Metacognition in Problem Solving

Adam Dossa

Abstract

This paper explores the role that metacognition has to play in problem solving. We focus primarily on its role as arbitrator between different aspects of the problem solving process, proposing a simple model of this process to help highlight these interactions. We suggest that it is this metacognitive aspect of problem solving which best explains our ability as humans to solve novel and previously unseen complex problems.

1. Introduction

Brownell (1942) gives the following as a definition of problem solving, which we use to guide our focus in this paper:
... problem solving refers (a) only to perceptual and conceptual tasks, (b) the nature of which the subject by reason of original nature, of previous learning, or of organization of the task, is able to understand, but (c) for which at the time he knows no direct means of satisfaction. (d) The subject experiences perplexity in the problem situation, but he does not experience utter confusion. ... problem solving becomes the process by which the subject extricates himself from his problem.

Whilst problem solving is by no means restricted to humans (it can be seen in both computer systems and the wider animal kingdom), our innate ability to quickly solve novel and previously unseen problems across varied domains does characterize our adaptability which is not easily replicated elsewhere. In this paper we discuss the framework within which this problem solving occurs, the role that metacognition has to play in this process, and what factors may improve or degrade the effectiveness of metacognition in this context.

Our conceptual model of metacognition for the purposes of this review is that of cognitive monitoring and cognitive control. In this sense we consider metacognition to be the process by which (either consciously or unconsciously) we direct our cognitive processes via a feedback loop moderated at a higher (meta) cognitive level. As an example within the context of problem solving, this might entail the monitoring of a problem solving strategy for efficacy, and if this monitoring determines that it is not an appropriate or efficient method of solving our problem, the control function could direct the choice of an alternative strategy.

Solving a non-trivial problem entails many different cognitive components, each related to the other in various complex ways. Whilst each of these components can be considered individually, and are independent in some limited ways, we argue that the best conception of the role that metacognition has to play in problem solving is to view it as the means by which the interrelated co-dependencies between these components are managed on a real-time basis, during the problem solving process. A choice made for any one of these components will necessarily impact the other components, and so in order to seek an optimal configuration of choices across all of these components we appeal to a higher level of cognitive processing whose role it is to access the checks, balances and trade-offs between these choices, and iterate towards an optimal solution.

Problem Solving

It is worth spending a little time on providing a more complete definition of what is meant by problem solving. Anderson (1980) says that problem solving is "any goal-directed sequence of cognitive operations", which is characteristic of the idea that problem solving is inherently an iterative process. There have been a number of problem-solving models proposed which try to generalize these concepts, such as the General Problem Solver (Newell & Simon, 1972) which splits the process into an understanding phase followed by a search phase, and the IDEAL problem solver (Bransford & Stein, 1984) which involves Identifying problems, Defining problems, Exploring strategies, Acting on those strategies and finally Looking back to evaluate success. The ACT-R Model of Operator Selection is more computational in approach, and describes the process by which problem-solving operators are chosen depending on metacognitive reasoning (such as the history of success of associated production rules, which are generalized versions of the specific operators being instantiated in the given problem).

One aspect of investigating problem solving which can lead to confusion is the sense in which a "problem" encompasses such a wide set of possible interpretations. This could be anything from a simple math or logic problem, through to a social problem (for example determining how to manage an employee) or anything in between. There have been efforts made to try and provide a more finely grained taxonomy of problems, for example Jonassen (1997) classifies problems as either well-structured (where you might expect the starting and end goals to be well-defined, and the solution to be obtained by the application of a finite number of rules or concepts to transform the initial problem to the required solution) or ill-structured (such as might be more commonly found in real-life, where there may be partial or incomplete problem specification, and no obviously defined solution). In terms of metacognition, it seems that the latter type of problem more obviously benefits from metacognitive strategies (Jonassen, 1997) such as comparing different solutions for relative quality and planning. Whilst we are not always explicit in our treatment, it is this type of ill-structured problem that we have in mind when discussing the role of metacognition below.

2. Model

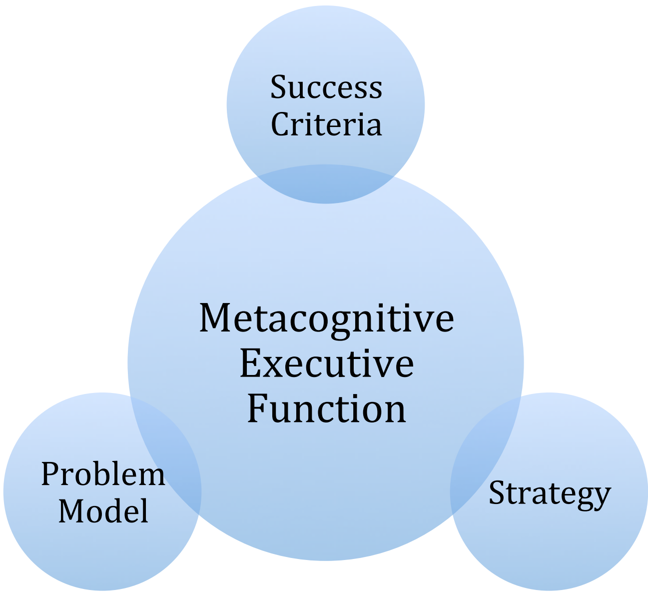
For the purposes of this paper we propose a relatively simplistic model of the problem solving process. The purposes of this model is not to provide an exhaustive computational account of the process of problem solving, but to provide a high-level framework that best exhibits the interplay of different problem solving components, and the management of these relationships (through metacognition).

We consider the problem solving process to have three high-level components; the formulation of a problem model, the choice of success criteria, and the selection and application of a problem solving strategy. As an illustrative example, consider the task of sorting a deck of playing cards. In this case a problem model could be that there are 52 distinct cards and a well-defined ordering between any pair of cards. Our success criteria could entail an assessment of what percentage of the cards are correctly ordered, and our strategy be to split the deck into two halves, sort them independently, and then merge the two halves maintaining their ordering.

We view the problem solving process then to be a series of iterations, where each iteration is defined by the choices made for each of these three components. Our metacognitive function is the executive process that directs these choices, and acts to continually refine and improve our choices until ideally we reach a solution to our problem (or determine that no solution is feasible).

Reder (1982) proposed two main classes of variables that impact our choices during problem solving for a given problem; those which are integral to the question itself (e.g. familiarity with the terms in the problem) and those which are more general in nature (e.g. prior history of success with different approaches). In the below we give examples of these types of factors that could be considered in isolation for each of our components, whilst in the subsequent section we consider the interplay of such factors across different components in our model.

2.1. Diagram



3. Independent factors influencing component choice

If one were to implement a problem solving mechanism computationally, it may be sufficient to simply specify each of these individual components, then hit start. The system would then run in a deterministic fashion until a solution is reached. For example, in the case of determining the prime factors of a number, we would specify the problem representation as being the domain of integer numbers under a multiplication operator, the success criteria as being that we have found two numbers which multiple to form our input prime number, and a strategy as simply trying every conceivable pair of integers until we find a valid pair (whose product is our input prime) or until we determine that no such pair exists.

Fortunately real life is usually more exciting, and instead of such a linear approach we would expect to have to dynamically modify the choice for each of our component pieces as we proceed, based on feedback from the efficacy of our approach thus far. It is in this interplay that we feel the most decisive role of metacognition is played, but none the less there remain metacognitive variables which have a role to play in each of our component choices which are (mostly) independent of the choices made for other components. In the below we elaborate on each of the components in our model, and mention some of these independent variables where appropriate.

3.1. Problem Modeling

When faced with a problem to be solved, the first step must be to try and understand the problem. This process of understanding a problem can be thought of as mapping the problem as presented to some internal and consistent representation of the problem. It is worth noting that this process may not be entirely exhaustive in the first instance, since we may choose to understand the problem to some limited level of abstraction, or to focus on our understanding of just one part of the problem initially, in the expectation that on further iterations we will refine our model.

This mental representation or model has been called the problem space (Newell & Simon, 1972), and can be thought to consist of some combination of structural knowledge, procedural knowledge, reflective knowledge, images and metaphors of the system, and executive or strategic knowledge (Jonassen & Henning, 1999). Our model may be informed by both intrinsic and extrinsic factors, with an appeal made to metacognition as being the engine through which these variables are weighed and a choice made. Research has often focused on trying to find problem settings that force one particular representation over another, but contrary to intuition, this seems very hard to achieve in practice. As an example a common psychometric test of spatial ability may involve extrapolating a two dimensional net to a three dimensional solid, but even in as well-defined case as this, research has shown (using introspection, gaze-tracking and so on) that some people will employ a spatial internal representation, whilst others will employ a more symbolic analytical representation (Barratt, 1952; French 1965, Just & Carpenter, 1985).

The factors that determine our choice of representation are likely some combination of both intrinsic factors (if the question is posed in a symbolic fashion, we may internalize it symbolically) as well as extrinsic factors such as the prior success we may have had using similar representation, or some facet of domain knowledge we have that may guide us towards one representation over another. Roberts (2000) suggests that individuals have a cognitive style, along with a degree of inflexibility, and that this plays a large part in how we choose to internally represent a problem, along with these more objective variables above.

3.2. Success Criteria

In order to have some goal-directed approach to problem solving we must have some criteria by which to measure the success of our problem solving process. This success criteria in a simple case may simply be a binary success or failure assessment, but in a more complex problem setting we would expect it to be considerably more nuanced.

For example it may be that we wish to optimize our problem solving process for speed over accuracy, which can be viewed as placing a small weight on accuracy and a large weight on speed when accessing the success of our problem solving approach. In a problem solving setting where there is not any absolutely correct or incorrect solution, we may have a bias towards solutions that fulfill certain criteria (such as practicality).

Especially relevant is the concept of a payoff matrix for certain solutions; this encapsulates the idea that not all outcomes or solutions will be of equal value to us, and that given a solution that may or may not be true we may wish to associate different payoffs with different situations. As an example, given the problem of medical diagnosis of cancer, we would recognize that a 100% accurate diagnosis is improbable, but that given an incorrect diagnosis it is far better that we have a false positive (where we diagnose a patient as possibly having cancer when they don't have cancer) as compared to a false negative (where we diagnose a patient with cancer as not having cancer). Metacognition should enable us to make smart choices for our payoff matrix, informing us when it may be beneficial to sacrifice speed for accuracy, the degree to which it is possible to interpret the accuracy of a solution, and the relative benefits of different types of solutions.

It is worth noting that in order to maximize our ability to leverage metacognition in the problem solving process, our success criteria should be applicable to interim stages of the problem solving process, and not just in the evaluation of a complete solution. As an example consider the problem of reaching the center of a maze. One possible success criteria could encompass the speed with which we solve the maze and the efficiency with which we solve it (i.e. the inverse of the number of times we need to backtrack). However, these criteria will only allow us to access our problem solving approach once we have actually reached the center, so be of limited use in terms of iterating through possible strategies during our search for a solution. In practice research has shown that people often make interim judgments of the success (or otherwise) of a problem solving approach by judging how much closer to their goal it has taken them, even if they have not yet reached their goal. For instance, in the water jars task problem setting college students were shown to prefer moves that took them closest to their goal state (Atwood & Polson, 1976). As such success criteria that incorporate the ability to make interim success judgments allow a faster iteration through possible strategies by providing fast and relevant feedback on an ongoing basis.

3.3. Strategy Selection

The process of strategy selection will be critical to our ability to resolve a problem. For a well-structured problem this can be thought of as being the rules and transformations that we apply to our problem model in order to determine the solution, and in the more general ill-structured problem, the methodology through which we seek to infer a feasible solution. It seems reasonable to assume that in general we have access to some library of strategies (which Roberts (2000) calls strategy availability) and that we make a choice from this library using metacognitive function to determine the best option at any given point in the problem solving process.

It is worth briefly noting that whilst the term strategy can widely encompass any number of possible methodologies, there are some common types of strategy (which are often leveraged in algorithmic problem solving) that we briefly mention here. A brute force strategy where, for some well-defined problem, we simply iterate through every possible option and check to see any of these options constitute a solution, is a strategy which is very simple, guaranteed to succeed (provided our model is correct), but which will generally be very inefficient. As an example consider solving a crossword clue by simply trying every combination of letter until we find a valid word that fits the clue – we are guaranteed to find a solution, but it may take us some time! An alternative smarter strategy is a greedy approach (also called a hill-climbing approach), where at each point where we need to make a decision from a set of finite options, we simply choose the decision that takes us closer to our stated goal. We note here that such a strategy relies on having success criteria that can access how close we are to our goal, rather than just simply whether or not we have reached our goal. We mention these examples as illustrations of when our strategy may be confined by our success criteria, in the sense that if we can only determine a measure of success once we have a solution, we may be constrained to relatively simplistic strategies as a consequence.

One last dichotomy of strategies that is worth considering is that of retrieval versus calculation. In the former case, we shortcut much of what has preceded this section, in the sense that if we recognize that we have seen a problem before it may be possible to simply retrieve the solution from memory, without going through the computationally expensive process of calculating the solution. Whilst it is certainly true that metacognition is the mechanism by which a decision to retrieve or calculate is made (Schunn, Reder, Nhouyvanisvong, Richards & Stroffolino, 1997), for the purposes of this paper we restrict ourselves to the case that the problem is novel, and hence simple retrieval of a solution is not an option.

4. Interplay between different components

In the above we have attempted to consider each component somewhat in isolation, in order to clearly delineate the role of each component. We now come to the area which we feel is most interesting in terms of the role of metacognition in problem solving, namely the management of each of these components simultaneously as an entire system. Whilst it is clear that each of these components is interrelated, we argue that the metacognitive choices made which optimize these relationships that is critical to the adaptability of our problem solving skills.

One area that we have not previously elaborated on, but which is relevant under some definitions of metacognition, is the degree to which the metacognitive role in decision-making is implicit or explicit. Much of the more recent research would seem to indicate the former, and that much of the metacognitive function in terms of the above components happens unconsciously, and further that we may also be unaware of the choices we have made at a conscious level (Reder & Schunn, 1996). For the purposes of this paper, given the definition of metacognition that encompasses both implicit and explicit processes, we do not dwell on these distinctions, although it is clearly of interest both in the context of the role of consciousness in strategic problem solving, and in how we may seek to improve metacognition in problem solving.

4.1. Problem Modeling and Success Criteria

Here we discuss some of the factors that influence the relationship between our definition of success and the choice of how to internalize our problem. That these two processes are interrelated seems intuitively clear, since the scope of each of these processes will be delineated by the other. To illustrate suppose we choose a model representation which is highly abstract or simplified, such that the solution to our problem can only be judged in a binary fashion (as either a success or failure). Then our corresponding success criteria must necessarily reflect this, allowing us only to judge performance in a strictly binary fashion, and allow no means of accounting for partial success. Conversely, if we have decided to access our success in terms of biasing heavily towards accuracy over speed, then this would imply that we ought to choose a model within which it is possible to make accuracy assessments. For example, in the water jars task mentioned above, and a similar building sticks task (Lovett & Anderson, 1996), participants favored strategies which reduced their distance to the stated goal at each step. In order to implement such a strategy it must be necessary to incorporate distance to the stated goal as part of the success criteria, and therefore have the model correctly represent this.

As mentioned it may also be the case that we wish to access the model not just in terms of success and failure, but in terms of the probabilities of each of these events, given a particular solution (in the case of an ill-structured problem where there may not be a well-defined absolute solution). If we wish further to incorporate a payoff matrix structure where we consider these probabilities in their relationship to different possible situations, then our model must also include this ability. As an example, suppose that our problem is to determine how much to pay for a particular product in a competitive transaction. If we wish, as part of our success criteria to consider both our outcome, and the outcome of our competitor (with whom we are transacting the purchase) them our model must include both sets of outcome, so that we can accurately determine the success of our decision as a function of both our and our competitors perspective.

It may be the case that our success criteria and problem model are even more intimately intertwined. For example it may only be possible to make a performance judgment within the context of our model (and in fact, generally since our model reflects our understanding of the problem, we would expect this to usually be the case). As an illustrative example, if we consider the problem above of optimizing the purchase price of a product, then in order for us to access our performance relative to our competitor we may need to appeal to our model to access how our competitor views different outcomes.

As a final point in this section we note that there may be various cognitive biases or misconceptions that lead to misinformed success criteria. This may vary from incorrect probability or risk assessments to failing to consider certain aspects of the problem even if these are present in our model (for example considering only one criteria of success when there may be multiple such criteria). This illuminates the codependency between problem modeling and success criteria in the sense that the determination of success is implicit in both areas – since our problem model defines the problem, implicit in this is some sense of solution, whilst at the same time our success criteria may represent a more detailed or multi-dimensional idea of success or partial success, along with other criteria such as speed, accuracy and payoff matrices.

4.2. Success Criteria and Strategy Selection

In order for our metacognitive function to be effective, we must have some method of accessing the efficacy of our chosen strategy through metacognitive monitoring. Our ability to make such an assessment depends on a number of factors related to the problem setting and chosen strategy. For example if our chosen strategy is of an "all or nothing" variety where the success or otherwise of the strategy is difficult to access until we have completed our strategic process, then we would expect to have a longer feedback loop, which may negatively impact our ability to quickly iterate through several possible strategies. Conversely it may equally be that the problem setting provides quick and immediate feedback on how successful our strategy is (for example if we are playing an adversarial game consisting of sequential turns, then after each turn we can use our success or failure to inform our future strategy).

Whilst for some familiar or simple problems it may be the case that we are able to immediately choose a successful strategy, in many cases we would expect our choice of strategy to evolve throughout the problem solving task. This evolution of strategy can be subtle, for example fine tuning the parameters of our model or strategy, or dramatic, where we entirely abandon one strategy and move to a very different model or strategy. Metacognitive control and its interplay with metacognitive monitoring would seem to play a critical part in this process – if we determine that a strategy is working correctly but not yielding a solution we may realize that our conceptual model is faulty and must be adjusted, or if we realize that our strategy is yielding incorrect solutions, perhaps that the strategy itself is faulty in this domain and must either be tweaked or abandoned.

When looking at the iterative nature of success strategy choices, it seems reasonable to consider two cases. The first is where we are making small, fine-tuning type (and probably implicit) modifications to our strategy as a result of ongoing metacognitive monitoring that relies on fast heuristic incremental calculations. As an example of such a process we can consider the probability learning experiment by Reber and Millward (1971) in which participants were able to learn to anticipate the changing probabilities of events over the course of each trial, without any explicit awareness that this change was occurring. This would correspond to fast heuristic appeals to our success criteria to provide ongoing feedback on the performance of our probability learning strategy (metacognitive monitoring), and consequent small tweaks to our strategy to reflect this (metacognitive control). These changes may not need any corresponding modification to our model, which continues to correctly reflect the problem.

The second case is where we make large, conceptual modifications to our strategy. In this case our success criteria, rather than providing small incremental feedback is suggesting that we need to modify our approach more dramatically. For these types of change it is more likely to be an explicit process in the sense that we are both aware of the reasons that we are changing strategy, as well as the changed strategy. An example might be changing from a brute force strategy to a more nuanced approach, if we determine via our success criteria that the brute force approach is unlikely to lead to a solution within a reasonable timeframe. Of course it could also go in the opposite direction if we determine that our complex strategy is not generating reasonable solutions, so default back to a simpler strategy as an interim measure until we can determine a better (and more efficient) strategy.

In summary, in order to have an efficient metacognitive feedback loop (consisting of both monitoring and control) we ought to favor strategies that allow for small fine-tuning based on constant feedback from our success criteria, and correspondingly favor success criteria that can provide relevant and accurate ongoing feedback with a low computational overhead (so that this process is fast, and probably implicit). Our success will also be influenced by overall strategy availability, in the sense that having a deep library of possible strategies (with corresponding success criteria and models) will allow us to both make a reasonable initial choice of strategy (based on the usual intrinsic and extrinsic factors) as well as update that choice based on metacognitive feedback, if necessary making large conceptual changes to all parts of our process.

4.3. Strategy Selection and Problem Modeling

There is clearly a deep relationship between our strategy and problem model. The choice of model and strategy cannot be made independently as certain strategies will only be applicable to certain models and vice versa. It is worth noting however that this is a many to many relationship in the sense that a given strategy may be appropriate for more than one model of a problem and also that a given model may lend itself to more than one choice of strategy to determine its solution.

As mentioned above strategy selection and problem modeling may reflect an individual cognitive style based on the individuals prior experience with that model and strategy combination. Roberts (2000) lists several possible cognitive styles, including whether people are field dependent or independent, assimilators or explorers, adaptors or innovators, visualizers or verbalizers, holists or serialists, and convergent or divergent thinkers. He further notes that there has been little attempt to unify these types of dichotomies into a single framework, and further that other personality traits (such as extraversion an neuroticism) may also influence our problem solving choices.

Domain experience will also play a large part in our strategy selection and problem model choice. In the case where we have relevant domain experience, it seems reasonable to expect this to be a dominating factor, and indeed the concept of strategy anchoring, termed the *Einstellung* effect has been observed in research. Experiments have replicated the effect that having seen a particular strategy be successful in earlier trials, participants are much more likely to apply that strategy to later problems, even if simpler more obvious strategies exist (Luchins, 1942). This is an example of metacognition having a negative effect, with the extrinsic factor of prior success overly dominating other intrinsic metacognitive criteria, which would otherwise guide us towards the simpler model and strategy.

When faced with a complex unfamiliar problem it seems reasonable to assume that our initial selected strategy may not end up being the correct choice, and so we would expect that metacognitive monitoring (and related control in the choice of alternative strategies) would play a larger part. If metacognition plays an important role in strategy selection, then we should prefer those strategies for which it is possible to receive relevant feedback. For example a strategy for which it is not possible to judge its efficacy on an ongoing (or at least frequent) basis may not be optimal unless we already know that it is the correct strategy for a problem. Through the interplay between our problem abstraction and strategy it may be that, for example by applying our strategy first to smaller and quicker abstractions of the same problem, we can improve our ability to gain useful feedback from our strategy, and hence iterate through problem conceptualizations and strategy selections faster.

5. Conclusion

From the above we can see that metacognitive functions play a variety of roles in the problem solving process. However, whilst assessment of individual cues to make reasonable choices around strategy, problem model and success criteria is essential, I argue that these functions are significantly less complex than those required to balance the relationship between these components in a dynamic fashion throughout the problem solving process. Whilst it seems conceivable to produce a deterministic algorithm that largely incorporates these types of intrinsic and extrinsic variables for one-off choices for each of these components, producing a system that can continually refine and finesse these choices, especially in their relationship to each other seems less likely. It is our ability as a species to do this, largely on an implicit basis and with little thought, that so sets us apart from state of the art computing systems, and the remainder of the animal kingdom.

6. References

David H. Jonassen (2000). *Toward a Design Theory of Problem Solving*

Namsoo Shin, David H. Jonassen & Steven McGee (1999). *Predictors of Well-Structured and Ill-Structured Problem Solving in an Astronomy Simulation*

Marsha C. Lovett & John R. Anderson (1996). *History of Success and Current Context in Problem Solving*

Lynee M. Reder & Christian D. Schunn (1996). *Metacognition Does Not Imply Awareness: Strategy Choice Is Governed By Implicit Learning and Memory*

H. Lee Swanson (1990). *Influence of Metacognitive Knowledge and Aptitude on Problem Solving*

Jeffrey D. Karpicke (2009). *Metacognitive Control and Strategy Selection: Deciding to Practice Retrieval During Learning*

Maxwell J. Roberts & Elizabeth J. Newton (2000). *Understanding strategy selection*

Christian D. Schunn, Lynee M. Reder, Adisack Nhouyvanisvong, Daniel R. Richards & Philip J. Stroffolino (1997). *Calculate or Not to Calculate: A Source Activation Confusion of Problem Familiarity's Role in Strategy Selection*

J. R. Anderson (1980). *Cognitive psychology and its implications*

A. Newel & H. Simon (1972). *Human problem solving*

J. Bransford & B.S. Stein (1984). *The IDEAL problem solver: A guide for improving thinking, learning, and creativity*

Hwa Tee Yong and Lau Ngee Kiong. *Metacognitive Aspect of mathematics Problem Solving*