Cognitive emulation in expert system design

PHILIP E. SLATTER

Telecomputing plc,
244 Barns Road,
Oxford, OX4 3RW,
England

SUMMARY

Cognitive emulation is an expert system design strategy which attempts to model system performance on human (expert) thinking. Arguments for and against cognitive emulation are reviewed. A major conclusion is that a significant degree of cognitive emulation is an inherent feature of design, but that an unselective application of the strategy is both unrealistic and undesirable. Pragmatic considerations which limit or facilitate the viability of a cognitive emulation approach are discussed. Particular attention is given to the conflict between cognitive emulation and established knowledge engineering objectives, detailed over 12 typical expert system features. The paper suggests circumstances in which a strategy of cognitive emulation is useful.

1 Introduction

Expert systems use Artificial Intelligence [AI] techniques to solve problems ordinarily requiring human expertise. Throughout their brief history expert systems have been loosely modelled on the behaviour of human experts. More systematic attempts at simulating experts' decision processes have only rarely been reported. Notable exceptions are: INTERNIST [60], which modelled the clinical reasoning of a diagnostician in internal medicine, and PSYCO [34], a system for diagnosing dyspepsia that incorporated various known principles of human information processing.

However, most of the better-known systems—including MYCIN [32], PROSPECTOR [33], DENDRAL [31], and R1 [35] were constructed with little or no explicit aim of modelling expert thinking. Outside research-orientated establishments there is, if anything, even less commitment to cognitive emulation. In commercial applications of today's technology the emphasis is squarely on achieving expert-level performance using formal problem-solving methods [22]. Despite this, it has been suggested that the future development of the field could gain considerably from being more closely modelled on expert cognition [8, 17, 19, 52].

Differences between psychological modelling and expert system design have occasionally been discussed [e.g. 13, 15], but the viability and implications of a cognitive approach in Knowledge Engineering has not been discussed in detail. This paper first reviews the major arguments for and against the cognitive emulation approach to knowledge engineering, based on a review of literature from Cognitive Psychology and expert systems. Pragmatic considerations are then identified that represent important constraints on the utility of emulation. Emphasis is given to exploring areas of conflict between cognitive emulation and other Knowledge Engineering objectives. However, other factors which should facilitate cognitive emulation, especially in the longer term are also discussed. Finally, some guidance is offered on when cognitive emulation should be explicitly considered in expert system design.

Note: Unless otherwise stated the term "expert system" refers to standard applications of commercially available tools and techniques. "Cognitive emulation" refers to a strategy in expert system design which seeks to emulate human cognitive processes.

2 Pros and cons of emulation

2.a Arguments for cognitive emulation

2.a.1 Cognitive emulation is inherent in knowledge engineering

As presently conceived expert systems almost inevitably resemble human expert thinking to a significant degree because expert systems make use of domain knowledge ordinarily elicited from a human expert for solving problems. Indeed, knowledge-based techniques are a defining characteristic of the current generation of expert systems. Cognitive research clearly shows how dependent human expert performance is on the use of large quantities of specialist knowledge acquired over many years [e.g. 81, 85].

Knowledge-based approaches evolved for sound pragmatic reasons and should thus endure. At first AI researchers tried to solve problems normally requiring human expertise using formal domain-independent methods. It was the ineffectiveness of these methods, coupled with the subsequent success of experimental systems employing domain knowledge, that led to the present emphasis on knowledge-based techniques. The history of the DENDRAL project [31], clearly illustrates this change. Many designers would go further and argue that it is necessary to capture something of how the expert represents his or her knowledge, and the reasoning strategies employed, in order to construct an effective system [28].

Rule-based approaches, and in particular production systems, can be seen in this light. Rule-based formalisms are currently the most popular approach to knowledge representation; indeed the term "expert system" is sometimes defined as using rule-based techniques [8]. Psychological justification has been claimed for this formalism [23], reference usually being made to Newell and Simon's [56] pioneering use of production systems for simulating human problem solving. (Production systems have subsequently been employed in modelling a variety of cognitive processes [68]. Anderson's [37, 38] ACT system is an ambitious attempt to model human cognition using an architecture based on production systems.) A large volume of psychological research indicates that certain types of knowledge may be mentally represented in a rule-like form. [e.g. 28, 69, 81, 85].

Other knowledge representation formalisms are of course supported by commercially available expert system packages [21]. For example, frames and semantic networks. However, as with production systems, these formalisms have also been taken seriously as psychological models [4, 37, 40].

Machine induction techniques for producing knowledge bases are also consistent with cognitive emulation. This is because a domain specialist's knowledge is usually employed in the initial selection of concepts, attributes and exemplars. Consequently these selections will reflect the experts' underlying conceptualisation of the domain. Moreover, at least one well-known machine induction product—Donald Michie's ExpertEase—is derived from an algorithm for concept learning originating in psychological research: Hunt's Concept Learning System [46].

Breuker and Wielinga [13, p. 24] summarise the present argument when they observe: "in general there is a large overlap between knowledge-based systems and psychological models of the same task for the simple reason that up to now it has been hard to conceive of more general, intelligent methods than those used by human (expert)s."

2a.2 Performance

Expert systems seek to achieve expert-level performance. Often the only available criterion for assessing the effectiveness of a system is to compare its behaviour with that of one or more human experts—the so-called "gold standard" (see e.g. [21]). In the absence of more objective criteria the expert's judgement is the only gold standard; it can be satisfied by simulating the (cognitive) processes of the expert. Gaschnig et al [19, p. 255] remark that "there is increasing realisation that expert-level performance may require heightened attention to the mechanisms by which domain experts actually solve the problems for which the expert systems are typically built. It is with regard to this issue that the interface between knowledge engineering and cognitive psychology is the greatest."

2.a.3 User acceptance

Gaining acceptance for a system is a major consideration in expert system design [16, 21]. This stems from the failure of such technically successful systems as MYCIN to be accepted in daily use by the intended users, leading to the realisation that achieving expert-level performance per se may not guarantee user acceptance. In so far as it leads to greater intelligibility of the system's knowledge and the processes using it, and to a greater tolerance of intervention by the user, cognitive emulation can be expected to promote user acceptance [13, 17].

A related argument is that cognitive emulation is morally or socially desirable. It could help humanise what might otherwise become an alien, machine-orientated technology [24, 25].

2.a.4 More effective knowledge acquisition and representation

Building a significant expert system is a difficult and time consuming process. It has recently been argued that a major contributory factor is the inappropriateness of techniques currently used by system builders to both elicit and represent experts' knowledge [18, 30]. Even within a narrow domain an expert's knowledge can be of many different types. Much of it is difficult to capture and express as rules. The use of psychologically-orientated techniques in both knowledge elicitation and knowledge representation has therefore been advocated (See also [19]).

2.b Arguments against cognitive emulation

Despite these clear claims to psychological plausibility rule-based formalisms have also evolved as a major technique principally because of their knowledge engineering virtues (e.g. modularity [1, 2, 15]). There are also a number of direct arguments against an exclusively cognitive approach.

2.b.1 Human cognitive weakness

Human cognition compares unfavourably with computer systems in a number of respects:

a. Human memory is prone to forgetting and distorting stored information [41, 43].
b. The rationality of human thinking (in the formal logical sense) can be questioned. For example, aspects of deductive reasoning such as negation prove problematical for human subjects [49, 50].
c. People experience difficulties with important mathematical concepts, e.g. probability and the combination of uncertain evidence [51].
d. Human information processing capacity is severely limited [54].

While these observations about the comparative weaknesses of human thinking are valid, the conclusion that formal computer systems are superior needs clear qualification. First, people are able to employ a wide variety of "heuristics" and other knowledge-based capabilities, which may effectively compensate for a lack of computational power [38, 51]. Second, the difficulties people experience with memory retrieval does not imply that computer-based storage systems are superior. Despite problems of forgetting, distortion, etc. human memory appears better adapted to maintaining and accessing prodigious databases of facts than existing computer systems [39].

2.b.2 Inefficient and suboptimal representations

It can be argued that people make suboptimal decisions because they are incapable of using certain types of optimal representations of problems. This assertion has yet to be demonstrated however [17]. Neither has it been shown that computer programs can invariably support more satisfactory problem representations.

A related and more plausible argument is that computer programs can represent parts of an expert's knowledge in a highly compact form and more efficiently than a human expert. A familiar knowledge engineering practice is to elicit knowledge from an expert in modular chunks (facts, rules, etc.), and at a later stage compile these into a more efficient representation such as a decision tree or network. This has lead some researchers [e.g. 14] to opt for efficiency rather than fidelity to an expert's representation in expert systems design. They recommend translation to and from a more natural representation purely for the users' benefit.

2.b.3 Improving on expert performance

Human experts are known to be inconsistent, unreliable and to disagree with their colleagues on important matters [19, 28]. Such observations suggest that a reasonable goal for expert system design is not merely the achievement of expert level performance but, ultimately, an improvement on human expertise. Results presented by Michalski and Chilausky [36], where machine-induced rules proved better at diagnosing soybean diseases than rules derived from an expert, indicate that this is already the case for some simple types of task. Logically, it is difficult to see how the objective of improving on human performance could be achieved in an expert system modelled purely on an expert's thinking, weaknesses as well as strengths.

2.b.4 Multiple expert systems

There is often a requirement to embody the expertise of specialists from different areas within a single expert program. This requirement is reflected, for example, in the "blackboard" architecture for expert systems.

Problem-solving in a blackboard system involves a set of independent "specialists" or knowledges sources co-operating in developing and checking possible problem solutions. Communication between the specialists is via a shared working memory: the blackboard. Even in the case where only the expertise of one particular type of domain specialist is required it is considered that the expert system should be capable of embodying the expert skill of several individuals [28].

This "multiple expert" requirement argues against cognitive emulation in its strongest form. Individual variation is a salient characteristic of human cognition (see e.g. [56]). Even within a particular speciality different experts can be expected to employ idiosyncratic reasoning strategies and knowledge representations [52]. However, while general principles underlying human thinking are discernable, the cognitive processes of different individuals cannot be combined, or averaged, very meaningfully.

2.c Conclusions

The greater power and accuracy of formal computer systems argues against an unselective strategy of cognitive emulation. However, it is not a conclusive argument against a more selective strategy which exploits the compensatory strengths of human cognition. Taken together the force of these arguments is that neither extreme attitude to cognitive emulation is tenable.

A pure, unselective application of the strategy is untenable because:

- it would entail a commitment to emulating human cognitive weakness [2.b.1]
- experts' representations of problems may be sub-optimal or inefficient [2.b.2]
- it implies that expert systems cannot aim to improve on human expert performance [2.b.3]
- some expert systems need to integrate the expertise of several domain specialists [2.b.4]

Equally, the opposite view that cognitive emulation should be avoided or ignored as a design strategy can also be rejected because:

- cognitive emulation seems inherent to knowledge engineering [2.a.1]
- a plausible strategy for achieving expert-level performance is to emulate the underlying cognitive processes [2.a.2]

- cognitive emulation can promote user acceptance [2.a.3]
- effective knowledge elicitation often requires psychologically-orientated techniques [2.a.4]

A more pragmatic attitude is called for, one that acknowledges both the potential usefulness and the limitations of cognitive emulation for knowledge engineering. The next two sections of the paper consider factors which are likely to further constrain or facilitate this approach.

3 Constraints on cognitive emulation

3.a The emulability of human expertise in artificial systems

It is not yet clear what limits there are (if any) on modelling the cognitive processes underlying expert behaviour in an artificial system. This issue is an instance of the wider debate about the comparative nature of artificial and human intelligence (e.g. [3]).

An influential development in this debate has been the notion of a "Physical Symbol System" proposed by Newell and Simon [5]. This defines a broad class of system capable of manipulating symbols or, more generally, symbolic structures, that are realisable as physical entities. Newell and Simon's central hypothesis is that Physical Symbol Systems have all the necessary and sufficient means for intelligent action. In their view, human beings and computers are prime instances of such systems. If this meta-theory is correct then in principle there should be no limitation on cognitive emulation in expert systems.

On the other hand it may be, as some critics of AI contend, that in essential respects human cognitive processes cannot be adequately modelled by a computer program [7, 10]. If these critics prove to be right then the prospects for an emulation approach are poor.

It seems doubtful whether this fundamental issue can be resolved to universal satisfaction solely on a priori arguments. Any "conclusive" argument is liable to be overthrown by empirical developments in cognitive psychology or technical ones in expert systems. An alternative intermediate view is that some aspects of human cognition can be emulated easily, others less easily. For example, cognitive scientists have found it relatively straight forward to simulate the acquisition of arithmetic skills using rule-based techniques (e.g. [69]). By contrast, reasoning by analogy and the representation of spatial concepts have proved difficult to model computationally [20, 59]. However, it is too early to tell whether aspects of expert thinking which are currently difficult to emulate can or cannot be dealt with in principle.

3.b The state of cognitive psychology

A source of difficulty for system builders adopting a strategy of emulation stems from the current state of development of Cognitive Psychology. Difficulties for a modelling approach include:

1 Ignorance. Despite undoubted progress in Cognitive Psychology many key questions about human cognitive processes have yet to be tackled, let alone answered to general satisfaction. Norman [58] maps out many of these gaps in our present knowledge of the human mind. Problem: One cannot emulate what is not known.
2 Diversity of approaches. In Kuhnian terms psychology is an immature science, in a "pre-paradigmatic" stage of development [3]. This is reflected in the varied range of theories, models, knowledge representation formalisms etc. in the cognitive literature which are often offered as explanations of the same phenomena. Problem: Which theory/model/formalism does one choose to emulate?
3 Scientific status. As in any science theories in Cognitive psychology are subject to constant revision, obsolescence and falsification by empirical research findings [6]. That is, our knowledge of human cognition is provisional in nature. Problem: The cognitive model emulated today may be refuted by later research.

3.c The state of expert systems technology

The existing technology for building expert systems—both hardware (e.g. Von Neumann architectures) and software (languages, shells and other system building tools)—is unequal to the task of modelling some influential and potentially useful ideas in Cognitive Psychology. (The tutorial article by Hayes-Roth [20] provides a clear summary of the state of the art in Knowledge Engineering, and what notions it can handle.)

One example is the "spreading activation" theory of human memory [e.g. 37, 38]. In this theory memory is represented as a network of concept nodes. When one or more nodes become active (perhaps due to a sensory input or memory search) activation spreads out in parallel to the nearest nodes, forming an expanding sphere of activation around the original one(s). Roughly speaking, activation decreases the further one gets from the initial source(s) of activation. Anderson [38] argues that activation can function as a "relevancy heuristic", on the reasonable assumption that knowledge which is closely associated with ideas which are being processed has a good chance of being relevant to the current problem. However, spreading activation is essentially a parallel mechanism requiring parallel machinery for its efficient implementation. "Until such machinery becomes available, this potentially good idea will not see extensive use in [AI] applications" [38, p. 88]. This restriction applies equally to other cognitive mechanisms that are believed to be parallel [see Table 2].

On the software side, Hayes-Roth [20] mentions analogies, meta-knowledge, naive physics and first principles as types of human knowledge not yet available in state-of-the-art expert systems. Table 2 gives other examples.

3.d Individual variation in expert thinking

Cognitive psychologists try to identify general principles of human thinking. Individual variation in cognitive performance is accounted for by these principles. A strategy of cognitive emulation could be applied to expert system design based on general principles of human (expert) cognition [cf. 34, 70], but it has its limitations.

The research which suggests generalisations about the human mind is usually based on experimental results gathered from many subjects and analyzed using standard statistical techniques. In computer simulations of mental processes psychologists may therefore model a typical or average subject [13]. This kind of application of general principles may fail to do proper justice to the known richness and variety of reasoning strategies, knowledge representations, etc. of individual experts [52, 56]. The problem-solving capability of an expert program might suffer accordingly.

An alternative methodology is to model a system on the thinking of particular expert(s) using suitable techniques to analyse that thinking (e.g. protocol analysis). Many knowledge engineers would claim to be doing this already [28].

Perhaps the two approaches can be combined? That is to say, acquire knowledge from individual experts using standard knowledge acquisition techniques like interviewing and protocol analysis, and then tailor knowledge formalisms and inferencing methods (selected from a "toolkit" of psychologically plausible alternatives) to reflect the distinctive cognitive processes of individual experts. Unfortunately, expert systems technology cannot support such an approach at present [see section 3.c]

3.e Knowledge engineering objectives

The final group of restrictions comes from the apparent conflict between cognitive emulation and other knowledge engineering requirements. Textbooks [1, 2, 21] and review articles [11, 14, 16] discuss various system design criteria. For present purposes only five are distinguished:

1 Efficiency. To minimise costs and response times, knowledges engineers aim to build computer systems that are maximally efficient in their use of computational resources, i.e. memory, time.

2 Modifiability. It is highly desirable that the knowledge embodied in an expert system be easy to add, delete and amend.
3 Simplicity. By keeping the techniques as simple as possible, the resulting system will be easier to develop and maintain, and cost less.
4 Understandability. Understandability facilitates all stages of expert system construction and helps promote user acceptance.
5 Correctness. Provably correct conclusions, consistency and completeness in the knowledge base are desirable. Users may express more confidence in the decisions of a system whose judgments are perceived as rational.

Table 1 suggests how these five design criteria, together with the constraints on emulation discussed earlier, favour design features commonly found in expert systems. In Table 2 the same design features are related to psychological research on corresponding aspects of human (expert) cognition. It is apparent from Table 2 that many typical expert system features are implausible models of how human experts think.

4 Factors facilitating cognitive emulation

4.a Developments in expert system technology

In section 3.3 the current state of expert system technology was identified as an important constraint on possibilities of cognitive emulation. However, the technology is still in an early phase of development [12] and the world-wide investment in fifth-generation research projects may remove some of the technological constraints on modelling human cognition. The development of parallel hardware is an obvious example. Research and development in software may lead to multi-representation systems, model-based reasoning, meta-level reasoning, meta-knowledge systems and learning by example [20].

4.b Cognitive research

Cognitive Psychology research can assist the refinement of the emulation strategy in some ways:

1 It has "heuristic value". The psychological literature on human thinking can serve as a useful source of ideas. For example, there is a degree of consensus amongst cognitive psychologists regarding certain broad principles of human information processing (some of these are referred to in Table 2) and human expertise. Domains investigated include algebra (82), geometry [71], physics [77, 80, 81], medicine [48, 84], programming [47, 70, 72], and chess and other board games [76, 78]. Many of these studies focus on the cognitive changes accompanying a person's "shift" from novice to expert status.
2 The advent of Cognitive Science. Some researchers in Cognitive Psychology and related disciplines (AI, philosophy, linguistics, etc.) now identify themselves as Cognitive Scientists. Amongst other things, this label denotes a commitment to implementing and testing cognitive ideas as computational models [66]. This requirement should make the task of emulating the capabilities of human experts easier.

4.c Cognitive emulation can coincide with knowledge engineering objectives

In Table 1 and 2 the emphasis is on areas of conflict. Nevertheless cognitive emulation criteria will sometimes coincide rather than conflict with other systems design criteria:

- Knowledge-based and rule-based techniques [see 2.a.1].

Table 1 Knowledge engineering factors favouring typical expert system features.

Expert system feature	*Knowledge engineering factor(s)*	*Comment*
1. Uniform Representation of Knowledge	Simplicity	Each formalism requires its own inferencing mechanism
	Efficiency	Computational resources needed to store and communicate between many formalisms
	Technology	Availability of uniform implementation languages e.g. PROLOG
2. Modular Representation of Knowledge	Modifiability	Enables knowledge to be added, deleted and amended with minimal effects on knowledge in system
3. Natural Representation of Knowledge	Understandability	A declarative formalism such as rules can make the individual items of knowledge easier to understand
4. Unrestricted Working Memory Capacity	Correctness	Potentially valuable facts, etc. could be lost maintaining limit on store size
	Understandability	Imposing a WM restriction might appear perverse or irrational to users
5. Boolean Representation of Patterns	Simplicity	Imposes constrained format
	Efficiency	Boolean algebra very efficiently processed by digital computers
	Correctness	Guaranteed "correct" reasoning using Boolean algebra
	Technology	Psychological nature of pattern representation poorly understood
6. Complete Pattern Matching	Simplicity	A more complex interpreter required to handle partial matching
	Correctness	Full matching requirement facilitates accuracy
7. Probabilistic Handling of Uncertainty	Correctness	Probabilistic methods offer the promise of accuracy (but see e.g. [29]).
	Technology	A range of methods is available (e.g. Bayes Thereom, fuzzy set theory).
8. Undirectional Control Structure	Simplicity	e.g. backward-chaining nature of PROLOG
	Technology	The flexible intermixture of forward/ backward reasoning characteristic of human thinking difficult to emulate
9. Serial Processing	Technology	Parallel architectures still at the prototype stage
10. Direct Manipulation of Knowledge	Modifiability	Ability to directly add, delete and alter units of knowledge required for ease of system development/maintenance
11. Static Knowledge Base	Technology	Absence of technique for dynamic changes to knowledge base whilst in use
12. Reasoning-Based Explanation	Correctness	Accurate picture of how conclusions were reached
	Technology	Easy to implement e.g. rule-tracing

Table 2 Typical expert system features and cognitive research findings.

Expert system feature	*Cognitive research findings*	*Selected references*
1. Uniform Representation of Knowledge	Expert knowledge of many distinct types—indicating multiple representations. Basic distinction between declarative[1] and procedural[2] representation often made. Evidence for specialised representations of visual and spatial knowledge.	[18, 38, 59, 81]
2. Modular Representation of Knowledge	Expertise in a domain characterised by integration of knowledge into larger units; chunks[3], scripts[4], themes[5], rules, etc. and by high interconnectivity between facts.	[62, 71, 76, 86]
3. Natural Representation of Knowledge	Human expertise seems to rely heavily on procedurally-embedded or "compiled" knowledge. Knowledge underlying cognitive performance (including experts') often not verbalisable.	[64, 71, 74]
4. Unrestricted Working Memory Capacity	Experts have larger WMs for domain knowledge than novices. Nevertheless, experts' WMs still severely limited in number and complexity of items that can be held—imposing constraints on human information processing capacity.	[15, 54, 56, 75]
5. Boolean Representation of Patterns	Generally, patterns and concepts do not appear to be represented in cognition as simple predicates linked in arbitrary arrangements by the Boolean operators "or", "and" and "not" e.g. negation rarely used, "integral" [6] and structural patterns difficult to represent.	[42, 45, 50]
6. Complete Pattern Matching	Humans adept at coping with missing or incorrect information in pattern matching e.g. classifying novel instances on the basis of their similarity to other category instances or category prototype[7].	[42, 44, 61]
7. Probabilistic Handling of Uncertainty	People, including experts, appear to rely heavily on nonstatistical heuristics e.g. availability[8] and representativeness[9] in making judgments under uncertainty. Human experts influenced by more qualitative (e.g. configurational) aspects of data.	[34, 51, 79, 83]
8. Unidirectional Control Structure	Compared with novices, the performance of human experts is often more dependent on data-driven processes e.g. pattern recognition. But generally expert thinking seems characterised by a flexible intermixture of forward and backward reasoning.	[38, 65, 77, 81]

Table 2 (*cont.*)

Expert system feature	*Cognitive research findings*	*Selected references*
9. Serial Processing	"Automated" cognitive processes[10] such as pattern recognition, memory retrieval and other more specialized skills, such as reading, appear essentially parallel in nature.	[53, 55, 65]
10. Direct Manipulation of Knowledge	Research suggests procedural knowledge first represented declaratively as facts, and only becomes proceduralized through repeated use or practice. Once "compiled", knowledge not readily deleted or modified.	[38, 71, 80]
11. Static Knowledge Base	Human knowledge is dynamic in that it has the ability to recode, restructure and generally modify itself	[43, 63, 77]
12. Reasoning-Based Explanation	Cognitive processes underlying expert performance often difficult to explain. Instead, experts may base their explanations on causal reasoning using domain models and principles, or even post-hoc rationalisations.	[57, 67, 74]

- Knowing the appropriate rule size for human reasoning can also help in knowledge elicitation from the expert [28].
- Providing an expert system with a sophisticated natural language interface may require the interface having access to the expert system's knowledge and operations, implying that these must be represented in an appropriate (humanlike?) form [26].
- Well-known cognitive principles are already being applied to "humanising" the man-machine interface of expert systems (see e.g. [28]). An awareness of the limitations on people's capacity to process information underlies Donald Michie's [24, 25] recommendation that expert systems be designed to include a "human window", i.e. employ a conceptualisation of knowledge and inferencing techniques that can be both understood and executed by a human user.

4.d Problems which are not amenable to formal methods

The intractability of some problems to formal techniques, could suggest more psychologically-orientated approaches.

To date, the limited range, complexity and size of tasks tackled by expert systems has enabled expert level performance to be achieved without significantly compromising design criteria like correctness, efficiency, etc. For example, the technology has proved sufficient for handling such "analytic" tasks as medical diagnosis and geological classification especially in small, well-defined domains. (Stefik et al. [27] have suggested a set of architectural prescriptions for building expert systems of this type.)

However, many practitioners [11, 14, 16, 20] suggest that current techniques may not prove adequate for much larger systems (knowledge bases containing several million facts), more varied tasks (e.g. "synthetic" tasks like design) and more complex capabilities (e.g. learning from experience) predicted for future applications. Since the human mind has proven equal to coping with such tasks, whereas formal AI techniques have not as yet, a cognitive approach is an obvious strategy.

Broadly speaking, the utility of the cognitive emulation principle might be expected to increase as the tasks tackled by expert system builders become more difficult.

IF	1.	Existing (formal) techniques cannot meet all major knowledge engineering requirements for an application satisfactorily
AND	2.	Cognitive process(es) can be identified that appear relevant to meeting the application requirements
AND	3.	Cognitive process(es) involved considered emulable in principle
AND	4.	Cognitive process(es) involved psychologically understood
AND	5.	Cognitive process(es) involved available as a computational model
AND	6.1	Knowledge engineering tools embodying cognitive process(es) commercially available
OR	6.2	Resources available to develop knowledge engineering tools embodying cognitive process(es)
AND	7.	Knowledge engineering tools embodying (generalised) cognitive process(es) can be tailored to expressively accommodate knowledge, elicited from individual domain experts
AND	8.1	Emulating cognitive process(es) DOES NOT significantly compromise other knowledge engineering objectives (e.g. efficiency)
OR	8.2.1	Emulating cognitive process(es) DOES significantly compromise other knowledge engineering objectives (e.g. efficiency)
AND	8.2.2	Problem unsolvable using alternative techniques
THEN		Cognitive process(es) involved should be emulated in an expert system

Figure 1 A decision rule for adopting a strategy of cognitive emulation in expert system design

5 Summary and conclusions

Cognitive emulation is a strategy for expert system design that can neither be explicitly rejected nor unselectively pursued. A significant element of cognitive modelling seems inherent in Knowledge Engineering, and is desirable in order to promote expert-level performance, user acceptance and effective knowledge acquisition. Against a pure emulation strategy are the known weaknesses, limitations and inefficiencies of human (expert) thinking, the desire to improve on expert-level performance, and the requirement that an expert system should be able to embody the specialist knowledge of several experts.

Constraints on cognitive emulation must be recognized. Some cognitive processes may not be amenable to emulation. Others may be emulable in principle, but not with available technology. Our current understanding of human cognition is incomplete, with many competing explanations and the prospect of established notions being falsified. Differences in expert's strategies may not map comfortably onto general principles derived from cognitive research. Cognitive emulation may conflict with established Knowledge Engineering objectives such as efficiency and modifiability.

Factors considered likely to facilitate a cognitive approach include: future developments in expert systems technology and the cognitive sciences, situations where cognitive emulation supports other design criteria, and the need to tackle problems which are intractable using formal methods.

Whether it is feasible to explicitly adopt a strategy of cognitive emulation will depend on the balance of facilitating and constraining factors operating in any given instance. Figure 1 summarises the points raised in this paper. Applying this rule, the state of the enabling technologies and disciplines would presently seem to weigh against all but a highly selective pursuit of cognitive emulation for commercial applications. In the longer term, however the rule implies that a combination of developments in expert system technology and Cognitive Science, the need to tackle larger and more difficult problems, the desire to further humanise the user interface, etc. will make a strategy of cognitive emulation increasingly attractive—if not essential.

Terms

1. DECLARATIVE KNOWLEDGE Most of the knowledge comprises a static collection of facts accompanied by a small set of general procedures for manipulating them.

2. PROCEDURAL KNOWLEDGE In this representation the bulk of the knowledge is contextually-embedded in procedures, e.g. computer algorithms.

3. CHUNKS "Chunks are learned configurations of symbols which come to act as a single

symbol" [37, p 97]. Classically a chunk occupies one of (about) seven available positions in short-term or working memory [54].

4. SCRIPTS A script is a frame-like structure describing an appropriate sequence of events in a particular context.

5. THEMES A term used to denote a thematically-integrated set of facts in declarative memory [38, 62].

6. INTEGRAL PATTERNS A class of stimuli that are not analysed into independent attributes (e.g. colour, size) in cognition, but are processed in a holistic manner [45].

7. CATEGORY PROTOTYPE The best or clearest example of a category. For example, a robin probably conforms to most people's idea of a typical "bird", whereas a vulture does not (see e.g. [61].

8. AVAILABILITY This term refers to the tendency of judging the likelihood of an event by how readily related information in memory can be retrieved. This can distort probability estimation as when someone overestimates the likelihood of being struck by lightning after reading a particularly vivid newspaper report on the subject [51].

9. REPRESENTATIVENESS This term is used to describe a human strategy for judgement under uncertainty where the likelihood of an item belonging to a particular class is based on how representative or typical an example of the class the item is. So, for example, a shy young woman might be judged more likely to be a librarian than a shop assistant, despite the fact that there are far more shop assistants and librarians in the population at large [51].

10. AUTOMATED COGNITIVE PROCESSES In contrast to "controlled processes", automatic processes: are unconstrained by the capacity limitations of working memory; are often parallel in nature; run to completion automatically once initiated; require considerable training to develop, being difficult to modify once learned; are often unavailable to introspection (see [65]).

Acknowledgement

I would like to thank John Fox for his help and encouragement in preparing this paper.

Suggested reading

Anderson, J.R. (1984) Cognitive Psychology. *Artificial Intelligence*, 23, 1–11.
Barr, A and E.A. Feigenbaum (Eds.) (1981/2) *The Handbook of Artificial Intelligence*, vols. 1 and 2. Los Altos, Calif: William Kaufman.
Boden, M.A. (1977) *Artificial Intelligence and Natural Man*. Brighton: Harvester.
Davis, R. and King, J. (1977) An Overview of Production Systems. In E.W. Elcock and D. Michie (Eds.), *Machine Intelligence*, vol. 8. Chichester: Ellis Horwood.
Norman, D.A. (Ed.) (1981) *Perspectives on Cognitive Science*. Hillsdale, NJ: Erlbaum.
Welbank, M. (1983) A Review of Knowledge Acquisition Techniques for Expert Systems. Martlesham Consultancy Services, British Telecoms Research Labs., Martlesham Heath, Ipswich.

References

1 Artificial Intelligence and philosophy

[1] Barr, A. and E.A. Feigenbaum (Eds.) (1981) *The Handbook of Artificial Intelligence*, vol. 1. Los Altos, Calif.: William Kaufman.
[2] Barr, A. and E.A. Feigenbaum (Eds.) (1982) *The Handbook of Artificial Intelligence*, vol. 2. Los Altos, Calif.: William Kaufman.
[3] Boden, M.A. (1977) *Artificial Intelligence and Natural Man*. Brighton: Harvester.

[4] Minsky, M. (1975) A Framework for Representing Knowledge, In P.H. Winston (Ed.), *The Psychology of Computer Vision*. New York: McGraw Hill.
[5] Newell, A. and Simon, H.A. (1976) Computer Science as Empirical Enquiry: Symbols and Search. (1975 ACM Turing Lecture) *Communications of the ACM*, 19 (3) 113–126.
[6] Popper, K.R. (1963) *Conjectures and Refutations*, London: Routledge.
[7] Searle, J. (1984) *Brains, Minds and Science*. Reith Lectures, BBC publications.
[8] Simons G.L. (1983) Towards Fifth-Generation Computers. Manchester: NCC Publications.
[9] Sloman, A. (1984) Why We Need Many Knowledge Representation Formalisms. In M.A. Bramer (Ed.), *Research and Development in Expert Systems*. Cambridge: Cambridge University Press.
[10] Weizenbaum, J. (1976) *Computer Power and Human Reason: From Judgement to Calculation*. San Francisco: Freeman.

2 Expert Systems: General

[11] Buchanan, B.G. (1982) New Research on Expert Systems. In J.E. Hayes, D. Michie and Y.H. Pao (Eds.), *Machine Intelligence 10*. Chichester: Ellis Horwood.
[12] Bramer, M.A. (1984) Expert Systems: The Vision and the Reality, In M.A. Bramer (Ed.), *Research and Development in Expert Systems*. Cambridge: Cambridge University Press.
[13] Breuker, J.A. and Wielinga, R.A. (1983) Analysis Techniques for Knowledge Based Systems Part 1. Report 1.1 Esprit Project 12 (Memorandum 10 of the Research Project "The Acquisition of Expertise").
[14] Davis, R. (1982) Expert Systems: Where are we? And where do we go from here? *The AI Magazine*, 3, 3–22.
[15] Davis, R. and King. J. (1977) An Overview of Production Systems. In E. Elcock and D. Michie (Eds.), *Machine Intelligence*, vol. 8. Chichester: Ellis Horwood.
[16] Duda, R.O. and Shortliffe, E.H. (1983) Expert Systems Research, *Science*, 220, 261–268.
[17] Fox, J. (1982) Expertise in Man and Machine. Paper given at the Colloquium on "Applications of Knowledge-based (or Expert) Systems" Institute of Electrical Engineers, Savoy Place, London.
[18] Gammack, J.G. and Young, R.M. (1984) Psychological Techniques for Eliciting Knowledge. In M.A. Bramer (Ed.), *Research and Development in Expert Systems*. Cambridge: Cambridge University Press.
[19] Gaschnig, J., Klahr, P., Pople, H., Shortliffe, E. and Terry, A. (1983) Evaluation of Expert Systems; Issues and Case studies. In F. Hayes-Roth, D.A. Waterman, D.B. Lenat (Eds.), *Building Expert Systems*. Reading, MA: Addison-Wesley.
[20] Hayes-Roth, F. (1984) The Knowledge-based Expert System: A Tutorial, *Computer*, September 1984, 11–28.
[21] Hayes-Roth, F., Waterman, D.A., and Lenat, D.B. (Eds.) (1983) *Building Expert Systems*. Reading, Ma: Addison-Wesley.
[22] Johnson, T. (1984) The Commercial Application of Expert System Technology. *Knowledge Engineering Review*, 1(1), 15–25.
[23] Leith, P. (1983) Hierarchically Structured Production Rules. *Computer Journal*. 26, 1–5.
[24] Michie, D. (1980) Expert Systems. *Computer Journal*, 23, 369–376.
[25] Michie, D. (1982) The State of the Art in Machine Learning. In D. Michie (Ed.), *Introductory Readings in Expert Systems*. London: Gordon and Breach.
[26] Sparck Jones, K. (1984) Natural Language Interfaces for Expert System In M.A. Bramer (Ed.), *Research and Development in Expert Systems*. Cambridge: Cambridge University Press.
[27] Stefik, M., Aikins, J., Balzer, R., Benoit, J., Birnbaum, L., Hayes-Roth, F. and Sacerdoti, E. (1982) The Organization of Expert Systems: A Tutorial. *Artificial Intelligence*, 18, 135–173.
[28] Welbank, M. (1983) A Review of Knowledge Acquisition Techniques for Expert Systems. Martlesham Consultancy Services, British Telecom Research Labs: Martlesham Heath, Ipswich.
[29] White, A.P. (1984) Inference Deficiences in Rule-based Expert Systems. In M.A. Bramer (Ed.), *Research and Development in Expert Systems*. Cambridge: Cambridge University Press.
[30] Young, R.M. (1985) Human Interface Aspects of Expert Systems. In J. Fox (Ed.), *State of the Art Report in Expert Systems*. Pergamon Infotech, 1984.

3 Expert Systems: Specific Applications

[31] Buchanan, B.G. and Feigenbaum E.A. (1978) DENDRAL and Meta-DENDRAL; Their Application Dimension. *Artificial Intelligence*, 11, 5–24.
[32] Davies, R., Buchanan, B.G. and Shortliffe, E. (1977) Production Rules as a Representation for a Knowledge-based Consultation Program. *Artificial Intelligence*, 8, 15–45.
[33] Duda, R.O., Gaschnig, J.G. and Hart, R.E. (1979) Model Design in the PROSPECTOR Consultant

System for Mineral Exploration. In D. Michie (Ed.) *Expert Systems in the Micro-electronic Age.* Edinburgh: Edinburgh Press.

[34] Fox, J., Barber, D. and Bardhan, K.D. (1980) Alternatives to Bayes? A Quantitative Comparison with Rule-based Diagnosis. *Methods of Information in Medicine*, 19, 210–215.

[35] McDermott, J. (1982) R1: A Rule-Based Configurer of Computer Systems. *Artificial Intelligence*, 19, 39–88.

[36] Michalski, R.S. and Chilausky, R.L. (1980) Knowledge Acquisition by Encoding Expert Rules Versus Computer Induction From Examples: A Case Study Involving Soybean Pathology. *International Journal of Man-Machine Studies*, 12, 63–87.

4 Cognitive Research: General

[37] Anderson, J.R. (1976) *Language, Memory and Thought.* Hillsdale, NJ: Erlbaum.

[38] Anderson, J.R. (1983) *The Architecture of Cognition.* Cambridge, MA: Harward.

[39] Anderson, J.R. (1984) Cognitive Psychology. *Artificial Intelligence*, 23, 1–11.

[40] Anderson, J.R. and Bower, G.H. (1983) *Human Associative Memory.* Washington, D.C.: Winston.

[41] Baddeley, A.D. (1976) *The Psychology of Memory.* New York: Harper and Row.

[42] Barslou, L.W. and Bower, G.H. (1984) Discrimination Nets as Psychological Models, *Cognitive Science*, 8, 1–26.

[43] Bartlett, F.C. (1932) *Remembering: A Study in Experimental and Social Psychology.* Cambridge, Cambridge University Press.

[44] Elio, R. and Anderson, J.R. (1981) Effects of Category Generalizations and Instance Similarity on Schema Abstraction. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 397–417.

[45] Garner, W.R. Interaction of Stimulus Dimensions in Concept and Choice Processes. *Cognitive Psychology*, 1976, 8, 98–123.

[46] Hunt, E.B., Marin, J. and Stone, P.J. (1986) *Experiments in Induction.* New York; Academic Press.

[47] Jeffries, R., Turner, A., Polson, P. and Atwood, M. (1981) The Processes Involved in Designing Software. In J.R. Anderson (Ed.), *Cognitive Skills and Their Acquisition.* Hillside, NJ: Erlbaum.

[48] Johnson, P.E., Duran, A.S., Hassebrock, F., Moller, J., Prietula, M., Feltovich, P.J., and Swanson, D.B. (1981) Expertise and Error in Diagnostic Reasoning. *Cognitive Science*, 5, 235–283.

[49] Johnson-Laird, P.N. (1982) Thinking as a Skill. *Quarterly Journal of Experimental Psychology*, 34A, 1–29.

[50] Johnson-Laird, P.N. and Wason, P.C. (Eds.) (1977) *Thinking: Readings in Cognitive Science.* Cambridge: Cambridge University Press.

[51] Kahneman, D., Slovic, P. and Tversky, A. (Eds.) (1982) *Judgment under Uncertainty: Heuristics and Biases.* Cambridge: Cambridge University Press.

[52] Kuipers, B. and Kassirer, J.P. (1984) Causal Reasoning in Medicine: Analysis of a Protocol. *Cognitive Science*, 8, 363–385.

[53] McCelland, J.L. and Rumelhart, D.E. (1981) An Interactive Model of Contextual Effects in Letter Perception: Pt 1, An Account of Basic Findings. *Psychological Review*, 88, 375–407.

[54] Miller, G.A. (1956) The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity to Process Information. *Psychological Review*, 63, 81–97.

[55] Neisser, U. (1967) *Cognitive Psychology.* New York: Appleton.

[56] Newell, A. and Simon, H.A. (1972) *Human Problem solving.* Englewood Cliffs, NJ: Prentice-Hall.

[57] Nisbett, R.E. and Wilson, T.D. (1977) Telling More Than We Can Know: Verbal Reports on Mental Processes. *Psychological Review*, 84, 231–259.

[58] Norman, D.A. (1981) Twelve Issues for Cognitive Science. In D.A. Norman (Ed.), *Perspectives on Cognitive Science*, Hillside, NJ: Erlbaum.

[59] Pinker, S. (1984) Visual Cognition: An Introduction, *Cognition*, 18, 1–63.

[60] Pople, H. (1977) The Formation of Composite Hypotheses in Diagnostic Problem Solving—An Exercise in Synthetic Reasoning. *International Joint Conferences on AI 5*, 1030–1037.

[61] Posner, M.I. and Keele, W.W. (1970) Retention of Abstract Ideas. *Journal of Experimental Psychology*, 83, 304–308.

[62] Reder, L.M. and Anderson, J.R. (1980) A Partial Resolution of the Paradox of Interference: The Role of Integrating Knowledge. *Cognitive Psychology*, 12, 447–472.

[63] Rumelhart, D.E. and Norman, D.A. (1978) Accretion, Tuning and Restructuring: Three Modes of Learning, In J.W. Cotton and R. Klatzky (Eds.), *Semantic Factors in Cognition.* Hillsdale, NJ: Erlbaum.

[64] Rumelhart, D.E. and Norman, D.A. (1981) Analogical Processes in Learning. In J.R. Anderson (Ed.), *Cognitive Skills and Their Acquisition.* Hillsdale, NJ: Erlbaum.

[65] Shiffrin, R.M. and Schneider, W. (1977) Controlled and Automatic Human Information Processing. II Perceptual Learning, Automatic Attending and a General Theory. *Psychological Review*, 84, 127–190.

[66] Slack, J.M. (1984) Cognitive Science Research. In T. O'Shea and M. Eisenstadt (Eds.), *Artificial Intelligence: Tools, Techniques, and Applications.* New York: Harper and Row.
[67] Wason, P.C. and Evans, J. St B.T. (1975) Dual Processes in Reasoning? *Cognition*, 3, 141–154.
[68] Young, R.M. (1979) Production Systems for Modelling Human Cognition. In D. Michie (Ed.), *Expert Systems in The Micro-electronic Age.* Edinburgh: Edinburgh University Press.
[69] Young, R.M. and O'Shea T. (1981) Errors in Children's Subtraction. *Cognitive Science*, 5, 153–177.

5 Cognitive Research: Studies of Expertise

[70] Adelson, B. (1984) When Novices Surpass Experts: The Difficulty of a Task May Increase With Expertise. *Journal of Experimental Psychology: Learning, memory and Cognition*, 10, 483–495.
[71] Anderson, J.R. (1983) Acquisition of Proof Skills in Geometry. In J.G. Carbonell, R. Michalski and T. Mitchell (Eds.), *Machine Learning, An Artificial Intelligence Approach*, San Francisco: Tioga.
[72] Anderson, J.R., Farrell, R. and Sauers, R. (1984) Learning to Program in Lisp. *Cognitive Science*, 8, 87–129.
[73] Arkes, H.R. and Freedman, M.R. (1984) A Demonstration of the Cost and Benefits of Memory. *Memory and Cognition*, 12, 84–89.
[74] Berry, D.C. and Broadbent, D.E. (1984) On the Relationship between Task Performance and Associated Verbalizable Knowledge. *Quarterly Journal of Experimental Psychology*, 36A, 209–231.
[75] Chase, W.G. and Ericsson, K.A. (1982) Skill and Working Memory. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation*, vol. 16. New York: Academic Press.
[76] Chase, W.G. and Simon, H.A. (1973) Perception in Chess. *Cognitive Psychology*, 4, 55–81.
[77] Chi, M.T.H., Feltovich, P.J. and Glaser, R. (1981) Categorization and Representation of Physics problems by Experts and Novices. *Cognitive Science*, 5, 121–152.
[78] Eisenstadt, M. and Kareev, Y. (1975) Aspects of Human Problem Solving: The Use of Internal Representations. In D.A. Norman and D.E. Rummelhart (Eds.), *Explorations of Cognition.* San Francisco: Freeman.
[79] Fox, J. (1980) Making Decisions Under the Influence of Memory. *Psychological Review*, 87, 190–211.
[80] Larkin, J.H. (1981) Enriching Formal Knowledge: A Model for Learning to Solve Textbook Problems. In J.R. Anderson (Ed.), *Cognitive Skills and Their Acquisition.* Hillsdale, NJ: Erlbaum.
[81] Larkin, J.H., McDermott, J., Simon, D.P. and Simon, H.A. (1980) Expert and Novice Performance in Solving Physics Problems. *Science*, 208, 1335–1242.
[82] Lewis, C. (1981) Skill in Algebra, In J.R. Anderson (Ed.), *Cognitive Skills and Their Acquisition.* Hillsdale, NJ: Erlbaum.
[83] Medin, D.L., Altom, M.W., Edelson, S.M. and Freko, D. (1982) Correlated Symptoms and Simulated Medical Classification. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 8, 37–50.
[84] Murphy, G.L. and Wright, J.C. (1984) Changes in Conceptual Structure with Expertise: Differences Between Real-World Experts and Novices. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 144–155.
[85] Simon, H.A. and Gilmartin, K. (1973) A Simulation of Memory for Chess Positions. *Cognitive Psychology*, 5, 29–46.
[86] Smith, E.E., Adams, N. and Schorr, D. (1978) Fact Retrieval and the Paradox of Interference. *Cognitive Psychology*, 10, 438–464.