CMU-CS-83-120

Towards a General Scientific Reasoning Engine

Jaime G. Carbonell, Jill H. Larkin
Carnegie-Mellon University

F. Reif
University of California, Berkeley

13 April 1983

Abstract

Expert reasoning in the natural sciences appears to make extensive use of a relatively small number of general principles and reasoning strategies, each associated with a larger number of more specific inference patterns. Using a dual declarative hierarchy to represent strategic and factual knowledge, we analyze a framework for a robust scientific reasoning engine. It is argued that such an engine could provide (a) the ability to reason from basic principles in the absence of directly applicable specific information, (b) principled knowledge acquisition by using existing general patterns to structure new information, and (c) congenial explanation and instruction in terms of general and familiar patterns of inference.

This work was supported by the Office of Naval Research an the Army Research Institute under contract number N00014-82-C-50767.

Joint Computer Sciece and Psychology technical report

C.I.P #445

Table of Contents

1. Objectives

Natural scientists believe that there are general patterns of reasoning used repeatedly in a variety of contexts. These repeated patterns occur both among scientific principles (e.g., conservation, equilibrium), and among problem-solving strategies (e.g., decomposition, analyzing extremes). We present an analysis of how to capitalize on this apparent redundancy to build a prototype of a potentially general scientific reasoning engine.

1.1. General Inference Patterns in Natural Science

The natural sciences contain two inter-related hierarchies of knowledge (see Figure 1-1). First, scientific principles often contain knowledge of varying generality. Conservation of charge, mass, and energy all share the principle of conservation and the inference rules it entails. Similarly, conservation of macroscopic kinetic energy (without dissapation) inherits many of the properties of conservation of energy. Of course, the more specific principles also provide additional information not present in from their more general counterparts. Second, there is a similar hierarchy of problem-solving strategies. Decomposition is a very general strategy, from which more specific strategies (vector or Fourier decomposition) inherit knowledge. Third, there are in each hierarchy horizontal links reflecting analogies among similar concepts in each hierarchy. This enables latteral (partial) inheritance [3, 7]. Finally there are links between the two hierarchies in both directions. A reasoning strategy may be based upon one or more principles, and, conversely, knowing that a particular principle applies to a given problem situation may suggest one or more appripriate strategies. For instance, noting that certain types of energy are conserved in some subsystems may suggest problem decomposition along the conservation boundaries.

1.2. Importance of General Inference Patterns to AI Models

Explicitly encoding the general inference patterns of the natural sciences should help alleviate some problems inherent to present expert system architectures, and should provide particularly good support for intelligent computer assisted instruction (ICAI).

1.2.1. Help for Common Expert-System Problems

While encoding new information for a currently interesting task it is hard to recognize knowledge that might be generally applicable in very different situations. Therefore general and specific knowledge are often inextricably encoded into the same rules. Moreover, strategic knowledge is often intermingled with factual knowledge, or worse yet, convolved with "certainty factors", "subjective probabilities" or other numerical parameters. This lack of separation gives rise to common deficiencies in the following areas:

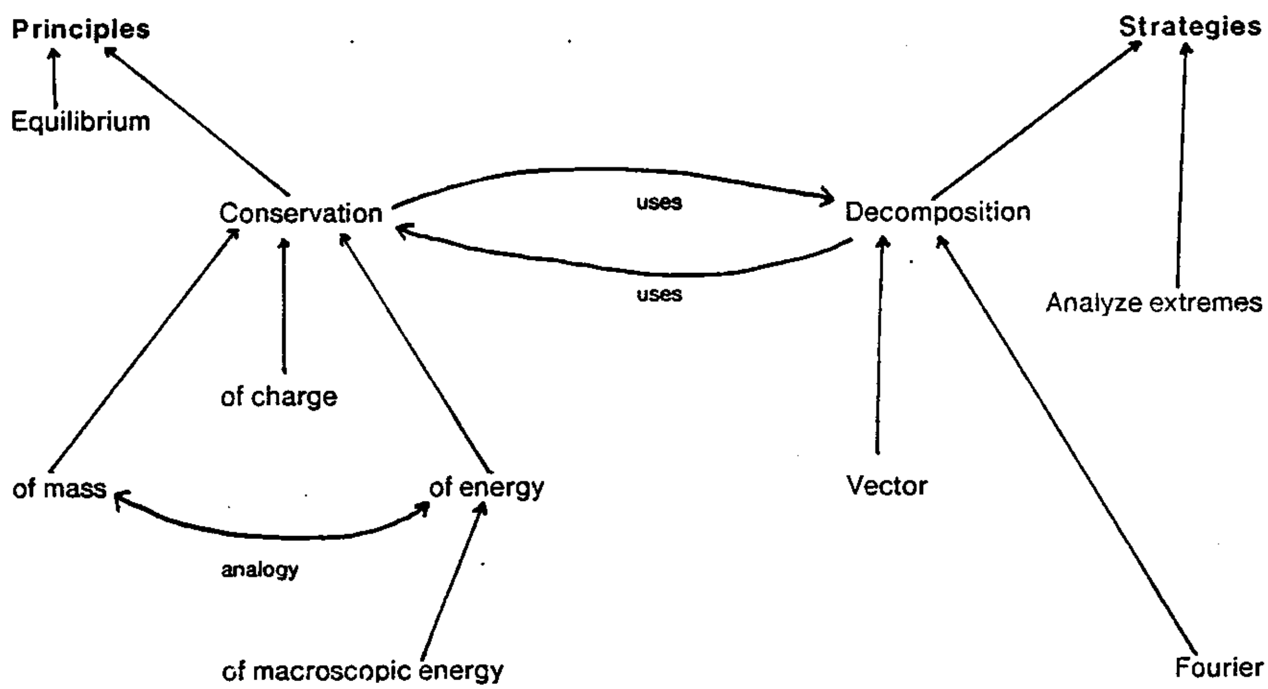

Figure 1-1: Fragment of the dual hierarchy illustrating the relation between principles and the strategies.

Portability. It is hard to apply or modicy knowledge encoded for use in one domain to another domain. Shared general knowledge or similarities among related domains cannot be exploited.

Robustness. It is hard to make an expert system behave gracefully and sensibly beyond the particular limited domain for which it was designed. In contrast graceful degradation rather than catastrophic failure is the hallmark of human expertise when faced with problems that depart somewhat from the expected.

Extensibility. It can be extremely difficult to revise past knowledge distributed among a large set of rules, rather than encapsulated at the most general appropriate level (and then inherited by lower levels of the hierarchy).

Congeniality. The reasoning of expert systems -- linear traces of hundreds or thousands of rules -- can prove extremely frustrating to humans. Moreover, when decisions depend upon combinations of certainty values, the process becomes even more obscure.

In the natural sciences there are general inference patterns that can be stated independently of other more specific knowledge. Encoding these patterns separately should help alleviate all of the preceding problems. This general knowledge can be used in extending the system to other scientific domains that rely on the same inference patterns. When taxed beyond its original limits, the system will at least have this knowledge available, enabling it to ellicit reasonable if suboptimal behavior. General patterns of inference, of the kind illustrated here, appear explicitly in human discussions of natural science (e.g., in textbooks) and there is some evidence [8] that experts use them in solving problems. Thus congeniality to humans should be improved by the explicit use of general inference patterns.

1.2.2. Implications for Intelligent CAI

The use of general inference patterns has definite benefits for ICAI systems which must satisfy particularly stringent demands for robustness, extensibility, and congeniality. Instructional systems must be robust -- deviation from particular anticipated paths is the rule, not the exception for learners. Instructional domains are always large (by AI standards), yet we must be able both to extend our systems at reasonable cost and to use instructional material in one domain to design instruction in a related domain. Lastly, the reasoning used by the expert system must be sufficiently congenial to students that they can learn to use it themselves.

In addition, encoding scientific knowledge in terms of repeated general inference patterns offers the opportunity for teaching directly a kind of reasoning that is ordinarily left implicit. Whereas textbooks and lecturers in physics and chemistry may use general inference patterns such as additivity or reduction of dimensionality, they are unlikely to make explicit the repeated use of these patterns in very different areas of science. An ICAI system with knowledge based on these patterns could make them explicit to students and illustrate their utility in solving different classes of problems. This explicit teaching could be an important educational contribution because there is evidence [8] that students may fail to solve science problems because they cannot reliably construct and use qualitative relations between scientific quantities. Many such relations (e.g., energies are conserved, forces balance) are included among the general patterns of inference.

2. The Reasoning Model

As outlined above, the reasoning model contains a linked dual hierarchy of factual or principle knowledge and strategic or problem-solving knowledge. This section describes each of these kinds of knowledge and then discusses their representation as linked declarative hierarchies.

2.1. Principles

The same patterns of inference are used repeatedly in different areas of natural science. For example, a system for reasoning about fluids at rest must include knowledge like the following:

- Pressure drops are additive along any path. Thus to determine the pressure drop between any two points, one can add the pressure drops along any sequence of segments joining those points.

- When a fluid is in equilibrium, the forces on any portion of fluid (and on objects in contact with it) are balanced. The total force in one direction is always balanced by a force of equal magnitude in the opposite direction.
- The magnitude F of a force due to a fluid on any small surface is proportional to the area A of that surface. Thus we can define a proportionality constant, called pressure p, that is equal to the ratio F/A of force divided by area, and that is independent of force and area.

If instead we were building a system to reason about chemical equilibrium, the system would need knowledge like the following:

- Concentrations are additive in any solution. Thus to determine the concentration of any species, one can add the concentrations due to any collection of sources for this species.
- When a reaction is in equilibrium, the reaction rates are balanced. The total reaction rate in one direction (e.g., producing a particular species) is always balanced by a reaction rate of equal magnitude in the opposite direction (destroying the species).
- The magnitude R of a reaction rate of a species is proportional to the concentration c of that species. Thus we can define a proportionality constant, called the reaction coefficient k, that is equal to the ratio R/c of reaction rate divided by concentration, and that is independent of reaction rate and of the concentration of that species.

Additivity, balance/equilibrium, and proportionality are examples of general inference patterns used repeatedly in natural science. Other examples are conservation and fields (unique functions of points). Many scientific principles thus form a hierarchy with a considerable amount of knowledge inherited from a relatively few general patterns of inference. By its very nature, an inheritance hierarchy allows one to encode facts or strategies at their most appropriate level of abstraction.

2.2. Strategies

Scientific strategic knowledge is also hierarchical, centering around a relatively small number of very general strategies, for example, the following:

Reduce dimensions

If an N-dimensional situation can be stated in N · K dimensions, reformulate in fewer dimensions. For example, calculating a parabolic trajectory of a projectile requires only the plane of the parabola. If all that matters is the maximum vertical height, then the one-dimensional vertical component of the motion will suffice.

Find dependencies

Consider using any applicable principle that relates unknowns to knowns or (second best) establishes dependencies among one or more unknowns.

Find analogies to solved problems

If a similar problem has a known solution, try the same solution procedure on the new problem (perhaps with minor changes to accommodate relevant differences) [5, 4]. For example, electrical circuitry was traditionally understood by analogy to steady-state fluid dynamics in closed systems.

Decompose Many apparently difficult problems become tractable when much simple instances of the problems are first solved and then composed into a full solution. For example, complex functions can often be decomposed into a sum or series of simple functions.

Analyze extremes
Analyze simple, often limiting, instances of a problem. The solution procedures for these simple problems may serve as a basis for the final solution. This method is essentially equivalent to Polya's "tame auxiliary problem" idea [14].

Perturbation analysis
Start from one or more known or easily computable steady-state solutions, then determine how the behavior of the system changes in response to minor perturbations of its independent parameters.

Exploit symmetries
Symmetry relations provide useful constraints on the form a result may take. For example, spatial symmetry requires that forces between particles depend only on the distance between them and not on the indivitual positions.

Reasoning strategies of the kind listed above are relatively few in number. However, there are potentially very many specific strategies (like the examples used above) which are essentially useful instantiations of the general strategies.

The idea and utility of a hierarchy of heuristics has been proposed by Lenat [9]. We add two new organizational principles: (a) basing this hierarchy on the available and powerful hierarchy of strategic methods in the natural sciences; (b) linking the strategic hierarchy to a hierarchy of factual knowledge as suggested in Figure 1-1. The result is a coordinated knowledge system in which strategies can call on principles and principles on strategies. Both can be used at any level of generality.

2.3. Knowledge Representation

The multiplicity of levels of abstraction among both principles and problem-solving strategies suggests representing both declaratively, allowing the more specific to inherit from the more general. This representation both allows general knowledge to be accessed naturally (through inheritance) in a variety of situations and has the following additional advantages: Using a declarative representation allows a learning system to directly access and so to modify its knowledge. Arranging this knowledge in a hierarchy allows the system to explain its reasoning at different levels of detail. Moreover, a hierarchy that clusters like strategies into groups with common parents (a parent being the more general strategy or strategies from which the various instantiations were derived) enables a system to compare solution attempts and to reason analogically [5].

Implementing this hierarchical knowledge representation across multiple scientific domains requires a flexible representation, one capable of encoding both problem-solving and principle knowledge at different levels of specificity and of bringing the appropriate knowledge to bear as new problems are encountered. In particular, we require a flexible inheritance mechanism [3, 7] that operates on structured objects representing problem solving methods and scientific principles. Lateral (or partial) inheritance is also required in analogical aspects of problem solving [15, 5]. We need to combine the better aspects of semantic networks [2, 6] and schemas (or frames) [12, 1]. The Schema Representation Language (SRL), presently in use in the Robotics Institute at CMU, presents a good basic structure in which to build declarative knowledge schemas that can be procedurally interpreted in the problem solving process.

To illustrate the implementational requirements, we describe the information encoded in the general principle of conservation, which must contain the following statement:

$$\mathrm{Conserve}(P,S) \equiv (\forall\, t_1,t_2)\ (\forall\, \{s_i\}, \{s_j\}) \sum_i P(s_i,t_1) = \sum_j P(s_j,t_2)$$

where,

- P is the property conserved (such as mass, volume, number, or energy). $P(s_i, t)$ is the value of P over subsystem s_i measured at time t.
- t_1 and t_2 are two arbitrary times at which the property P is measured
- S is the closed system where conservation applied (loosely stated, the circumscription of the world over which conservation is postulated). S is treated computationally as a set.
- $\{s_i\}$ and $\{s_j\}$ are two partitions of the set S, corresponding to the possibly different groupings of elements in S at times t_1 and t_2 respectively. For instance, if the atoms in a chemical process (such as oxidation) are rearranged into different molecules, the first partition is the substances before the reaction (e.g., at t_1) and the second partition is defined by the products of the reaction (at t_2).

Conservation of mass is represented as an instance of the conservation principle (with P = MASS) with the set of preconditions necessary for it to apply (e.g., that there be no nuclear reactions with mass-energy conversion -- $E = Mc^2$ -- in the time interval $[t_1, t_2]$). Additionally, information on *when* it may prove useful to apply this principle is represented[1], as is knowledge of which strategic reasoning

[1] We have some notions of acquiring this heuristic knowledge from experience, in much the same way that Mitchell does in his LEX system [13] -- but applied to our more general problem solving context.

methods can utilize the conservation of mass principle and in what manner.

3. Conclusion

Whereas the reasoning system discussed is yet to be built, we believe that our analysis and design suggest a powerful new framework for building reasoning engines. Most expert systems, such as MYCIN [16] and R1/XSEL [10, 11], do not exhibit the more significant properties in our design, as summarized below:

- A separation of strategic problem-solving knowledge from knowledge in the domain.
- Declarative, modifiable, representation of problem-solving strategies as well as of domain principles and facts.
- Representation at multiple levels of generality with vertical and lateral inheritance, allowing economy of representation for general knowledge and exhibiting non-brittle behavior when a problem differs somewhat from the set foreseen by the system designer.
- The ability to use more general and familiar reasoning patterns in generating an explanation to an external user.

In addition to its intrinsic AI interest, this project is relevant to questions central to philosophy of science, to psychology, and to education. For example, to what extent does the structure of science use a limited number of basic inference patterns? A model that explicitly encodes these patterns and uses them to make scientific inferences can shed some light on this question. Do human scientific experts use general patterns of inference? A model of how this process might occur can be compared with human expert data. Can students be helped to learn science by explicit teaching of general inference patterns? A general reasoning system could support explanatory ICAI systematically acquainting students with powerful re-usable patterns of inference.

4. References

1. Bobrow, D. G. and Winograd, T., "An Overview of KRL, a Knowledge Representation Language," *Cognitive Science*, Vol. 1, No. 1, 1977 , pp. 3-46.

2. Brachman, R. J., "On the Epistemological Status of Semantic Networks," in *Associative Networks*, N. V. Findler, ed., New York: Academic Press, 1979.

3. Carbonell, J. G., "Default Reasoning and Inheritance Mechanisms on Type Hierarchies," *SIGART, SIGPLAN, SIGMOD Joint volume on Data Abstraction*, 1980 .

4. Carbonell, J. G., "A Computational Model of Problem Solving by Analogy," *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, August 1981 , pp. 147-152.

5. Carbonell, J. G., "Learning by Analogy: Formulating and Generalizing Plans from Past Experience," in *Machine Learning, An Artificial Intelligence Approach*, R. S. Michalski, J. G. Carbonell and T. M. Mitchell, eds., Tioga Press, Palo Alto, CA, 1983.

6. Fahlman, S. E., *NETL: A System for Representing and Using Real World Knowledge*, MIT Press, 1979.

7. Fox, M. S., "On Inheritance in Knowledge Representation," *Proceedings of the Sixth International Joint Conference on Artificial Intelligence*, 1979 , pp. 282-284.

8. Larkin, J. H., "The Role of Problem Representation in Physics," in *Mental Models*, D. Gentner & A. Collins, eds., Lawrence Erlbaum Associates, Hillsdale, N. J., 1983.

9. Lenat, D. B., "The Role of Heuristics in Learning by Discovery: Three Case Studies," in *Machine Learning, An Artificial Intelligence Approach*, R. S. Michalski, J. G. Carbonell and T. M. Mitchell, eds., Tioga Press, Palo Alto, CA, 1983.

10. McDermott, J., "R1: A Rule-Based Configurer of Computer Systems," Tech. report, Dept. of Computer Science, Carnegie-Mellon University, 1980.

11. McDermott, J., "XSEL: A Computer Salesperson's Assistant," in *Machine Intelligence 10*, Hayes, J., Michie, D. and Pao, Y-H., eds., Chichester UK: Ellis Horwood Ltd., 1982", pp. 325-337.

12. Minsky, M., "A Framework for Representing Knowledge," in *The Psychology of Computer Vision*, P. Winston, ed., McGraw-Hill, New York, 1975.

13. Mitchell, T. M., Utgoff, P. E. and Banerji, R. B., "Learning by Experimentation: Acquiring and Refining Problem-Solving Heuristics," in *Machine Learning, An Artificial Intelligence Approach*, R. S. Michalski, J. G. Carbonell and T. M. Mitchell, eds., Tioga Press, Palo Alto. CA, 1983.

14. Polya, G., *How to Solve It*, Princeton NJ: Princeton U. Press, 1945.

15. Polya, G., *Induction and Analogy in Mathematics*, Princeton NJ: Princeton U. Press, 1954.

16. Shortliffe, E., *Computer Based Medical Consultations: MYCIN*, New York: Elsevier, 1976.