular chunk. *Understanding is a property of a process*. Let us see an example in chess.

Consider the example in Fig. 5. This denotes what is meant by information-processing trajectories: could a tree-searching computer fall for Simagin's blunder? Since this is a shallow mate in three, there is no computer program of the tree-search paradigm that would fall for such mistake; as a small tree search leads to perfect play. But would its information-processing cone reflect Simagin's thought process? Vladimir Simagin's clear advantage resided in the passed pawn, and he was primed to promote it. The overconfident rush toward promotion shaped his information-processing cone into one much narrower than the computer's cone; that is, a cone that did not include the opponent's imminent checkmate. Under the definition of understanding as 'relevance within an information-processing cone', we can say that the computer's, however unnatural, 'understanding' of the combinatorial landscape involved in this chess scenario, surpassed Simagin's: there was a highly relevant adjacent counterfactual that Simagin overlooked and a machine wouldn't.

The computer here resembles an automatic assembly line, effortlessly working on each required step, *ad nauseam*. The chess player, on the other hand, resembles the process of craftsmanship; gradually converging into an interpretation of the situation—and as soon as the blunder is made it is perceived. At that point focus rapidly shifts from the shallow combinatorial world of chess into the vastness of the real world: the perceiving of inevitable defeat, the response from an overlooking audience, and a host of other questions far beyond the combinatorics of chess, becomes inevitable. This shift of focus back to the real world cannot be accomplished by the machine, and while it is wired up in a form that understands the combinatorics required to win, it does not understand what 'to win' means, or what 'understanding the combinatorics' means—these are states outside of its adjacent possible, and are not encompassed in its information-processing

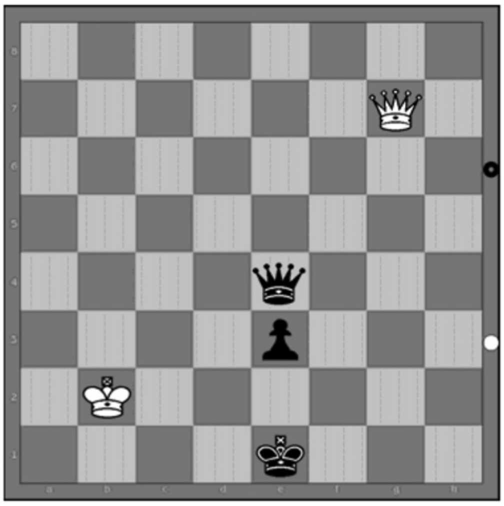


Fig. 5. Batuyev versus Simagin, in Riga, 1954. Black to move. Simagin, a player of extraordinary skill, playing black, committed a blunder which no tree-searching machine ever could. With material advantage of a passed pawn, it seemed that Simagin could handle a winning endgame. However, in rush to promote, he played pawn to e2. Since his material advantage consisted on the passed pawn, one can only resonate with Simagin's attempt to promote. This led to a mate in three: white queen to g1 check, king escapes to d2, white queen to c1 check, king escapes to d3, queen to c3 checkmate.

cone. A static state in time can represent an anticipation of adjacent counterfactuals, but it is the process in which these anticipated counterfactuals change through time, the process of creation and removal of constraints of the information-processing itself, and the relevance of these adjacent counterfactuals, that should define whether or not a model is psychologically plausible. Cognition anticipates the adjacent possible using solely chunks, combinations of chunks, and knowledge of the different forms in which chunks may bind to each other (linear, hierarchical, 2-dimensional, 3-dimensional, circular, etc.; see Kemp, 2007). Path-dependence molds the information-processing cone, by creating and releasing constraints into the next cycle of anticipation of the adjacent possible.

Internal encoding should take precedence to raw external data (patterns)—this was the main argument of Linhares and Freitas (2010). A war movie may bring different experiences to a veteran than the experiences of his friends: The same raw data outside of any understanding can lead to vastly different experiences in individuals.

Path-dependence and the constant molding of the adjacent possible bring us to some questions we propose should be addressed in evaluating whether cognitive (computational) models have a semblance of "experience recognition": contextual influence, cultural influence, domain-dependence, temporal-dependence and the ability of a model to readily self-criticize and self-organize (as we conjecture Simagin must have gone through). These are discussed in the following subsection.

3.1. Methodological questions to evaluate "experience recognition"

There are at least five questions to probe the path-dependence of a cognitive computational model's reflection of "experience recognition".

3.1.1. The information-processing cone must be "context-influenced"

A pattern is never seen outside of a context. From anchoring to availability to framing, people are subject to strong influence from context. Consider the ubiquity and enormous influence of priming, for example: Williams and Bargh (2008) recently conducted a striking experiment, in which subjects were asked to interview a person and decide whether or not the person should be hired. The experimental setting had the subject meeting with a researcher, and both going up in an elevator to meet the interviewee. Inside the elevator, the subject was casually asked to hold a warm cup of coffee, or a cold cup of iced tea. The results imply that the physical warmth experienced during these brief moments transfers to interpersonal warmth: subjects that held a cold iced-tea cup are less likely to recommend hiring the interviewee. Context may radically alter the adjacent possible and each "pattern" is highly context-sensitive.

3.1.2. The information-processing cone must be "culture-influenced"

Imagine a one-year old baby, laying down, and staring up. What would be in its field of view? Due to the availability heuristic, many readers are likely to imagine a room, toys, a cradle, walls and a ceiling, doors, windows, cabinets, and so forth. Yet, because the reader is very likely to come from a Western, Educated, Industrialized, Rich, and Democratic background—or, as Henrich, Heine, and Norenzayan (2010) put it, "the WEIRDest people in the world"—caves, igloos, dense forests, ant trails, deserts, or houses made with straw, such as those from the Atr Tribes of Eritrea, are unlikely to be imagined. At this very moment, a large number of babies is most likely starring at things beyond our imagination—and such experiences will profoundly shape their psychology.

In practice, consider visual illusions, such as those of depth perception. The Ponzo illusion superimposes two lines of exact same size over the background of a "train track", and, unsurprisingly, we tend to see the lines with different sizes: for our entire lives, we have lived in a carpentered world, with right angles at walls, doors, windows, cabinets, television sets, computer screens and keyboards, tables, roads and streets, buildings, books and articles, and so forth. Our experience has taught us to use these angles to see perspective, so that objects "farther away" in our projection must be larger if they have the same size in their light projection.

But this is not universal. Those from extremely different backgrounds, who have lived most of their lives in a world that is not carpentered, do not fall prey to such illusions, and can point out that the Ponzo illusion lines do have similar size (Shiraev & Levy, 2007). In fact, people with lower IQ are less susceptible to such illusions (Frederick, 2005).

> "…these findings suggest that visual exposure during ontogeny to factors such as the "carpentered corners" of modern environments may favor certain optical calibrations and visual habits that create and perpetuate this illusion. That is, the visual system ontogenetically

adapts to the presence of recurrent features in the local visual environment. Because elements such as carpentered corners are products of particular cultural evolutionary trajectories, and were not part of most environments for most of human history, the Müller-Lyer illusion is a kind of culturally evolved by-product." (Henrich et al., 2010).

For some people, there is no illusion. The exact same pattern, projected into different eyes, creates vastly different representations.

A converse version of this statement also holds: Extremely different patterns can, through analogy, create strikingly similar psychological experiences. French (1995), Hofstadter (2001), Hofstadter and FARG (1995) and Mitchell (1993) study this issue in depth.

3.1.3. Not all experiences are the same

People can develop a choice set skewed toward some options (or against others). At the extreme, opiates (or the stimulation of the pleasure centers of the brain) rapidly become addictive, causing a shutdown of most alternative options of behavior that detract from further stimulation. At another extreme, people face strongly negative experiences, which can induce irrational overestimation of its chances and enormous effort in order to avoid a recurrence. At the very least, experiences should lie somewhere between undesirable, indifferent, and desirable. As we seem to be far from a information-processing theory of emotion (or of pain or pleasure), the issue here is not that traditional pattern-recognition systems cannot be strongly influenced by a new pattern. The issue is with claims that static pattern-recognition systems be used to reflect human psychology. (In the next section, we will see how some current pattern-recognition systems can generate extremely inappropriate results, from a human viewpoint.)

3.1.4. Knowledge and behavior are domain-dependent

Consider Wason and Shapiro's (1971) selection task: given a set of four cards, with numbers on one side and letters on the other, and a rule, such as "Every card which has a vowel on one side has an even number on the other side" (p. 64), subjects find it difficult to select which cards should be turned to test the rule. If, however, the same logical test is applied in a setting in which subjects have experience, the task becomes "trivial" (i.e., "underage students cannot drink alcohol" leads to the correct test of those drinking alcohol and those underage). The exact same logic underlies both problems, yet, if presented within a setting in which we do not have experience, less than 10% of subjects respond correctly. Lack of experience is a lack of chunks; or, as Wason and Shapiro (1971, p. 70) write: "The abstract material has no unifying link: each card is distinct and separate rather than being parts of a whole".

The university is one domain, while real life is another. Kahneman and Tversky have famously shown how even professional statisticians, when detached from the classroom, fall for basic probability biases. Shane Frederick (personal communication; also discussed in Postrel, 2006) found that numerous university-educated people prefer 100% chances of $500 to a 15% chance of a million dollars—breaking the 'expected utility' viewpoint by many orders of magnitude.

Fig. 6. State-of-the-art pattern recognition can yield inappropriate experiences. Le Monde presents some of the first reports of the AF447 tragedy: an Air France flight from Rio de Janeiro to Paris collapsed into the Atlantic, in June 2009. Three selected advertisements from Korean Air are displayed, with soothing messages such as "le monde est ma destination" or "excellence in flight".

Fig. 7. CNN reports the Virginia Tech Massacre. An advertisement on display mentions "even when empty, it's fully loaded".

Fig. 8. The Washington Post reports that the late Steven Jobs, CEO of Apple, announces a leave of absence due to his cancer treatment. An advertisement starts with the message that "Jobs become obsolete. Talent doesn't". The intended meaning of the advertisement is clearly orthogonal to the story, yet the pattern-matching system is unable to perceived the inadequacy of the construed meaning.

The implication for cognitive models seems clear: experience gained in a contextual setting does not immediately transfer to other domains. It takes work to create the analogies and metaphors that will provide the necessary parallels between domains. This conclusion is anathema to many computational models, such as Newell, Shaw, and Simon's (1959) 'General Problem Solver' system.

3.1.5. The time dimension is necessary, but not sufficient

Though we propose that the temporal dimension is crucial for analysis, it is not, by itself, sufficient. There are numerous computational models that deal with patterns changing through time; for instance in pattern recognition in videos or sound waveforms. What we are concerned with here is how the internal processes and representations and mental states evolve in time, are affected by prior experiences and how they could have evolved. When we parse a phrase like:

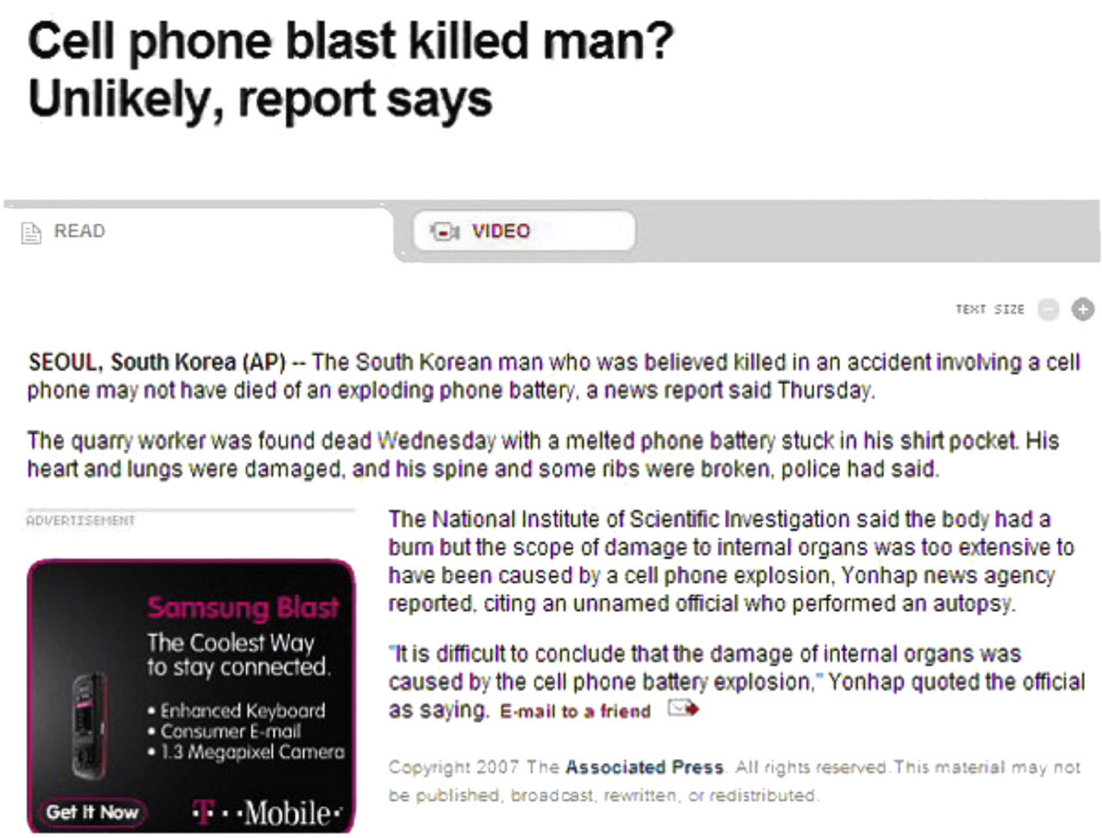
"Prostitutes appeal to Pope"

There is a temporal dimension to our processing, in which we deal with the cognitive dissonance of such a phrase and re-parse it. The first anticipation brought to mind is clearly untenable; the adjacent possible demands that we reject large amounts of previous knowledge. Here one is forced into an interpretational "double-take", the original input must be re-parsed, taking into account that the most likely pattern is to be explicitly ignored and rejected. Any temporal model will require input from culture and context to process the above example. This cannot be achieved through temporal pattern recognition alone. This temporal processing is different from the application of pattern-recognition methods to temporal data, followed by measurement of results obtained.

Furthermore, the temporal aspect must not be simply translated into adaptability within the model. Here we posit that the interpretation of the current pattern (say, at time t) must be most influenced by the most recent patterns experienced ($t-1, t-2 \ldots t-n$). A quick visual of this process would be the old children's gag, where one has to say the word 'fork' out loud at least ten times then is asked: 'What do we use to eat soup?'. While this is fairly straightforward to model in certain computational models (such as connectionist models), we believe it remains essential in any experience recognition model.

Some models have partly employed the concepts which fill our proposed characteristics. Recent models which address episodic memory and bias acquisition in cognitive systems potentially incorporate path-dependence, context-and-domain-dependence and the considerable influence of previous experience, though none all at once (for examples see Kemp, 2007; Marshall, 1999; Ramamurthy, D'Mello, & Franklin, 2004, 2006). Note also that these points resemble Piaget's notion of the assimilation—accommodation cycle (Piaget, 1962, 1977; see also the discussion in Indurkhya, 2007, p. 29).

In the next subsection we show that it is possible to have high-quality pattern-recognition and low-quality experience recognition. This not only emphasizes the distinction between the two concepts in an applied scenario, but also is



Cell phone blast killed man? Unlikely, report says

READ VIDEO

TEXT SIZE

SEOUL, South Korea (AP) -- The South Korean man who was believed killed in an accident involving a cell phone may not have died of an exploding phone battery, a news report said Thursday.

The quarry worker was found dead Wednesday with a melted phone battery stuck in his shirt pocket. His heart and lungs were damaged, and his spine and some ribs were broken, police had said.

ADVERTISEMENT

Samsung Blast
The Coolest Way to stay connected.
• Enhanced Keyboard
• Consumer E-mail
• 1.3 Megapixel Camera
Get It Now
T · · Mobile ·

The National Institute of Scientific Investigation said the body had a burn but the scope of damage to internal organs was too extensive to have been caused by a cell phone explosion, Yonhap news agency reported, citing an unnamed official who performed an autopsy.

"It is difficult to conclude that the damage of internal organs was caused by the cell phone battery explosion," Yonhap quoted the official as saying. E-mail to a friend

Copyright 2007 The Associated Press. All rights reserved. This material may not be published, broadcast, rewritten, or redistributed.

Fig. 9. A story questions whether a cell phone battery blast killed a man in South Korea. The contextual advertisement on display is for a cell phone named "Samsung Blast".

a display of a domain in which static pattern recognition can be crucially orthogonal to experience recognition.

3.2. High-quality pattern-recognition and low-quality experience recognition

It is possible to have high-quality pattern recognition and low-quality experience recognition. By this we mean that humans would be able to perceive a high-quality matching of patterns, simultaneous with a lack of understanding of at least one of the principles pointed out above (context-sensitivity, cultural sensitivity, desirability, domain-dependence, and the temporal dimension). A natural domain in which this arises is in the industry of “contextual advertising”.

Evans (2008) shows a clear example of how network effects favor a winner-take-all industrial organization: Google, the market leader, has more pageviews, more advertisers (and hence a greater advertisement inventory). By comparing the results of a search for “Germany SIM cards”, for instance, 8 out of 10 of Google advertisement inserts are directly relevant to the search, contrasted to only 2 out of 10 from Microsoft’s system. Evans concludes that “search-ad platforms with more advertisers will generally deliver more relevant advertisements to the searcher”. This should raise the incentive for advertisers to stick with the market leader. Notice that this higher advertisement inventory effect is not only tied to web searches, but also to inserts placed on web pages. Google’s massive access to web pages tends to increase their advertisement scope and relevance. Evans mentions that “like modern finance, online advertising relies heavily on advanced economic and statistical methods”. Economic models (such as auctions) determine specific keyword advertisements pricing, while statistical methods are used for the automatic insertion of advertisements. Contextual advertisement targeting has been a topic of increasing scientific and technological importance (Bhatnagar & Papatla, 2001; Chickering & Heckerman, 2003). A growing number of studies and patents is dedicated specifically to ad-placement (e.g., Bhatnagar & Papatla, 2001; United States Patents, 2005, 2007).

In contextual advertisements, a webpage with baby-care content should not carry advertisements for sophisticated violins, pig medication, or SIM cards in Germany. There is a very small likelihood that the reader of such a webpage will have an *a priori* interest in violins. Even if that is the case, in the particular “baby-care” primed state in which readers may find themselves in, their interest in violins will be, in all likelihood, considerably diminished.

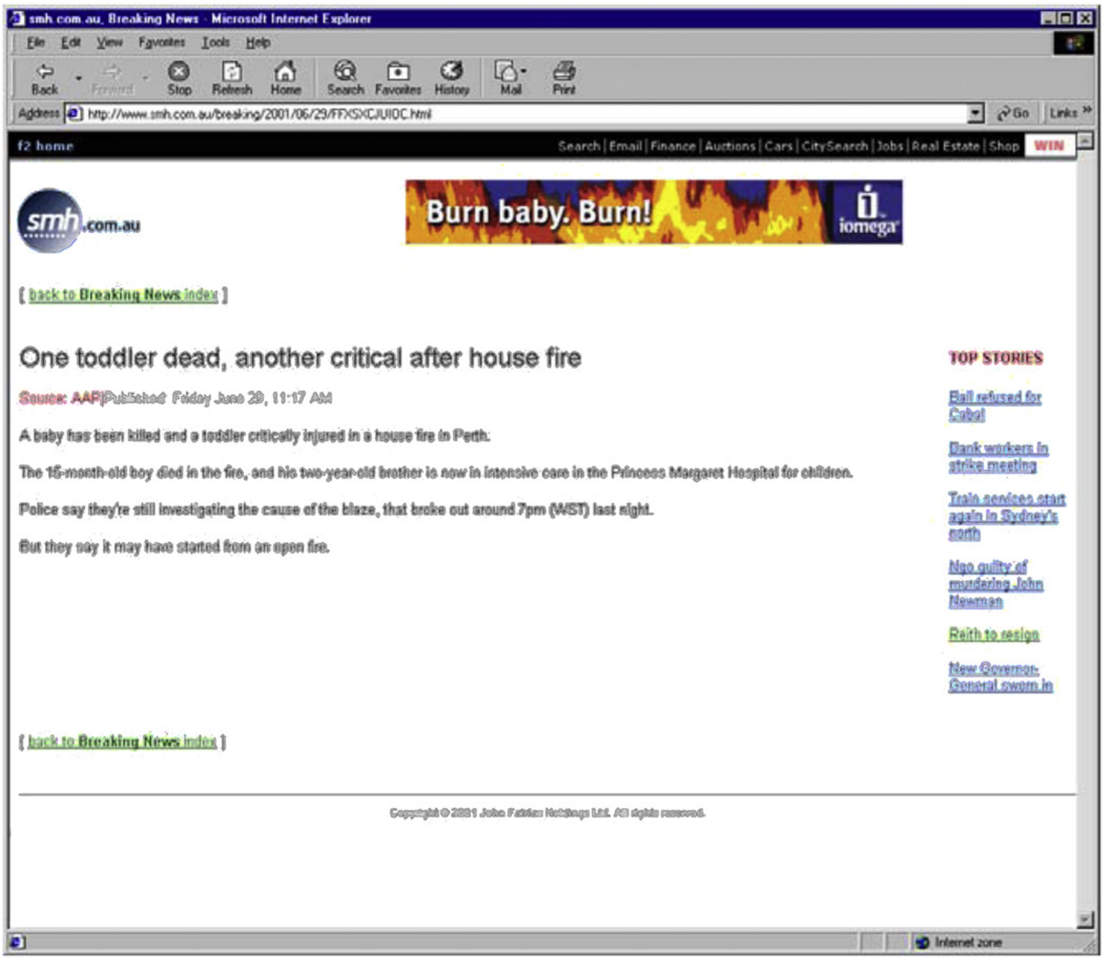


Fig. 10. A story mentions a house fire that took the life of a toddler, and another lies critical. The contextual advertisement, referring to cd-burners, mentions “Burn baby. Burn!”

Contextual advertisement by pattern-recognition systems automatically selects an advertisement that should, in principle, maximize revenue—by maximizing relevancy for users. Yet, they can be extremely inappropriate. Despite being relatively rare, these inappropriate inserts might still occur in large absolute numbers due to sheer number of daily advertisements displayed on the web.

3.2.1. Examples of inappropriate advertisements

As examples of high-quality pattern recognition, and low-quality experience recognition, we refer readers to Figs. 6–10.[4] These examples demonstrate advertisement systems that are able to produce high-quality pattern-matching, while being oblivious to the inadequacy of the construed meaning. Billions of dollars are spent in these high-quality pattern-matching machines, yet they occasionally present incredibly inappropriate ads. Despite the name "contextual advertisements", these messages are oblivious to context.

4. Conclusion

We cannot escape our experience. Every new experience we live is shaped by its predecessors. Yet, most extant cognitive models do not reflect the biases born out of accumulated experience or the pervasive influence of context. We postulate that the mind's pattern-recognizing process holds the following properties: it is a highly path-dependent process; it prioritizes internal encodings; it is a self-organizing process in constant change; and it constructs its future information-processing pathways by continuously recognizing the possibilities that lie within the adjacent possible. We believe these are required traits of "experience recognition".

Our vocabulary for the future was written in the past. The History of science and technology is filled with accidents of experience recognition; when one's personal experience is transferred to a new, poorly understood, domain, and the words originally used to describe this new domain survive far into the future: meet the new Porsche, with enormous 'horsepower'. In Biology, the term "cells" was given by Robert Hooke because, to him, their view under a microscope resembled rooms on a monastery. Later, parts of the cell were named as Chromosomes ("colored bodies"), for dyes made them colorfully visible under a microscope. In Astronomy, Ptolemy rejected the idea of a moving Earth because that would imply unimaginably powerful winds, whereas the geostationary Earth was consistent with the lack of such winds. Ptolemy's personal experience prevented him from believing the Earth could be moving across a space without any wind. As Indurkhya (2007) says, "When we consider changes in ontology, [an] issue to ponder is whether there are any criteria that identify certain ontology-changing moves rational and others irrational. There is obviously difficulty in positing any such criterion *a priori*. This is because in projecting a new theory onto a phenomenon, one cannot determine beforehand what the new ontology of the phenomenon would be and what consequences it might have". (p. 33)

As we attempt to model human information processing, a productive avenue is, perhaps, to study how people match their surroundings to previous pathways of information processing—"experience recognition". If we focus on the experience, the information-processing trajectories, we may be able to better understand how our computational models reflect, or fail to reflect, human cognition. Through careful comparisons between human information-processing trajectories with trajectories generated by computational models, we may eventually be able to glance at advances in computational theories of intelligence, skill, and, perhaps, inner experience.

[4] The interested reader may contact us for additional examples. See also French and Labiouse (2001).

Acknowledgments

This work has been generously supported by grants from the FAPERJ Foundation (grants E-26/110.540/2012 and E-26/111.846/2011), grants from the CNPq Foundation (grants 401883/2011-6 and 470341/2009-2), the Pro-pesquisa program of FGV Foundation.

References

Bhatnagar, A., & Papatla, P. (2001). Identifying locations for targeted advertising on the Internet. *International Journal of Electronic Commerce, 5*, 23–44.
Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*, 462–479.
Chalmers, D. J., French, R. M., & Hofstadter, D. R. (1992). High-level perception, representation, and analogy: a critique of artificial intelligence methodology. *Journal of Experimental and Theoretical Artificial Intelligence, 4*(3), 185–211.
Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81.
Chickering, D. M., & Heckerman, D. (2003). Targeted advertising on the web with inventory management. *Interfaces, 33*, 71–77.
Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science, 41*(6), 391–407.
Evans, D. S. (2008). The economics of the online advertising industry. *Review of Network Economics, 7*, article 2.
Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*, 25–42.
French, R. M. (1995). *The subtlety of sameness*. Cambridge: MIT Press.
French, R. M. (1997). When coffee cups are like old elephants or why representation modules don't make sense. In A. Riegler, & M. Peschl (Eds.), *Proceedings of the international conference new trends in cognitive science* (pp. 158–163). Vienna: Austrian Society for Cognitive Science.
French, R. M. (2008). Relational priming is to analogy-making as one-ball juggling is to seven-ball juggling. *Behavioral and Brain Sciences, 31*, 386–387.
French, R. M., & Labiouse, C. (2001). Why co-occurrence information alone is not sufficient to answer subcognitive questions. *Journal of Theoretical and Experimental Artificial Intelligence, 13*, 419–429.
Gobet, F. (1998). Expert memory: a comparison of four theories. *Cognition, 66*, 115–152.
de Groot, A., & Gobet, F. (1996). *Perception and memory in chess*. Assen: Van Gorcum.
Gobet, F., & Jackson, S. (2002). In search of templates. *Cognitive Systems Research, 3*(1), 35–44.
Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science, 24*, 651–682.
Hawkins, J. (2005). *On intelligence*. St. Martin's Griffin.
Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61–83.
Hofstadter, D. R. (2001). Epilogue: analogy as the core of cognition. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind:*

Perspectives from cognitive science (pp. 499–538). Cambridge, MA: MIT Press.

Hofstadter, D. R., & Fluid Analogies Research Group (FARG). (1995). *Fluid concepts and creative analogies: Computer models of the fundamental mechanisms of thought*. New York: Basic Books.

Indurkhya, B. (2007). Rationality and reasoning with metaphors. *New Ideas in Psychology, 25*, 16–36.

Kaufmann, S. (2000). *Investigations*. New York: Oxford University Press.

Kemp, C. (2007). *The acquisition of inductive constraints*. PhD Thesis: MIT.

Klein, G. (1999). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Lane, P. C. R., & Gobet, F. (2011). Perception in chess and beyond: commentary on Linhares and Freitas (2010). *New Ideas in Psychology, 29*(2), 156–161.

Linhares, A. (2000). A glimpse at the metaphysics of Bongard problems. *Artificial Intelligence, 121*, 251–270.

Linhares, A. (2005). An active symbols theory of chess intuition. *Minds and Machines, 15*, 131–181.

Linhares, A. (2008). Dynamic sets of potentially interchangeable connotations: a theory of mental objects. *Behavioral and Brain Sciences, 31*, 389–390.

Linhares, A. The emergence of choice: decision-making and strategic thinking through analogies, manuscript submitted for publication.

Linhares, A., & Brum, P. (2007). Understanding our understanding of strategic scenarios: what role do chunks play? *Cognitive Science, 31*, 989–1007.

Linhares, A., & Freitas, A. E. T. A. (2010). Questioning Chase and Simon's (1973) "Perception in chess": the "experience recognition" hypothesis. *New Ideas in Psychology, 28*(1), 64–78.

Linhares, A., Chada, D. M., & Aranha, C. N. (2011). The emergence of Miller's magic number on a sparse distributed memory. *PLoS One, 5*, e15592.

Linhares, A., Freitas, A. E. T. A., Mendes, A. S., & Silva, J. S. (2011). Entanglement of perception and reasoning in the combinatorial game of chess: differential errors of strategic reconstruction. *Cognitive Systems Research*. http://dx.doi.org/10.1016/j.cogsys.2010.12.006.

Marshall, J. B. (Nov. 1999). *Metacat: a self-watching cognitive architecture for analogy-making and high-level perception*. PhD Thesis: Indiana University.

Mitchell, M. (1993). *Analogy-making as perception*. Cambridge: MIT Press.

Newell, A., Shaw, J. C., Simon, H. A. (1959). Report on a general problem-solving program. In *Proceedings of the international conference on information processing* (pp. 256–264).

Piaget, J. (1962). In C. Gattegno, & F. M. Hodgson (Eds.), *Play, dreams and imitation in childhood*. New York: W.W. Norton, (Original work published 1945).

Piaget, J. (1977). In M. Cook (Ed.), *The origin of intelligence in the child*. New York: Penguin, (Original work published 1936).

Postrel, V. (January 26 2006). Would you take the bird in the hand, or a 75% chance at the two in the bush? *The New York Times, Economic Scene Column*.

Ramamurthy, U., D'Mello, S. K, & Franklin, S. (Oct. 2004). Modified sparse distributed memory as transient episodic memory for cognitive software agents. In *Proceedings of the IEEE international conference on systems, man and cybernetics – SMC2004*. The Hague, Netherlands.

Ramamurthy, U., D'Mello, S. K., & Franklin, S. (Jul. 2006). Realizing forgetting in a modified sparse distributed memory system. In *The 28th annual conference of the cognitive science society* (pp. 1992–1997). Vancouver, Canada.

Saariluoma, P., & Kalakoski, V. (1998). Apperception and imagery in blindfold chess. *Memory, 6*(1), 67–90.

Shiraev, E., & Levy, D. (2007). *Cross-cultural psychology* (3rd ed.). Boston: Pearson Education.

Simon, H. A. (1986). The information processing explanation of Gestalt phenomena. *Computers in Human Behavior, 2*, 241–255.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675–735.

United States Patent #20050149395 (2005). *System and method for real-time web page context analysis for the real-time insertion of textual markup objects and dynamic content*.

United States Patent #20070156519. (2007). *Method and system for providing sponsored content based on previous provided content*.

Wason, P. C., & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology, 23*, 63–71.

Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science, 322*, 606–607.