

An evaluation of ACT-R and the rules of it's mind

Introduction to Cognitive Modeling, 1999/2000 - Assignment 2

by: Lucas Dixon

tutor: Dan Yarlett

In this short paper I shall evaluate the ACT-R theory of cognitive processes, as described by Anderson, in his paper, "Rules of the mind" 1993, with respect to the theoretical and practical implications. ACT-R is a theory in aid of understanding human cognition. Anderson proposes that Production Rules are the realization of cognitive skills, and ACT-R is the correct production system theory. Prior to any evaluation of ACT-R itself, I shall outline what ACT-R is, it's main features distinguishing it from other production systems and discuss some of the theoretical and implementational details.

a description of ACT-R

ACT-R is theory of the Production Rule Framework, similar to SOAR, and based on PSG. As such Anderson claims that it correctly captures cognitive skills using a set of rules. ACT-R is also an implementation of a production rule system. Production rule systems involves the existence of a set of production rules that are a condition-action pair, where if the given precondition is satisfied then the rule action is activated. The general operation of a production system involves pattern matching the contents of working memory to the preconditions of rules. Many rules may match the preconditions, so a second process is performed: conflict resolution, which determines which production rules are performed and which aren't. This process is termed a cycle.

ACT-R has two types of knowledge - procedural and declarative - and two types of storage/access to knowledge - working memory and long term memory. Declarative knowledge is defined by Anderson as knowledge that is accessible by introspection and represents conscious awareness of facts, where procedural knowledge is not available by introspection, and represents knowledge about performing a function. However Anderson also claims that sometimes these distinctions are broken, so what was a clear idea of types of knowledge is could be criticized as less clear and seemingly to be defined by Anderson's intuitions, and thus failing to be empirically falsifiable. For an understanding of ACT-R this is sufficient but I shall discuss in more depth the distinction drawn by the ACT-R theory between declarative and procedural knowledge.

It is the separation of long term and working(short term) memory that distinguishes ACT-R from other production rule systems and the differentiation between declarative and procedural knowledge is one of the main differences between production rule and other frameworks. The memory distinction between working memory and long term memory is a representation of the psychological notion of long and short term memory. ACT-R allows limits on the size of working memory, but the limit is not one of quantity of knowledge, nor of procedural knowledge, but of access to declarative knowledge - the notion is one of activity. ACT-R has all declarative knowledge and doesn't lose it, but at any time it has access to limited subset of that declarative knowledge. Procedural knowledge is not subject to any working memory limits, but instead an activation constraint by the storage of working memory. So it ACT-R is also theory of the "There are two memory stores: short-term and a long term" framework.

The ACT-R theory also defines a data structure for memory: the 'Chunk'. This term has been taken from psychology in particular G. A. Miller 1956 and N. F. Johnson 1970. In ACT-R a chunk is defined as '... the units in which knowledge is represented in working memory ...'. Three characteristics of chunks are then given: Firstly a limited number of components can be combined in a single chunk. Presumably the components are either declarative or procedural units of knowledge (either some quantity of declarative knowledge or a production rule). Secondly, chunks have configurable properties such that different components have different roles. Thirdly Chunks are capable of being organized into a hierarchy such that chunks can appear as components of another chunk.

ACT-R is presented by Anderson as a system with which to model skill learning, where the claim is that learning is a process of knowledge compilation, a function from declarative to procedural knowledge. This compilation has two sub processes, proceduralisation and composition.

The theory is that there are three distinguishable stages of expertise from skill learning: Interpretative, Compiled, and Automatic. Interpretative is the first stage and most prone to errors. Declarative knowledge about the task process may be incomplete or even incorrect. Interpretative procedure involves conscious manipulation of the declarative representation of the method for performing a task. As a result of this processing a task interpretatively tends to be slow and halting, and often verbalized. Mike Ramscar gives the example of driving, where the learner must consciously remember each action to perform in sequence, 1. checking the mirror... etc. During the compiled stage parts of the skill are chunked into a procedure for performing the skill. As such demands on working memory are lower and there is less errors and because less declarative knowledge is used it also becomes more difficult to verbalize the procedure. The automatic stage of expertise involves a task process being optimized. The procedure to perform becomes less conscious and less affected by a change in attention level on the task. Using Mike Ramscar's example of driving this could be interpreted as the stage where talking and driving can be done without any decrease in performance on driving. However there is no consensus on whether the the automatic stage implies a task can be performed without any attention.

In ACT-R the proceduralisation process is equivalent to the change from interpretative to compiled process of task performance. The Interpretive stage is represented in the ACT-R theory by declarative knowledge used in conjunction with hill-climbing heuristics and means end analysis to perform a tasks. After repetition of a task the proceduralisation process uses the repetition of contexts to the action taken, to create procedural knowledge in the form of "if... then ..." rules (production rules). The proceduralisation process improved performance in ACT-R by the decreasing in required active declarative knowledge. The composition process is another method of knowledge compilation that involves removal of unnecessary production rules that lead to useless search paths.

The tuning of production rules and the composition process is aimed at reflecting the process by optimization of a task. ACT-R has three methods for optimizing: strengthening- The rule becomes more likely to fire given the preconditions required; generalization- part of the if condition of a rule is replaced by a variable, making the rules applicable to a larger variety of stimulus/preconditions; Discrimination- given two rules with similar preconditions then if one rule leads to failure, it's activation is decreased (less likely to fire given the preconditions in the conflict resolution phase), where the other rule succeeds, the successful rule's activation is increased (it is strengthened).

The ACT-R theory claims that the process of skill acquisition from novice involves proceduralisation, and optimization of

the results of the proceduralisation. This is the change from hill-climbing and means end analysis to the encoding of production rules, specific to the given domain. It seems then that ACT-R also embodies the framework: “people solve problems by means end analysis”.

interpretation of and justification for ACT-R

If ACT-R does embody the framework: “people solve problems by means end analysis”, this poses a question about the nature of the means end analysis knowledge in cognitive skills, because the knowledge of means end analysis is required for the process of skill acquisition, then it must be apriori knowledge. If it wasn't then we would never learn anything as we require means end analysis for learning. If it is apriori knowledge, then philosophical arguments about apriori knowledge apply. If we assume that people can learn in other ways than by using means end analysis, then we have to question the boundaries between procedural and declarative knowledge and the functions which take declarative knowledge to procedural knowledge. Would this system have to in turn be a cognitive skill acquisition model? If so then the requirements for ACT-R are circular, and unsatisfiable.

Anderson defines the interpretation of a production rule in ACT-R as:

“Each production rule is thought of as a modular piece of knowledge in that it represents a well defined step of cognition”

The concepts of “modular piece of knowledge” and “well defined step of cognition” are not well defined at all. However assuming that they are taken at face value then if it is possible to acquire a cognitive skill without having any clear well defined step of cognition involved, then this interpretation of production rule systems would prove ACT-R as a representation of cognitive skills, false. Unfortunately Anderson hasn't attempted this, or looked at examples where this appears to be the case. From personal experience it seems intuitive that one can acquire skills without the skill being describable as a well-defined step of cognition. This then suggests that either production rule system are not a valid model of cognitive skills, the cognitive task is not in the subset of cognitive skills, or the process is in well defined steps of cognition, but we aren't aware of them.

This also poses question about the nature of declarative knowledge: If declarative knowledge is not limited to well defined concepts in cognition, then how can we have a production rule which is a well defined step of cognition, where it applies to something that is not well defined and possibly produces something that is not well defined, surly for it to be well defined it's components must be well defined? if declarative knowledge is limited to well defined concepts in cognition, then we end up with problem of grounding, are there any ground terms in cognitive processes? This would also make production rule systems very similar to a hierarchical categorization systems, where rules represent branches in the tree.

Assuming that this interpretation of production rules holds, then the ACT-R theory claims that:

“Complex cognitive processes are achieved by stringing together a sequence of such rules by appropriate setting of goals and other writing to working memory, and by reading from working memory”

That is to say a complex task is performed by a number of cycles of the production system operation. (pattern matching and conflict resolution). One interesting implication of this interpretation is that it would become impossible to a cognitive skill to be mathematically complete. This is because production rule systems are equivalent of a subset of formal systems, which are incomplete (Godel's theorem).

Production rules are asymmetric, given the consequence of a production rule the precondition is not made true by a rule. Anderson justifies this by relating it to human asymmetric problem solving behavior.

A final feature of ACT-R's production rules is that they are to some extent abstract. Generally this is implemented by variables. This feature seems to blur the technical distinction between schema based and production based frameworks. If a production framework can be context dependent then a schema based theory can be represented in terms of abstract production rules. Of course it would be questionable whether the production rule was still representing a well defined modular step of cognition. But this question arises anyway: can a defined modular step of cognition involve a variable? If so is it still well-defined? a more in depth analysis of this is probably required.

There is some confusion as to the exact meaning of a number of terms Anderson uses. Of particular importance is: "cognitive skills" and "cognitive architecture".

"... cognitive skills does not encompass all of cognition." (Page 4 : Rules of the Mind - John R. Anderson 1993) however that's begs the question, what does cognitive skills encompass? Anderson gives no direct answer in this paper, the only answer that can be perceived from "Rules of the mind" is an implicit understanding that cognitive skills encompasses the examples from the experiments he draws to support his claim. But this would mean that the original claim is not empirically falsifiable! So one must generalize from the examples given and assume that cognitive skills encompass, perceptual, motor and conceptual problem solving tasks.

"Cognitive architectures are relatively complete proposals about the nature of human cognition." p3 Anderson seems to be defining a cognitive architecture as a theory for a large set of cognitive behavior, where a theory itself only specifies a small portion of the cognitive process. However by the use of cognitive architecture it is unclear where in the hierarchy of framework, theory, model, architectures fit. From the computer science field an architecture is a way of building a system, and is independent of the algorithms involved, this seems to relate more closely to framework level than theory. However because Anderson only complements the nature of cognitive architectures and so ACT-R by being an implementation from a cognitive architecture, this can be ignored with simple request for a more exact definition.

evidence for ACT-R

Because ACT-R is a theory of on the Production Rule framework we have to question the validity of this framework. However framework's fail the most fundamental requirement of scientific theory: empirical falsifiability. And as such Anderson claims that a framework can only be evaluated in terms of the best theory it supports : "The evidence for a framework always comes down to the success of the best theory specified with it" (page 11: Rules of the Mind - John R. Anderson 1993).

The ACT-R theory specifies two types of long term memory, declarative and procedural. Declarative memory is

remembered by a process of a production rules - this is an important implementational detail, because it's not clear whether the ACT-R theory specifies this or it is an implementational change from the method used (spreading activation) in the original ACT system. Anderson claims that the two systems behave the same, but doesn't provide justification for this detail, leaving it's nature unclear.

However Anderson does continue to justify the split in long term memory between procedural and declarative. The differences between the behaviors of long term procedural and declarative memory is defined by Anderson as:

“Reportability”: procedural knowledge is not reportable, where declarative is. And while Anderson does discuss this later, he has earlier claimed that not all declarative knowledge is reportable, if so then there either there must be a third attribute that helps make the procedural declarative distinction (this is what Anderson claims) and evidence is required, or the two types of memory are not distinguishable by this attribute.

“Associative Priming”. The example given is:

“... when one hears the word ‘computer’, there is priming for the word ‘programming’ (e.g., one can read the word ‘programming’ more rapidly [directly after reading the word ‘computer’]), but not for ones’s computer programming skills (i.e., one cannot program more rapidly).”

This seems like a poor example as it could be explained that programming is not related to the hearing the word ‘computer’, why should there be any associative priming in this case? Where it is easily explainable that the task of writing the word ‘programming’ having just heard it, is a related task (writing and hearing words are related).

Anderson also claims that there is no corresponding associative priming in procedural memory. This claim is also possibly refuted by a change in the interpretation of the same example: the task of writing and hearing seem describable in terms of procedural knowledge, so associative priming does occur between procedural knowledge. This points out two important observations about the nature of justification by experimental research in this field: Experiments can be interpreted very differently yielding completely different results; and also that, assuming a distinction between declarative and procedural knowledge, within these experiments it's very difficult to test declarative knowledge without testing procedural.

“Retrieval Asymmetry” is the claim that there is an asymmetry of access to procedural knowledge that isn't true for declarative. At first sight this seems to be a direct contradiction to “Reportability” and “Associative Priming” combination, procedural knowledge is not reportable, so how can access be measured given that there is no associative priming?

Anderson provides the results of an experiment in using the programming language LISP. Two tasks are given, one to interpret LISP code and the other to write LISP code. The experiment shows that there is no significant improvement in writing LISP code, if the subject had only prior experience with interpreting LISP code. And secondly the subject's experience of programming in LISP did not significantly improve interpreting LISP. This experiment would seem to show that a procedure learnt doesn't allow one to perform the inverse procedure, hence support the distinction between procedural and declarative knowledge by the asymmetric nature of procedural knowledge, and therefore support ACT-R's claim that cognitive skills are realized through asymmetrical production rules. However it is questionable that interpretation of LISP code is really the inverse procedure of writing LISP code, and not simply a largely unrelated

complex cognitive task. Another important issue is the difficulty of learning LISP, it is conceivable that if the subject has only a basic understanding of LISP, let's assume they only had an Interpretative skill level (as discussed earlier), then the subject would not be using procedural knowledge, but means end analysis with declarative knowledge.

“Retention” is the claim that the retention function of procedural knowledge is different from that of declarative knowledge:

“The most striking case of this is when people get better as using the procedural knowledge but worse at recalling the declarative knowledge.” - p22.

However in ACT-R the recall of declarative knowledge is by a production rule, which would suggest that in ACT-R the function of memory loss between declarative and procedural must be linked. intuitively this seems correct, however more evidence on all type of cognitive skills would be required to propose a strong justification.

“Dissociation”. This is the concept that procedural knowledge and declarative knowledge can be lost independently. Evidence for this is given by studies of amnesia. Anderson claims that amnesia in patients that can no longer recall new events have lost the ability to acquire new declarative information. However amnesia patients have been taught new skills. The interpretation of studies on amnesia patients is controversial, some people point to studies and claim that it is evidence against the procedural declarative distinction (Nissan, Knopman, & Schacter, 1987) and some point to the same empirical data and claim that it provides an argument for the procedural declarative knowledge distinction (Anderson). Given the lack of consensus it seems hard to draw a conclusion.

However if it is assumed that the data does support procedural-declarative distinction as implemented in ACT-R, then there is the question of how is this modeled in the ACT-R theory? The ACT-R theory does not involve the loss of declarative structures, so ACT-R interprets the loss of knowledge as “week initial memory traces” however it is unclear what these are and whether they are at the implementational or theory level.

in conclusion

Anderson proposes to evaluate frameworks by the best theory within it, and to evaluate a theory by how well it fit's to existing data and how likely it is to fit with all possible data. Some evidence and justification is given in support of the ACT-R theory, however the interpretation of this evidence is difficult to agree upon, and as such it is hard to agree with the ACT-R theory, especially as some of the evidence seems under an equally valid interpretation to provides a counter example for concepts in the theory.

A second major concern with the ACT-R theory is that the only obvious distinctions between the production rule system and a classical first order logic forward deduction system is 1. the limit on working memory, and 2. that a production rule must represent a well defined step of cognition. 3. The exact nature of the conflict resolution step. It would seem that ACT-R theory is then proposing that cognition is logical derivation in terms of well defined conceptual steps, where there is an imperfection on memory limits. This means that classical criticisms of modeling cognition in terms of predicate logic probably apply to production systems. In particular questions about vagueness of representation.

This is closely tied to the question of the level of representation of declarative knowledge. If declarative knowledge is a

clear concept in cognition then there must be some base concepts, and the question of grounding reoccurs. If not then production rules don't represent well defined pieces of cognitive skill.

As ACT-R attempts to cover such a large scope of cognitive tasks, it is not clear what the alternatives are. Within some domain's better alternatives exist, but no complete system mentioned in Rules of the mind matches the extent of the proposition of the ACT-R theory. Within smaller domain's categorization and analogy for example, alternative do exist, for example. One alternative is schemata based approaches, however this has only been shown to match empirical data within a small domain. Another is connectionist approaches, but they don't model the same scope as ACT-R. Given that no theory, as an implementation of a any framework, is very convincing, it seems that either that cognitive processes are too varied to be be modeling within the given frameworks. So either there is a theory possible that explains cognition, or there isn't, and some combination of theories and frameworks will have to be used to approximate cognitive processes.

This brings us to the scientific question of what cognitive science is supposed to be doing, Anderson seems to believe that there is a 'Correct' theory of cognitive ability and even sites the possibility of non-unique solutions as a problem. Any solution will probably involves different theories dependent on the modeling context. However it seems that the task is not one of finding a correct solution, but rather a theory that approximates cognition accurately as is possible - cognition may be understood by approximate theories.

Comments about this assignment: It is not clear what the memory based approach to cognitive modeling is, there is not formal definition in any of the lecture notes. I have assumed that it is the implicit notion of how to model cognitive modeling that has developed throughout the course. As such I have not referred to it directly as case/memory based approach, but simply as another approach. I have talked to several other students about this and they shared by difficulty in understanding what exactly this distinction is, but seeing as I have had no choice but to leave this assignment until this point in time (I have has over 60 hours of work per week) I have no answer to this question.