COORDINATED SCIENCE LABORATORY
College of Engineering

A MODEL OF ATTENTION FOCUSSING DURING PROBLEM SOLVING

Jude W. Shavlik
Gerald F. DeJong

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
Unclassified	None

2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT
N/A	Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
N/A	

4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
UILU-ENG-86-2220	N/A

6a. NAME OF PERFORMING ORGANIZATION	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION
Coordinated Science Lab University of Illinois	N/A	National Science Foundation

6c. ADDRESS (City, State and ZIP Code)	7b. ADDRESS (City, State and ZIP Code)
1101 W. Springfield Ave. Urbana, IL 61801	1800 G Street Washington, D.C. 20550

8a. NAME OF FUNDING/SPONSORING ORGANIZATION	8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
National Science Foundation	N/A	NSF IST 85 11542

8c. ADDRESS (City, State and ZIP Code)	10. SOURCE OF FUNDING NOS.			
	PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
1800 G Street Washington, D.C. 20550	N/A	N/A	N/A	N/A

11. TITLE (Include Security Classification) A Model of Attention Focussing During Problem Solving

12. PERSONAL AUTHOR(S)
Jude W. Shavlik, Gerald F. DeJong

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Technical	FROM _____ TO _____	March 1986	10

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	attention focussing, mathematical problem solving, novice-expert differences, problem-solving operators, computational systems

19. ABSTRACT   We propose that three qualitatively different strategies help focus attention during problem solving. The first strategy is to apply operators that will lead to definite progress toward the goal. Attention will be focussed by this strategy as long as some operator in this class is applicable. When clear progress can not be achieved, the problem solver must decide how best to proceed. It then invokes the second strategy to select operators that preserve important characteristics of the current problem. These operators are likely to keep the problem solver from diverging sharply from the goal while possibly enabling the application of operators by the first strategy. When the problem solver can follow neither of the first two strategies, it invokes the third strategy of arbitrarily applying legal operators. We see the second strategy as an essential difference between novice and expert problem solvers. It is easy to recognize definite progress toward a goal and it is easy to recall which operators can be legally applied. Expertise involves knowing which characteristics of a situation should be preserved (or created) when no way to definitely progress toward the goal is known. This three-strategy theory has been implemented and tested in a system that performs mathematical calculations in the course of solving physics problems. We describe a number of mathematical calculation operators used under each strategy.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED X SAME AS RPT. ☐ DTIC USERS ☐	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
		None

A MODEL OF ATTENTION FOCUSSING
DURING PROBLEM SOLVING [*]

Jude W. Shavlik[†]
Gerald F. DeJong

Artificial Intelligence Research Group
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, IL 61801

ABSTRACT

This paper also appears on pages 817-822 of the *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, Amherst, Massachusetts, August 1986.

We propose that three qualitatively different strategies help focus attention during problem solving. The first strategy is to apply operators that will lead to definite progress toward the goal. Attention will be focussed by this strategy as long as some operator in this class is applicable. When clear progress can not be achieved, the problem solver must decide how best to proceed. It then invokes the second strategy to select operators that preserve important characteristics of the current problem. These operators are likely to keep the problem solver from diverging sharply from the goal while possibly enabling the application of operators by the first strategy. When the problem solver can follow neither of the first two strategies, it invokes the third strategy of arbitrarily applying legal operators. We see the second strategy as an essential difference between novice and expert problem solvers. It is easy to recognize definite progress toward a goal and it is easy to recall which operators can be legally applied. Expertise involves knowing which characteristics of a situation should be preserved (or created) when no way to definitely progress toward the goal is known. This three-strategy theory has been implemented and tested in a system that performs mathematical calculations in the course of solving physics problems. We describe a number of mathematical calculation operators used under each strategy.

A MODEL OF ATTENTION FOCUSSING
DURING PROBLEM SOLVING

INTRODUCTION

A novice problem solver attacks a problem in one of two ways. He may immediately notice a way to progress toward the solution. Alternatively, he may flounder around performing legal, but aimless, operations in an attempt to transform the problem into a familiar form. An expert, on the other hand, can perform in a qualitatively different manner. If the solution is not immediately apparent, he can focus his efforts in a much more guided way. Instead of simply thrashing around, he has an appreciation of what kinds of transformations are likely to change the current problem into a soluble problem.

Consider the problem of evaluating the expression

$$mass_1\ velocity_1\ +\ mass_2\ velocity_2$$

when $mass_1$ and $mass_2$ are known but $velocity_1$ and $velocity_2$ are not. This expression cannot be evaluated directly. A valid approach might be to substitute equivalent expressions for the unknowns. A novice might perform the unappealing substitutions of figure 1. While these transformations are valid, they are unlikely to yield a solution. An expert will appreciate this, and be more likely to perform the pleasing problem transformations of figure 1. In this example, there is something about the parallel structure of the problem that makes a parallel substitution more appealing. Yet parallel substitutions are not always aesthetically appealing. For example, consider the substitutions of figure 2. In this example, parallel substitution using Newton's third law (every action has an equal and opposite reaction) misses a useful variable cancellation. If only one instantiation of this formula is used, the two forces can be eliminated from the calculation.

$$mass_1 \ velocity_1 \ + \ mass_2 \ velocity_2$$

$$mass_1 \ \frac{d}{dt} \ position_1 \ + \ mass_2 \ velocity_2 \qquad\qquad mass_1 \ \frac{d}{dt} \ position_1 \ + \ mass_2 \ \frac{d}{dt} \ position_2$$

$$mass_1 \ \frac{d}{dt} \ position_1 \ + \ mass_2 \ \int acceleration_2 \ dt \qquad mass_1 \ \int acceleration_1 \ dt \ + \ mass_2 \ \int acceleration_2 \ dt$$

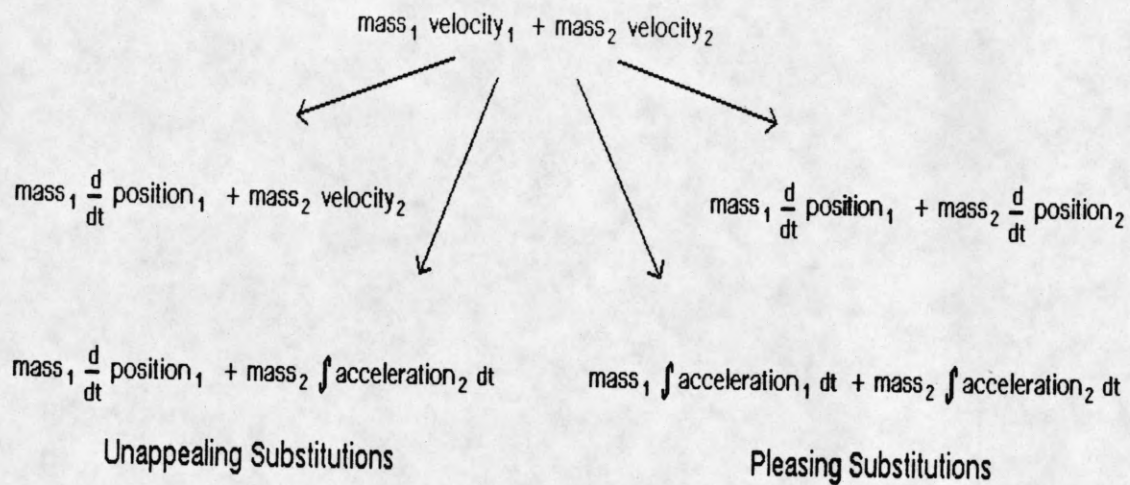**Unappealing Substitutions**   **Pleasing Substitutions**

Figure 1. Sample Mathematical Substitutions

We propose three qualitatively different strategies for human problem solving. Attention is selectively focussed according to one of these strategies. *Strategy 1* is hill climbing. Attention is focussed on this strategy as long as some operator moves the problem solver closer to its goal. Occasionally the problem solver will reach a local maximum; there will be no way to move closer to the goal. At these times, the problem solver must diverge from his goal, in the hopes of transforming the current situation into one where hill climbing can occur. We propose there are two qualitatively different ways by which a problem solver re-focusses his attention during this divergent phase. A novice problem solver merely selects an arbitrary legal operator. This can lead to aimless floundering, due to the large number of possible combinations of operator sequences. An expert can often wisely choose how to transform the current situation. Characteristics that are

$$\int force_{1,2} \ dt \ + \ \int force_{2,1} \ dt$$

$$\int (- force_{2,1}) dt \ + \ \int (- force_{1,2}) dt$$

Figure 2. Inappropriate Parallel Substitutions

believed to be of general or domain-specific problem-solving importance (e.g., symmetry) are to be maintained or introduced; introduction of troublesome characteristics is to be avoided. Such motivated diversions compose *strategy 2*. However, when an expert exhausts his expertise he also resorts to any legal operator. This unmotivated application of operators is termed *strategy 3*.

AN APPLICATION OF THE MODEL IN MATHEMATICAL CALCULATION

Our three-strategy theory has been implemented and tested in a system called **Physics 101** [Shavlik85]. This system performs mathematical calculations in the course of solving physics problems. We have identified a number of mathematical calculation operators used by the three strategies.

Our model is a *state-space* model. The problem solver is provided an initial state and a goal description. The role of the problem solver is to successively apply legal transformations to the current state until a state satisfying the goal description is reached. This process in the example domain of mathematical calculation is illustrated is figure 3. Notice that more than one substitution is allowed during the transformation from one state to the next.
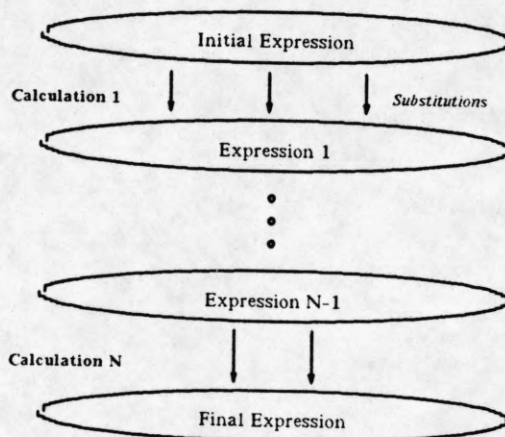
```
              ┌─────────────────────────────────────┐
              │         Initial Expression          │
              └─────────────────────────────────────┘
Calculation 1     │         │            │  Substitutions
                  ▼         ▼            ▼
              ┌─────────────────────────────────────┐
              │            Expression 1             │
              └─────────────────────────────────────┘
                              •
                              •
                              •
              ┌─────────────────────────────────────┐
              │           Expression N-1            │
              └─────────────────────────────────────┘
Calculation N          │            │
                       ▼            ▼
              ┌─────────────────────────────────────┐
              │          Final Expression           │
              └─────────────────────────────────────┘
```

Figure 3. The Structure of a Calculation Sequence

Strategy 1 - Definite Progress.

In our **Physics 101** system, often the goal is to produce an expression that only contains variables whose values are known. (Once in this state, the expression can easily be evaluated.) The hill-climbing measure is the number of variables in the expression whose value is not known (these variables will be called *unknowns* from now on).

Physics 101 contains two basic techniques for reducing the number of unknowns. One, unknowns can be replaced using known formulae if these formulae introduce known-valued terms or lead to the cancellation of unknowns. An example of this technique is shown in figure 4 (discussed below). Two, values of variables can used if doing so eliminates unknowns. If our current expression is $A * B * C$, and A's value is zero, both B and C can be cancelled by replacing A with its numerical value. Similarly, given the expression $A * B * C - B * C$, where A equals one, a numerical replacement can lead to a reduction in unknowns.

A problem-solving operator applied by strategy 1 is illustrated in figure 4. The goal is to evaluate the top expression, but the values of the two inter-object forces ($force_{1,2}$ and $force_{2,1}$) are not known. A "substitute-to-cancel-unknowns" operator can detect that the two inter-object

Consider the Expression

$$\int force_{1,2} \, dt \ + \ \int force_{2,1} \, dt$$

Choose a Variable-Cancelling Substitution

$$\int (- force_{2,1}) \, dt \ + \ \int force_{2,1} \, dt$$

Bring Cancellers Together

$$\int (- force_{2,1} \ + \ force_{2,1}) \, dt$$

Cancel Variables

$$\int 0 \, kgm \, / \, s^2 \, dt$$

Figure 4.   Example of an Operator Applied by Strategy 1

forces cancel due to Newton's third law. This operator is a complex operator. We allow operators to comprise a sequence of problem-solving steps that achieve some goal. (Operators comprising other operators have variously been called macro-operators [Fikes72], frames [Minsky75], scripts [Schank77], and schemata [Chafe75].) These "super" operators possess the desirable property that once one is selected, several problem-solving steps can be carried out without the need for intervening search. The "substitute-to-cancel-unknowns" operator first applies Newton's third law to replace one of the inter-object forces. It then brings these potential cancelling terms into a position where cancellation can take place. This requires that the two integrals be combined. Finally, the two troublesome variables can be eliminated.

Bundy's meta-level solution methods [Bundy81] follow strategy 1. He considers solving complicated equations containing a single unknown variable (there may be multiple occurrences of the unknown, however). His *attraction*, *collection*, and *isolation* methods always bring one closer to the goal of having the only occurrence of the unknown isolated on the left-hand side of an equation.

Strategy 2 - Motivated Diversions.

There are several techniques used in **Physics 101** that follow the second strategy. The major ones are presented below, in the order they are used by the system.

(1) *Elimination of variables whose values are known.* Even when is not possible to reduce the number of unknowns in an expression, it is a good idea to cancel terms, even those whose values are known. Eliminating these terms may allow productive cancellations that had been prevented by the presence of known-valued terms. For example, suppose we have the following formulae, and the values of A and C are known.

$$A = B / C \quad \text{and} \quad B = -D$$

If our current expression is $A * C + D$, we cannot reduce the number of unknowns. However if we use the first of the above formulae, we can cancel some terms, while

momentarily increasing the number of unknowns. Fortunately, in the next step both of the unknowns can be cancelled.

(2) *Elimination of integrals and derivatives.* Performing calculus is harder that performing algebra. When it is not possible to eliminate terms, it is often a good idea to eliminate calculus structure. An example where **Physics 101** removes calculus from an expression is shown in figure 5. Here the program detects that it can eliminate the derivative because it knows the derivatives of all the terms being differentiated. Applying this operator adds four steps to the calculation sequence. After this operator is applied, direct progress toward the goal state can be made. For instance, object 1's mass times its acceleration equals the net force on it (by Newton's second law). Hence, two variables can be replaced by one. (Continued calculation leads to this expression becoming zero, a consequence of the principle of conservation of energy.)

(3) *Preservation of expression type.* Assume we have the following two equations.

$$(i) \; A = (D * E) \qquad (ii) \; A = (F + G)$$

If our current expression is $A * B * C$, equation (i) would be preferred as this would maintain the property that the expression is a product of terms. Conversely, given $A + B + C$, the second equation is preferred, because now the expression continues to be a

$$\frac{d}{dt} ((1/2) \; mass_1 \; velocity_1{}^2 + mass_1 \; g \; position_1)$$

SeparateCalculus $\qquad = \frac{d}{dt} ((1/2) \; mass_1 \; velocity_1{}^2) + \frac{d}{dt} (mass_1 \; g \; position_1)$

ConstantsOutOfCalculus $\qquad = (1/2) \; mass_1 \; \frac{d}{dt} \; velocity_1{}^2 + mass_1 \; g \; \frac{d}{dt} \; position_1$

SolveCalculus $\qquad = (1/2) * 2 \; mass_1 \; velocity_1 \; \frac{d}{dt} \; velocity_1 + mass_1 \; g \; \frac{d}{dt} \; position_1$

SubstCalculus $\qquad = (1/2) * 2 \; mass_1 \; velocity_1 \; acceleration_1 + mass_1 \; g \; velocity_1$

Figure 5. An Application of the Substitute-Calculus Operator

sum of terms. There is a strong reason for preserving expression type, one involving more than aesthetics. In the first example we can produce

$$D * E * B * C \quad \text{or} \quad (F + G) * B * C.$$

In the result on the left, all the terms are equally accessible. Future substitutions involving B, for example, can cancel D or E. (Recall that cancellation of variables is the mechanism that leads to the goal state.) The right-hand result requires that a replacement for B cancel F and G together.

(4) *Preservation of structural symmetry.* When similar additive or multiplicative structure is present, the same general rule should be used repeatedly whenever possible. For example, given the following expression, substitutions involving all three of the A's or all three of the B's would be favored.

$$A_1 B_1 + A_2 B_2 + A_3 B_3$$

This accounts for the unpleasing quality of the first transformations of figure 1. It would be better to replace both of the velocities either by the derivatives of position or by the integrals of acceleration. Mixing the two does not seem right.

The mathematical problem-solving methods learned by Silver's **LP** program [Silver84] follow strategy 2. **LP** acquires information that constrains the choice of applicable operators. The learned operators are not guaranteed to bring the problem solver closer to a final solution.

Strategy 3 - Floundering Around.

In **Physics 101** strategy 3 looks for the first legal substitution and applies it. Only one substitution is made, in order to minimize this undirected perturbation of the calculation.

The **LEX** system of Mitchell [Mitchell83b] acquires heuristics that estimate when it is wise to apply an integration operator. It learns how operators previously used only under strategy 3 can be applied by strategy 2. The later **LEX2** system [Mitchell83a] learns under what conditions a

specific integration operator will lead to a solution. This can be viewed as learning how to apply, under strategy 1, an operator previously used only under strategy 3.

CONCLUSION

We propose that three qualitatively different strategies help focus attention during problem solving. The first strategy is to apply operators that will lead to definite progress toward the goal. Attention will be focussed by this strategy as long as some operator in this class is applicable. When clear progress can not be achieved, the problem solver must decide how best to proceed. It then invokes the second strategy to select operators that preserve important characteristics of the current problem. These operators are likely to keep the problem solver from diverging sharply from the goal while possibly enabling the application of operators by the first strategy. When the problem solver can follow neither of the first two strategies, it arbitrarily applies operators. This three-strategy theory has been implemented and tested in a system that performs mathematical calculations in the course of solving physics problems. We have identified a number of mathematical calculation operators used by the three strategies.

When several operators are seen as being equally viable, a problem solver must choose which to apply. Under strategy 3 the choice is made arbitrarily. In the other two cases, the choice is made by using the second strategy. For example, if there are a number of ways to cancel two unknowns, a way that preserves the symmetry of the situation is preferred.

These three strategies can be viewed in terms of simulated annealing problem-solving models [Hinton84, Kirkpatrick83]. Progress using strategy 1 involves the movement toward a solution space minima. Strategy 2 is analogous to slightly increasing the system "temperature" when stuck at a local minima that is not a goal state. Here it is hoped that the problem solver does not drift too far in the problem space. Strategy 3 potentially involves much greater increases in system temperature, and, hence, much greater jumps in the problem space.

We see the second strategy as an essential difference between novice and expert problem solvers [Chi81, Chi82, Larkin80]. It is easy to recognize definite progress toward a goal and it is

easy to recall which operators can be legally applied. Expertise involves knowing which characteristics of a situation should be preserved (or created) when no way to definitely progress toward the goal is known.

REFERENCES

[Bundy81]   A. Bundy and B. Welham, "Using Meta-level Inference for Selective Application of Multiple Rewrite Rules in Algebraic Manipulation," *Artificial Intelligence 16*, 2 (1981), pp. 189-212.

[Chafe75]   W. Chafe, "Some Thoughts on Schemata," *Theoretical Issues in Natural Language Processing 1*, Cambridge, MA, 1975, pp. 89-91.

[Chi81]   M. T. Chi, P. J. Feltovich and R. Glaser, "Categorization and Representation of Physics Problems by Experts and Novices," *Cognitive Science 5*, 2 (1981), pp. 121-152.

[Chi82]   M. Chi, R. Glaser and E. Rees, "Expertise in Problem Solving," in *Advances in the Psychology of Human Intelligence, Volume 1*, R. Sternberg (ed.), Lawrence Erlbaum and Associates, Hillsdale, NJ, 1982, pp. 7-75.

[Fikes72]   R. E. Fikes, P. E. Hart and N. J. Nilsson, "Learning and Executing Generalized Robot Plans," *Artificial Intelligence 3*, (1972), pp. 251-288.

[Hinton84]   G. F. Hinton, T. J. Sejnowski and D. A. Ackley, "Boltzmann Machines: Constraint Satisfaction Networks that Learn," Technical Report CMU-CS-84-119, Carnegie-Mellon University, Pittsburgh, PA, May 1984.

[Kirkpatrick83] S. Kirkpatrick, C. D. Gelatt and M. P. Vecchi, "Optimization by Simulated Annealing," *Science 220*, (1983), pp. 671-680.

[Larkin80]   J. H. Larkin, J. McDermott, D. P. Simon and H. A. Simon, "Models of Competence in Solving Physics Problems," *Cognitive Science 4*, 4 (1980), pp. 317-345.

[Minsky75]   M. L. Minsky, "A Framework for Representing Knowledge," in *The Psychology of Computer Vision*, P. H. Winston (ed.), McGraw-Hill, New York, NY, 1975, pp. 211-277.

[Mitchell83a]   T. M. Mitchell, "Learning and Problem Solving," *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, West Germany, August 1983, pp. 1139-1151.

[Mitchell83b]   T. M. Mitchell, P. E. Utgoff and R. Banerji, "Learning by Experimentation: Acquiring and Refining Problem-solving Heuristics," in *Machine Learning: An Artificial Intelligence Approach*, R. S. Michalski, J. G. Carbonell, T. M. Mitchell (ed.), Tioga Publishing Company, Palo Alto, CA , 1983, pp. 163-190.

[Schank77]   R. C. Schank and R. P. Abelson, *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*, Lawrence Erlbaum and Associates, Hillsdale, NJ, 1977.

[Shavlik85]   J. W. Shavlik and G. F. DeJong, "Building a Computer Model of Learning Classical Mechanics," *Proceedings of the Seventh Annual Conference of the Cognitive Science Society*, Irvine, CA, August 1985, pp. 351-355. (Also appears as Working Paper 65, AI Research Group, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign.)

[Silver84]   B. Silver, "Using Meta-level Inference to Constrain Search and to Learn Strategies in Equation Solving," Ph.D. Thesis, Department of Artificial Intelligence, University of Edinburgh, 1984.