Scientific Discovery without Search

KENNETH MAN-KAM YIP
Department of Computer Science
Yale University
P.O. Box 2158, Yale Station
New Haven, CT 06520, USA.
(yip-ken@cs.yale.edu)

Abstract

We discuss progress in the design of an apprentice program S3 whose goal is to recapitulate certain aspects of the conceptual development in an important branch of physics, hydrodynamics. The program is initially endowed with a bag of "mathematical tricks" and hydrodynamics knowledge as known in the scientific community around 1830. The program not only solves water wave problems but also acquires new concepts and theories in light of its problem solving experiences. We show how the problem solver might discover the concept of group velocity by examining the solution trace of a typical wave problem. The discovery process is based on a combination of careful observations and justifiable mathematical speculations. Little search is required.

Introduction

Can computer make significant, qualitative new findings in science? Although recent years have seen an increase in our understanding of concept formation in elementary number theory [Lenat, 1978] and empirical generalization in chemistry ([Lindsay *et al.*, 1980], [Langley *et al.*, 1987]), a predictive theory of scientific discovery has not yet been established. Very few computer programs have advanced far enough to produce new scientific results publishable in professional journals. [1] Our continuing inability to understand and simulate the intellectual mechanisms responsible for progress in science seriously limits the advancement in the design of a new generation of intelligent programs that can deal with the challenging tasks of theory and concept formation, and directing controlled experiments for new observations and theory confirmation.

Anyone who is introduced to the subject of computer discovery for the first time quickly encounters the supposition that the discovery process can be cast as searches guided by a large collection of heuristics. Intelligence, according to this view, amounts to the possession and efficient access to this collection of heuristics which constrains searches in an a priori immense rule space or concept space. Until recently, much effort has been focussed on finding heuristics that can be applied to a larger and larger class of scientific problems. However there is no guarantee that such effort will lead to a predictive theory of discovery, and the goals of this effort remain large unrealized.

There are indications that the discovery-as-heuristic-search theory might not be necessary and might even be wrong-headed in fields like physics, and that the formulation of an adequate discovery theory will require a deeper understanding of the physics and mathematics of the subject matter. Even in mathematics, practitioners do not in general systematically combine mathematical concepts and operations and then filter out those that are less interesting. Rather they attempt to prove theorems and criticize the failed attempts. New concepts and theorems develop through attempts to fill in the gaps in "proof sketches" and in successive refinements of refuted conjectures [Lakatos, 1976]. Substantial amount of mathematical knowledge is needed to do such proof analysis.

Our view is that scientific discovery, as least during normal science, is largely a combination of careful empirical observations and justifiable mathematical speculations. There is little search. Our approach requires every result obtained by the problem solver has a justification which describes the set of all simplifying assumptions and all applicability conditions of mathematical operations used for its derivation. Solutions to a good mathematical problem not only lead to new concepts but also suggest new problems to solve. This approach of concept learning is closely related to the goal-directed learning of new vocabulary terms proposed by Mitchell [Mitchell, 1983] and the dependency mechanism of "problem solving by debugging almost-right plans" by Sussman [Sussman, 1977]. Our contribution is to show how such ideas when combined with deep knowledge of physics and mathematics can be the basis of problem solving programs powerful enough not only to solve problems in a difficult branch of science

[1] See however [Tsai *et al.*, 1990; Yip, 1991].

but to increase its problem solving expertise as new concepts and theories are discovered.

Our goal is to build an apprentice system that builds up its knowledge in light of its own experience in solving increasingly difficult problems. Our programs are not just smart number crunchers nor symbolic calculators although they possess such capabilities. Rather we view them as models of what some scientists do when they are investigating physical phenomena. We want our computer programs to simulate how scientists analyze these phenomena; they should be able to make observations, propose working hypotheses, perform qualitative and heuristic analyses, make predictions, and direct controlled experiments to verify its theories.

Our objective is to develop a suite of computer programs – collectively known as S3 [2] – that recapitulate certain aspects of the conceptual development in a particular branch of physics, hydrodynamics. Given the hydrodynamical equations as known in the scientific community around 1830 and a bag of applied mathematics "tricks", these programs will be able to (i) observe and measure important quantities in wave tank experiments, (ii) simplify governing equations by asymptotic techniques, (iii) derive qualitative and quantitative predictions, (iv) compare predictions with controlled experiments, (v) create useful concepts that summarize observed features, and (vi) successively refine the equations to improve the fit between theories and experiments.

Our work is rooted in the tradition of focusing on the problem-solving behavior of articulate professionals in well-structured domains and formalizing their methods so that a computer can exhibit similar behavior on similar problems. But why do we want to do this? Why do we want to make machines do something that some good scientists already know how to do? Shouldn't we just focus on the difficult problems arising from numerical simulation of fluid flows? There are several reasons for this research effort. First, as workers in the field of Artificial Intelligence, we are interested in the epistemology of problem solving in many domains. We want to understand what sort of computer representations and reasoning mechanisms are necessary for capturing the knowledge that expert scientists have so that we can eventually build increasingly sophisticated programs that rival the effectiveness and versatility of human experts.

In addition, this kind of research is likely to have important consequences for the development of scientific and engineering curricula. Scientific and engineering curricula today make almost no effort to formally teach students reasoning and problem solving patterns that we observe in accomplished professionals. We believe it is important to teach students not only the fundamental ideas of a subject but also *how to think like scientists or engineers*. (See [Sussman and Stallman, 1975] for a similar view.)

Finally, we also see the potential of computer programs as intelligent problem solving partners for increasing the effectiveness of human scientists. By raising the conceptual level of interaction and hiding the details of "routine" problem solving, such systems will allow the human expert to focus on the more challenging tasks of problem formulation and theory formation. Consequently, the human scientist will be able to tackle new classes of problem, and not just solve some old problems faster.

Role of Mathematics and Physics in Discovery

Witness the obvious fact that great discoveries in science are in general not made by people previously ignorant of the field. It is only the master of his subject guided by a combination of empirical observations and mathematical speculations who contributes most to the growth of physical understanding.

It is well known that empirical sciences have greatly influenced the development of mathematics. Physics sets problems for which mathematical solutions are required – the invention of calculus for studying body motion, vector analysis for electromagnetics, and Riemannian geometry for relativity, just to name a few. These examples illustrate the role of physics in the invention and development of mathematical instruments. We are interested, however, in the reverse process: whether the solution of a mathematical problem leads to discovery of new physical concepts and theories.

For a physicist, mathematics is more than a calculational tool for the mere purpose of deriving consequences from the physical theory; it is a main source of concepts and principles by means of which theories can be created [Dyson, 1964]. There are two important ways mathematics can further the discovery of physical theories.

First, mathematical formulation gives precise meanings to physical terms, removing ambiguities and vague images associated with them. Second, by insisting on giving physical interpretations to mathematical quantities and operations used in the scientific theory, a scientist can sometimes predict new physics (e.g., Dirac's discovery of positron by insisting on the reality of the negative energy solutions to a relativistic equation).

There is however a down side to the role of mathematics in physics. Mathematical prejudices can set back scientific development for decades or even centuries. It took Kelper years to shake off the strong grip of perfect circles and sphere and finally postulate elliptical motion. This is where physical observations and experiments come into help. Mathematics may give meanings to physical theory but its truth cannot be determined until its predictions are consistently confirmed by experiments and accurate measurements.

[2]The acronym stands for three important families of wave: Simple progressive wave, Stokes wave, and Solitary wave.

Design of S3

Fig. 1 displays the structure of the S3 problem solver. Given a water wave problem, it simulates the wave motion in a numerical wave tank, summarizes the simulation output, and tries to give a quantitative description of the motion based on its knowledge of mathematical techniques and hydrodynamical equations. By examining the resulting solution trace, it creates new concepts and new simplified equations. The design of S3 is based on six modules: (1) knowledge base, (2) experiment, (3) interpretation, (4) modeling, and (5) analysis, and (6) learner.

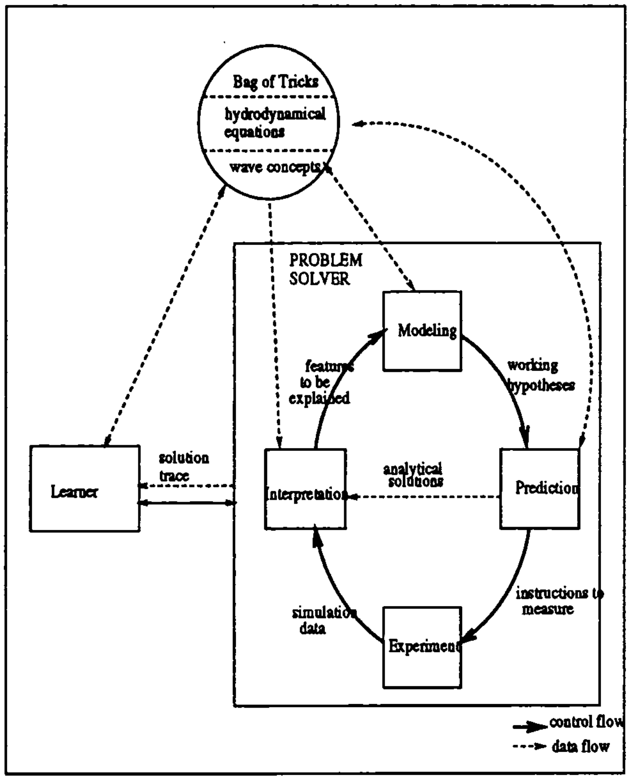

Figure 1: *The Structure of the S3 Problem Solver.*

The flow is assumed to be incompressible, inviscid, and irrotational so the Laplace's equations must hold:

$$\nabla^2 \phi = 0$$

where ϕ is the velocity potential.
There are three types of boundary conditions:
(1) A fluid particle at the free surface ($z = \eta(x, y, t)$) always remains at the surface:

$$\phi_z = \eta_t + \phi_x \eta_x + \phi_y \eta_y$$

where the subscripts denote partial derivatives.
(2) The pressure within the fluid motion must conform to Bernoulli's equation:

$$\phi_t + g\eta + \frac{1}{2}\nabla\phi \cdot \nabla\phi = 0$$

where we have omitted the effect of surface tension.
(3) On any fixed boundaries, there must be no normal velocity component:

$$n \cdot \nabla\phi = 0$$

Figure 2: *The full hydrodynamical equations are Laplace's equation with nonlinear free surface boundary conditions.*

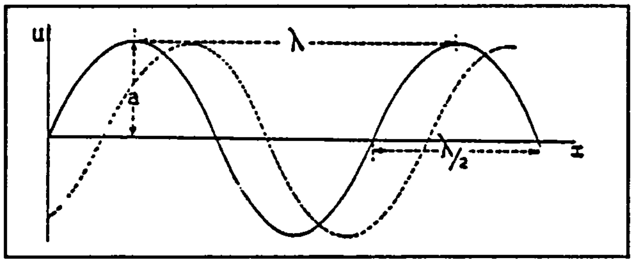

Figure 3: *Two snapshots of a simple progressive wave. The wave is traveling to the right and extends to infinity on either side. The dotted profile is what the wave looks like at a slightly later time. The wave is usually written in exponential form $ae^{i(kx-\omega t)}$. Terminology: λ is the wavelength and the wave number k is defined to be $\frac{2\pi}{\lambda}$. The amplitude is denoted by a and frequency by ω.*

The **knowledge base** consists of mathematical and physical knowledge. Its mathematical knowledge comprises a subset of modern applied mathematical techniques. Its hydrodynamical knowledge is what was known in the science community around 1830. The knowledge base has three components: (a) a bag of tricks consisting of general applied mathematical techniques (such as the Fourier method, standard asymptotic approximation techniques, and standard methods to solve differential and integral equations), (b) the full hydrodynamical equations (Fig. 2), and (c) the concept of a simple progressive wave (Fig. 3).

The **experiment module** is a numerical wave tank (see next section). A problem is given by specifying the initial disturbance and boundary conditions on the tank and the wave maker. Its output is the numerical values of the resulting wave profile as a function of space and time.

The **interpretation module** describes the outstanding features of the wave profile and compared them with predicted values, if there are any. The description uses either general geometric language (e.g., smooth curve, cusp, sine curve, etc.) or comparison with the shape of known concepts (e.g., the simple progressive wave). If the match between predicted and measured values is good, then the control is passed to the learner module. If the match is poor, it will compile the list of discrepancies and call the modeling module.

The **modeling module** checks its knowledge base for the simplest equations that have not been tried to describe the current situation. If the equation has been solved before, then the remembered solutions will be used to give a quantitative description of the wave profile. If the equation has not been solved before, then it checks if it can be solved by some standard techniques. When the equation cannot be solved exactly, it will first be simplified. These simplification techniques include linearization, asymptotic analysis based on small parameters, and assumed form of solution. The simplified equation constitutes the current working model and is sent to the next module. As the problem solving progresses, a lattice of mathematical models will be developed until a good match is obtained between the predictions from the current working model and the observation data.

The **prediction module** checks its bag of tricks for techniques to solve the equation. Even in the rare event that an explicit formula can be found for solution of the equation, it might be necessary to further simplify the solution in order to arrive at a useful interpretation of the result. Then it translates the approximate analytical solution into measurement procedures involving physical quantities such as wave amplitude, phase period, phase velocity, phase zero points, amplitude envelope zero points, etc. The predicted output is remembered and passed onto the interpretation module.

The **learner module** analyzes the solution trace produced by the problem solver. It collects constraints and conditions on the applicability of solution methods, and notices singular behaviors of the solution. It spots any novel combination of constraints and uses it to create a new concept.

An Example of Concept Discovery

Real ocean waves and its effects on beaches and ships are complicated problems. So let us begin the research into the nature of water waves by experimenting with wave tanks in laboratories. Actually we don't even use a physical wave tank; we numerically simulate the waves in a "numerical wave tank." The advantage of numerical simulation is obvious: we can measure any quantities and set up arbitrary initial conditions easily. The drawback is that only few physical effects and rather simplified boundary conditions can be reliably simulated. In fact, the subject of numerical simulation of water waves is a major research area in its own right. Nevertheless, there are still many interesting problems to be studied even for two-dimensional wave motion with fairly simple geometry.

Suppose you fill the wave tank with several inches of water and assume the tank is long enough so that wave reflection does not have to be considered. Now drop a box (a weight or something) at one end of the tank. What will you see? That depends on the dimensions of the box and the depth of water in the tank. If the box is not too large and the water is not too shallow, the initial hump of water displaced by the box will travel down the length of the tank in the form of oscillatory waves. If we record the wave height at various locations along the tank, we will see a surface profile similar to the one shown in Fig. 4. Near the box the profile is quite irregular. For larger time and farther away groups of waves appear. More interestingly, as the wave crest near the front fades out, a new wave crest begins to grow in the rear. The new wave crest will continue to grow until it reaches the front and then it fades to nothing. This whole process goes on continually. The individual wave crests inside a group seem to travel faster than the group itself.

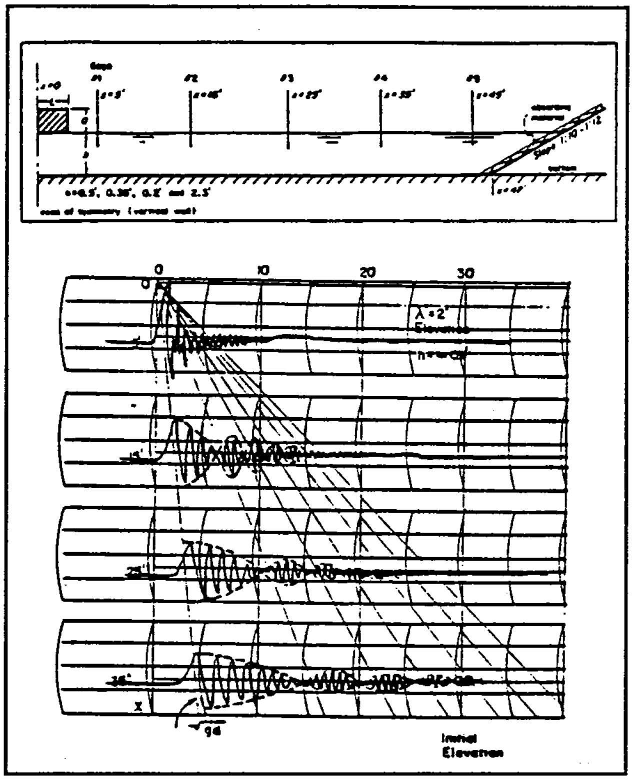

Figure 4: *Top: The wave tank. Bottom: Wave profile at various locations along the length of the tank. The vertical axis is space; the horizontal, time*

This curious behavior of water waves may be new to you. It does not happen in sound waves, nor with waves traveling down a string. If you measure the speed of the group, you will find that it is very close to half the speed of the individual wave crests inside a group. Physicists give this group speed a name, the **group velocity**. Group velocity is an important concept in many branches of physics involving wave propagation; a whole book has been written about it [Brillouin, 1960]. The task for our computer program is to discover the concept of group velocity and explain the shape of the wave profile. We should emphasize that the pattern for discovery of the concept of group velocity is by no means special. Many hydrodynamical concepts, such as solitary wave, Stokes wave, and Benjamin-Feir instability, can be discovered in a similar vein.

We recognize that the encoding of the qualitative features of the output of the wave tank simulator is not obvious; part of the research is to determine the appropriate language for such descriptions. The program first checks its knowledge base for a concept that might fit the descriptions. An expert physicist will of course have such a concept. But our apprentice program starts with only the full hydrodynamical equations and a few simple concepts about waves. It has the notion of a simple progressive wave, infinite in extent and constant in amplitude. This concept does not match what it sees. "Match" is a rather complicated idea – features must be extracted from experimental profiles and compared with those of the simple progressive wave.

Now the program must go back to the hydrodynamical equations (see Fig. 2). The full equation is too complicated to solve analytically. Approximations and simplifications must be made. In this case, a good heuristics is to linearize the equations, i.e., assuming the wave amplitude is small and product terms like $\phi_x\eta_x$ and $\phi_y\eta_y$ are small. The linearized hydrodynamical equations are easy to solve. A simple program implementing the separation of variable technique is built on top of a conventional computer algebra system like Mathematica to solve such equations. A special solution, the simple progressive wave, is found to satisfy the equations provided that its frequency ω and wave number k are related by the following relation:

$$\frac{\omega}{k} = \sqrt{\frac{g}{k}\tanh kh}$$

This relation is known as the **dispersion formula**, and the quantity $\frac{\omega}{k}$ is the **phase velocity**, i.e, the speed at which individual wave crests of the simple progressive wave travel. For the case of deep water, i.e., when kh is large, the relation simplifies to:

$$\frac{\omega}{k} = \sqrt{\frac{g}{k}}$$

since $\tanh kh \approx 1$.

The problem solver's bag of applied mathematics tricks consists of the Fourier Method which tells it to compose a general solution by superposing the special solutions with different amplitudes. The form of the solution can be expressed as:

$$\eta(x,t) = \int_{-\infty}^{\infty} A(k)e^{i(kx-\omega t)}dk$$

where $\eta(x,t)$ is the wave height at location x and time t, and $A(k)$ is determined by the initial wave profile at time zero.

Now that we have exhibited an explicit formula for the solution, we might think the work is over. But no such luck. No physical insights can be directly grasped from the formula. It shouldn't be surprising. After all, the wave evolution from the initial profile is fairly complicated as witnessed in the experimental observations. To discover the essentials of the solutions, we must delegate details into the background and look for revealing behaviors. Asymptotic approximation does just this: it concentrates on the singularities of the function. In this case, the singularity of the formula is in the limit $t \rightarrow \infty$. That it would be informative to find an approximation to the solution for large time is already suggested by the observations. The behavior in small time is dominated by the effect of initial conditions, whereas in long time the intrinsic nature of water waves will prevail.

The problem solver again checks its bag of tricks for an approximation method whose pattern of applicability matches the integral. For this integral, there is a good approximation method available, the method of stationary phase [Bender and Orszag, 1978], which applies to integral of the form:

$$\int_{-\infty}^{\infty} f(k)e^{ith(k)}dk$$

for large values of the parameter t. The basic idea behind the approximation is this: as t becomes large, the exponential term oscillates rapidly and therefore its contributions to the integral will be cancelled by the alternate positive and negative oscillations. However, cancellation does not occur at (i) points where the function $h(k)$ varies slowly, and (ii) at the endpoints (when finite). Since the integration is over an infinite interval, only points of type (i) can have significant contributions. These points are called stationary points because they are the zeroes of the derivative of h. Assuming the function h has one stationary point at k, we apply the method of stationary phase to get following result:

$$\int_{-\infty}^{\infty} f(k)e^{ith(k)}dk \sim \sqrt{\frac{2\pi}{t\mid h''(k)\mid}} f(k)e^{ith(k)+\frac{\pi}{4}sgn(h''(k))}$$

Applying the result to water wave and using the deep water dispersion formula, we get for large values of t:

$$\eta(x,t) \sim \frac{4Q}{t}\sqrt{\frac{x}{\pi g}}\sin(\frac{gt^2L}{4x^2})\cos(\frac{gt^2}{4x}-\frac{\pi}{4})$$

where Q and L are the dimensions of the box. This formula compares favorably with the observations.

Now comes the crucial step. Having achieved a good match between prediction and experiment, the problem solver steps back and look at the record of its symbolic reasoning. In particular it looks for conditions under which the solution is valid and places where the method breaks down.

First, the problem solver notices that the method is valid only when k is a stationary point satisfying:

$$\frac{d}{dk}h(k) = \frac{d}{dk}(k\frac{x}{t} - \omega) = 0$$

$$\Longrightarrow \frac{d\omega}{dk} = \frac{x}{t}$$

It therefore suggests the term $\frac{d\omega}{dk}$, which has the dimension of velocity, as an important concept to remember. This is how the concept of group velocity is discovered.

Second, looking at the approximation formula, the problem solver notices the significance of the second derivative of h at k. If it is zero, then the formula blows up and the whole asymptotic analysis fails. For water waves, it means that the result only applies to the class of waves for which $\frac{d^2\omega}{dk^2} \neq 0$. The condition becomes the definition of **dispersive waves.**

After a concept is discovered, the problem solver tries to find out how it relates to existing concepts. For example, in the case of deep water, it finds that the group velocity is given by:

$$\frac{d\omega}{dk} = \frac{1}{2}\sqrt{\frac{g}{k}}$$

That is, the group velocity is half the phase velocity of the individual wave crests. Since the formula matches observation well, the fact is incorporated into the concept of group velocity.

Work in Progress

Pieces of the S3 problem solver have been built. These include the numerical wave tank, procedures implementing standard applied mathematical tricks, and part of the mechanism for keeping track of justification records. The major unfinished business is the representation and manipulation of concepts.

How does the apprentice program "grasp" a physical concept? Once a concept is created, it must be related to physical quantities, to existing concepts, and to explanatory procedures who might employ the concept to describe new events and phenomena. For instance, to grasp the concept of energy, the apprentice must be able to (1) measure it, (2) calculate with it (using conservation principle, for example), and (3) recognize particular problem situations to which it can apply.

We will characterize a concept by three aspects: (1) operational definitions, (2) symbolic relations, and (3) applicability. Operational definitions consist of procedures for measurement. Symbolic relations refer to the mathematical definitions, laws or equations relating the concept to other concepts, and rules of calculation (e.g., arithmetic, matrix operations, differentiation or composition of function, etc.). Applicability are recognition procedures that determine the scope of application within which the concept has relevance.

Understanding of a concept is a evolving process. As new problems are solved, new concepts, new facts, or even new laws might be found. Old concepts are then revised, modified, or even discarded in light of new experiences.

Conclusion

Discovery systems are often likened to an explorer wandering in a vast search space punctuated by occasional sign posts pointing to promising directions or places to avoid. In this paper, by showing how an important scientific concept – group velocity – might be discovered by a combination of careful empirical observations and justifiable mathematical speculations based on solution traces, we recast the discovery process so that it completely avoids the search problem. The effectiveness of the proposed discovery method depends on an arduous process of mutual adjustment and feedback between the problem solver and the learner. As the problem solver learns new concepts and theories from its experiences, it becomes a better problem solver. As the problem solver increases its expertise, its opens up more new problems to solve and thus helping the learner to learn better. We often hear human professional researchers remarking that finding the right problem to solve is half the success. Our claim is that the remark applies equally well to a computer discovery system.

References

Bender, Carl and Orszag, Steven 1978. *Advanced Mathematical Methods for Scientists and Engineers.* McGraw-Hill.

Brillouin, Leon 1960. *Wave propagation and group velocity.* Academic Press.

Dyson, Freeman 1964. Mathematics in the physical sciences. *Scientific American.*

Lakatos, Imre 1976. *Proofs and refutations : the logic of mathematical discovery.* Cambridge University Press.

Langley, P.; Simon, H.; Bradshaw, G.; and Zytkow, J. 1987. *Scientific discovery: computational explorations of the creative processes.* MIT Press.

Lenat, Douglas 1978. The ubiquity of discovery. *Artificial Intelligence* 9.

Lindsay, Robert; Buchanan, Bruce; Feigenbaum, Edward; and Lederberg, Joshua. 1980. *Applications of Artificial Intelligence for Organic Chemistry.* McGraw-Hill.

Mitchell, Tom 1983. Learning and problem solving. In *Proceedings IJCAI-83.*

Sussman, Gerald and Stallman, Richard 1975. Heuristic techniques in computer-aided circuit analysis. *IEEE Transactions on Circuits and Systems* 22(11).

Sussman, Gerald 1977. electrical design: A problem for artificial intelligence research. In *Proceedings IJCAI-77.*

Tsai, Wu-ting; Yue, Dick; and Yip, Kenneth 1990. Resonantly excited regular and chaotic motions in a rectangular wave tank. *Journal of Fluid Mechanics* 216.

Yip, Kenneth 1991. Computer aided heuristic and qualitative reasoning in fluid mechanics. In *Proc. of 13th IMACS World Congress on Computations and Applied Mathematics.*