# ON THE MECHANIZATION OF ABDUCTIVE LOGIC

Harry E. Pople, Jr.
Graduate School of Business
University of Pittsburgh
Pittsburgh, Pennsylvania 15260

Abstract

Abduction Is a basic form of logical inference, which is said to engender the use of plans, perceptual models, intuitions, and analogical reasoning - all aspects of Intelligent behavior that have so far failed to find representation in existing formal deductive systems. This paper explores the abductive reasoning process and develops a model for its mechanization, .which consists of an embedding of deductive logic in an iterative hypothesis and test procedure. An application of the method to the problem of medical diagnosis is discussed.

Key Words

Planning, perceptual-models, tbduction, deduction, induction, hypothesis-formation, linear-resolution, synthesis.

## Introduction

There has been growing criticism lately concerning the methodology of artificial intelligence. While differing in the specifics of their analyses of the problem, most thoughtful observers seem to feel that the current stock of deductive machinery is simply not up to the task at hand. A deficiency often cited with regard to the present deductive procedures is their inability to represent and make use of the sort of plans, perceptual models, intuitions, and analogical reasoning processes that characterize at least some phases of virtually all human problem solving activity.[1,2,3] As remedy, various authors have suggested replacement of the existing formalisms by alternative representations such as higher logics, analog models, or general purpose programming languages.

In this paper, we offer yet another diagnosis of the problem, and propose a somewhat different solution. In our view, the principal deficiency of existing systems is their reliance on a single form of logical inference - deduction - which, though essential, is inadequate for many types of problem solving activity. Our proposed remedy is to extend the existing formal systems to embrace additional forms of inferential reasoning, especially that of abduction.

Abductive inference is one of the three fundamental modes of logical reasoning - the others being deduction and induction - characterized by Peirce[4] as the basis of scientific inquiry. According to Peirce, abduction underlies "all the operations by which theories and conceptions are engendered;" moreover, "abductive inference shades into perceptual Judgment without any sharp demarcation line between them." By his account, it is "the only kind of reasoning that supplies new ideas, the only kind which is, in this sense, synthetic." Finally, according to Peirce, abduction plays a key role in the process of analogical reasoning.

Since these are the very aspects of intelligent behavior that have been found wanting in the methodology of artificial intelligence, it has seemed to us important to explore the abductive reasoning process and to develop a model for its mechanization.

The essence of abductive inference is the generation of hypotheses, which, if true, would explain some collection of observed facts. This concept is illustrated by the following example from McCulloch[5]:

> "...abduction starts from the rule and guesses that the fact is a case under that rule: All people with tuberculosis have bumps; Mr. Jones has bumps; perhaps Mr. Jones has tuberculosis. This, sometimes mistakenly called 'inverse probability,' is never certain but is, in medicine, called a diagnosis or, when many rules are considered a differential diagnosis, but it is usually fixed, not by a statistic, but by finding some other observable sign to clinch the answer."

It is important to observe that the hypothesis formulated in an abduction is typically "quite different from anything observed, something that is in the majority of cases empirically unobservable." Thus the abductive generation of hypotheses is clearly set apart from that of induction, wherein we typically "generalize from a number of cases of which something is true, and infer that the same thing is probably true of a whole class."[6]

The various modes of logical inference can be characterized as alternative forms of argument based on the following syllogistic schema:

| | | | |
|---|---|---|---|
| I. | Major Premise | (rule) | $\forall(x)\ (P(x) \supset Q(x))$ |
| II. | Minor Premise | (case) | $P(a)$ |
| III. | Conclusion | (fact) | $Q(a)$ |

Specifically:

A) The process of <u>deduction</u> reasons from a major premise (T) and a minor premise (II) to a conclusion (III); thus, from the rule "all things that are P are also Q" and the specific case "a is a P", we conclude via deduction inference that "a is a Q."

B) <u>Induction</u>, on the other hand, comprises the process of reasoning from a case (II) and a fact relating to that case (III) to a hypothesis of the rule (I) that describes thia relation in general terms. Thus from an observation "a is a P" and the fact "a is a Q," one form of inductive inference would be to hypothesize that "perhaps all things P are also Q."

C) The process of <u>abduction</u> is characterized by the form of reasoning that takes a given fact (III) in conjunction with a rule (I) to hypothesize a case (II) that could account for or "explain" the observed fact. Thus, from an observation "a is a Q" and the rule "all things P are also Q", one might hypothesize, via abductive inference, that "perhaps a is a P."

These forms of reasoning all entail the same underlying representation of information. Whatever language is used to express the axioms and conjectures of a deductive system can also be used to express the observations and hypotheses required for induction or abduction. However, although they employ a common information structure, each of these inference schemes requires the development of a separate computational procedure for its mechanization.

Some work has been reported in the literature dealing with such procedures; these *have,* for the most part dealt with approaches to the mechanization of induction, • although there has also been report of at least one abductive system that takes a 'British Museum' approach to the generation of all possible hypotheses.9

The real problem in this area, of course, is the selective generation of hypotheses that have some reasonable prospect of being valid. Our approach to this problem has been guided by our interpretation of the traditional maxim of Occam's Razor, which requires adoption of that hypothesis which is the simplest - in the sense that it contains the smallest number of independent types of elements, adding least to what has been observed. We have taken this to mean that the preferred hypothesis is the one that reflects to the greatest extent possible a synthesis of ideas evoked from the observations. This heuristic criterion of acceptability will be discussed further in a later section, where procedures for testing and validation of hypotheses will also be described.

Before pursuing these matters in more detail, it will be necessary to develop first the framework within which hypothesis formation takes place in our system. It will be shown in the following section that the process of abductive inference can be accomplished by means of a procedure that makes use of much of the machinery already developed for deductive logic. Although our abduction procedure has been implemented using GOL[10] as the basic deductive processor, for purposes of explication we describe the procedure in the context of the better known deduction scheme of S-linear resolution.[11]

In the following, we *review* first the linear deduction method, recasting the exposition somewhat so that the intrinsic problem-reduction structure of the method is exposed. In this context, we present the notion of partial trees as a way of characterizing the AND/OR search procedure underlying this form of deduction. We then show that abductive hypothesis formation can be effected by a synthesis process that merges partial trees into complete abduction graphs.

## 1. Linear Resolution Viewed as a Problem Reduction Procedure

In this section, we provide a somewhat unorthodox description of the axioms, conjectures, and intermediate structures that arise in the course of a typical linear deduction. Our purpose is to show that linear resolution can usefully be viewed as a problem-reduction procedure (cf. Nilsson ), which uses backward search to develop direct derivations of conjunctive sets of subgoals. It is a relatively small step then to show how the same sort of search, operating on the same form of data, can be used abductively to develop direct 'explanations' of conjunctive sets of observations.

### 1.1 The Information Structure

A deduction problem is typically stated in terms of a collection of axioms A and a conjecture C which is to be deduced from the axioms. It is conventional in resolution based systems to conjoin the negation of C to the set A, convert the resulting set of expressions to quantifier free form, then to search for evidence of a contradiction that would render the expression AAC unsatisfiable. An alternative formulation of the problem would be to search for direct proof of the disjunction AVC. In this section, we discuss the procedure from yet a third point of view - namely, the demonstration of At-C; i.e., taking the expressions of A as axioms, to derive C by means of a direct proof procedure. These are all essentially equivalent views of the problem; they merely require somewhat different handling of the conversion to quantifier free form and subsequent interpretation of the information structure.

The first step in our procedure is to convert the axiom set A to conjunctive-normal quantifier-free form, using the techniques first outlined by Davis,-'-[3] Ordinarily, the conjecture *C.* would then be negated and added to this axiom set. In our case, however, as in the direct proof procedures of Cooper[14] and Maslov[15] the conjecture is transformed into disjunctive normal form on the basis of the following revised rules:

1. elimination of implication; (as described by Davis)
2. reduction of the scope of negation; (as in Davis)
3. replacement of each universally quantified variable by a skolem function having as arguments the variables of any existential quantifiers occurring before the universal quantifier.
4. elimination of existential quantifiers.
5. transformation of the resulting expression to disjunctive normal form.

The result of this process will be a disjunctive expression consisting of one or more conjunctive sets of literals.

Consider for example the following set of axioms:

1. $\forall x(P(x) \wedge Q(x) \supset R(x))$
2. $\forall w(\overline{P(w)} \supset R(w))$
3. $\forall y\, P(y)$

and the conjecture:

$$\exists z\ (R(z) \wedge Q(z)) \vee \forall s\, \overline{Q(s)}$$

Converting the axioms, we obtain the clauses:

1. $\{\overline{P(x)}, \overline{Q(x)}, R(x)\}$
2. $\{P(w), R(w)\}$
3. $\{P(y)\}$

and the conjecture is transformed into the following disjunctive form:

$$(R(z) \wedge Q(z)) \vee \overline{Q(a)}$$

where 'a' is a constant skolem function.

Each disjunct in a conjecture can be considered to be an alternative formulation of the derivation problem, as illustrated by the following:

Assume that the converted conjecture comprises n disjuncts $(D_1, D_2, \ldots D_n)$. For each $D_i$, we can formulate the expression: $(B_i \vee D_i)$, where $B_i$ represents the disjunction of all $D_j$ for $j \neq i$. This expression can then be converted to implicative form $(\overline{B_i} \supset D_i)$; $\overline{B_i}$ can be added to the axiom set A; and the problem is then reduced to derivation of the conjunctive expression $D_i$. One such alternative statement of the problem can be formulated on the basis of each $D_k$ in the conjecture. In the example above, the conjecture would give rise to the two alternatives:

I. Assuming:
1. $\{\overline{P(x)}, \overline{Q(x)}, R(x)\}$
2. $\{P(w), R(w)\}$
3. $\{P(y)\}$
4. $\{\overline{R(z)}, \overline{Q(z)}\}$

Prove: $\{\overline{Q(a)}\}$

II. Assuming:
1. $\{\overline{P(x)}, \overline{Q(x)}, R(x)\}$
2. $\{P(w), R(w)\}$
3. $\{P(y)\}$
4. $\{Q(a)\}$

Prove: $\{R(z), Q(z)\}$

These formulations of the problem are surrogates in the sense that the solution of either one of them is equivalent to solution of the original problem. Collectively, therefore, they constitute an "OR" branch in the search tree, each successor of which is an "AND" node (consisting of the conjunction of literals in the conjecture associated with that sub-problem).

Any unit clause in the axiom set is considered to be unconditional assertion of fact; all others are conditional. For example, the clause {L1 L2,...Ln} can be thought of as asserting the implication (M 3 Lj), where M is the conjunction of all Tj for j J= i. This can be interpreted aa expressing a condition (namely M) under which Li can be deduced (by means of Modus Ponens perhaps with substitution). Similar implicative expressions can be developed to assert the conditions under which each of the other literals of the clause can be derived.

These implicative expressions can be thought of as productions, or 'rewrite rules, that are the transformations by which a problem may be 'reduced' - i.e., replaced by an equivalent set of sub-problems. In the running example, if we reformulate the axioms to emphasize the conditional nature of axioms 1 and 2, the resulting set of productions would be:

1a. $\{\overline{P(x)} \leftarrow Q(x) \wedge \overline{R(x)}\}$

1b. $\{\overline{Q(x)} \leftarrow P(x) \wedge \overline{R(x)}\}$

1c. $\{R(x) \leftarrow P(x) \wedge Q(x)\}$

2a. $\{P(w) \leftarrow \overline{R(w)}\}$

2b. $\{R(w) \leftarrow \overline{P(w)}\}$

3. $\{P(y)\}$

here the symbol '←' (actually '⊂') is to be thought of as a replacement operator.

### 1.2 Interpretation of the deduction procedure

The sub-problems described in the preceding section correspond in an obvious manner to the various ways in which an initial clause may be selected from the Bet of support (where this consists of clauses arising from the conjecture) in a resolution deduction.

It should also be readily apparent from a comparison of the two procedures that the problem-reduction search process described below, which results from our interpretation of the axioms as productions, is essentially the same as the S-linear resolution method described by Loveland11.

A linear deduction is a procedure that 'solves' the conjunctive problem posed by a conjecture by systematically dealing with each of the conjuncts (sub-problems) in turn, Proceeding typically from right to left, each literal of the conjecture is processed in the following manner:

First, the literal is 'matched' (using Robinson's unification procedure 16*)* against each unit clause and against the literal to the left of '-' in each production of the axiom set. Every successful match gives rise to an alternative successor of the starting conjecture.

Any match that succeeds with respect to a unit (unconditional) axiom provides a solution to the sub-problem posed by the subject literal (which is true 'by assumption'). Thus this term may be deleted from the list of literals of the successor conjecture (after appropriate instantiation of any bindings of variables established in the matching process).

For those matches that succeed with respect to the left hand side of a production (conditional axiom), the successor conjecture is formed by replacing the subject literal by the appropriately instantiated collection of literals comprising the right hand side of that axiom.

The set of successors constructed as above, taken collectively, constitute an 'OR' branch in the search tree - each element of which is a surrogate for the original conjecture.

If a successor node is null (i.e., contains no further sub-problems), the portion of the search tree by which it was discovered constitutes a solution to the original problem. A non-null node with no successors cannot be solved.

It sometimes happens that while developing the search tree for a particular literal (say S), a sub-goal is generated that is the negation of that which is to be proved (in this case: s). In such cases, it can be shown that, provided that all other conjunctive subgoals in the tree are satisfied, a proof of S has been established by reductio ad absurdura, and that the contradictory subgoal can be deleted from further consideration, immediately upon its generation. This refinement to the basic problem-reduction scheme can be seen to be related to the subsumption condition which is required for completeness of linear resolution strategies.

To illustrate the linear deduction procedure, we give below a derivation of version II of the previous example.

The starting form of the conjecture is:

{R(z), Q(z)}

To show that this is derivable we must now show that each literal is either an instance of an axiom (in which case it can be deleted from the set of sub-problems) or else is implied by some conjunction of literals, each of which is derivable.

In order to prove 'Q(z)' we can use Axiom 4 . By substituting 'a' for the existentially quantified variable V of the conjecture, we Bee that the right-most sub-problem In the set (R(a), Q(a)} is solved trivially, and we are left with the successor problem (R(a)I.

Since the literal 'R(a)' matches two of the conditional axioms, there will be two successor nodes:

a: {P(a), Q(a)}, and

As candidate hypotheses, our procedure selectB those literals that are abandoned by deduction, in the sense that they fail to generate successor nodes. A candidate hypothesis is entertained seriously if it arises in the partial search trees that are developed on the basis of two or more of the data making up the conjunctive observation being explained; the more data accounted for by a candidate hypothesis, the more highly regarded it is by the abduction processor. This is our Implementation of the maxim of Occam's Razor,

Data contribute to *a* hypothesis by means of an operation that we refer to as synthesis, which is actually analogous to the operation of factoring in resolution. This process of synthesis, or factoring across partial trees, is at present the only mechanism by which abductlve hypotheses are generated in our system. Other useful heuristic criteria will undoubtedly be forthcoming, but what form these may take is still an open question.

The combined abduction/deduction procedure is illustrated by the following example, based on our application of the method to the problem of clinical diagnosis.

Assume that we have available a pathophysiological data base, structured along the lines suggested in the preceding Bection, that includes the following sort of axioms: (note that the replacement operator "-" should rightfully be interpreted as "could be caused by" rather than as "implied by" in this context.)

Axioms

A1l {chills *■ presence (P.S) A Inflammatory (P))

A2 {pain (R) *■ presence (P,S) *A* Located-in (S.R)}

A3 {inflammatory (abscess)}

A4 {located-in (liver, right-upper-quadrant)}

A5 {jaundice ■ presence (P, liver)}

Here, we have indicated only those conditional axioms that are considered relevant for the application at hand. Other variations of the conditional axioms Al and A2 which are not useful for purposes of diagnosis are not explicitly displayed. (In our work, this form of information structure is implemented by means of the GOL DELTA function (cf, laner^-[7].)

Assume that the following conjunctive set of symptoms has been observed:

{chills, pain (right-upper-quadrant)}

The diagnostic task 1B to formulate some hypothesis of the form:

{presence (P,S)}

where P is some abnormal process, and S is some structure such that the presence of P at S could account for the observations given.

Using the synthesis criterion suggested above to control the adduction of hypotheses (which are displayed to the left of the vertical bar in lines 4-6 below) a diagnostic model of the pathology can be derived as follows:

<table>
<tr><td>1.</td><td>{|chills, pain (right-upper-quadrant}</td><td>observation</td></tr>
<tr><td>2.</td><td>{|presence (X,Y), located-in (Y, right-upper quadrant), chills}</td><td>1, A2</td></tr>
<tr><td>3.</td><td>{|presence (U,V), inflammatory (U), presence (X,Y), located-in (Y, right-upper-quadrant)}</td><td>2, A1</td></tr>
<tr><td>4.</td><td>{presence (U,V)|inflammatory (U), located-in (V, right-upper quadrant)}</td><td>3, synthesis</td></tr>
<tr><td>5.</td><td>{presence (U, liver)|inflammatory (U)}</td><td>4, A4</td></tr>
<tr><td>6.</td><td>{presence (abscess, liver)|}</td><td>5, A3</td></tr>
</table>

Except for the introduction of a tentative hypothesis in line 4 above, and the fact that literals are processed on a first-in-first-out basis, this derivation follows the usual form of a linear deduction *as* previously described. The abduction task is completed when no remaining subgoals occur to the right of the vertical bar; any terms occurring to the left of the bar at that juncture constitute the basis for a diagnostic model.

It may be of some interest to follow the course of development of the diagnosis in the case illustrated above. In line 4, where the synthesis step occurs, what is hypothesized is that some as yet unknown process affecting an unspecified structure is responsible for both of the observed symptoms. The instantiations that then take place in the final two steps of this derivation entail contributions from the search trees developed from each of the root nodes; thus the proposed diagnosis represents a true synthetic inference.

A hypothesis developed on the basis of such a procedure is, in general, not a unique explanation of the data, and the problem then becomes one of discrimination, among contending diagnoses. This phase of the problem entails the *use* of deductive, as well as abductive logic. Once a diagnostic model has been proposed, it can be used to generate predictions of additional unreported manifestations of the assumed pathology. Thus, for example, in the case illustrated above, the presumption of a liver abscess can be used to <u>deduce</u> the prediction of jaundice (on the basis of axiom 5). Such predictions can then be subjected to empirical verification. Any new observational data can then be fed back via the abduction procedure through another iteration of the cycle - giving rise, perhaps, to revised hypotheses that may in turn generate new predictions, leading to new observations, and so forth.

3. <u>Discussion</u>

The iterative hypothesis and test procedure outlined in the preceding section is, of course, one of the basic paradigms *of* human cognitive, and problem solving activity. The performance of any task that is basically synthetic in nature entails the use of this procedure. We would include tn the list of such tasks those episodes of comprehension and planning that arise in the course of any real problem-solving act.

A number of studies are presently being conducted at Pittsburgh to evaluate and further develop this concept. These include the biomedical theory formation project which has been reported previously,[18] and a study of planning in organic synthesis, recently completed by Smith, [19]which uses abduction to develop strategies for deductive planning of a synthesis. Another interesting project, now nearing completion, uses abductive logic to develop perceptual models for use in natural language comprehension; another exploits the planning and model-building capabilities of the system in the automatic programming task environment.

There would appear to be a number of ways to go from here. As practical systems of higher logic become available, we may want to raise the sights of the abduction processor accordingly; our GOL implementation already has limited 'higher level' capabilities that have been found extremely useful in some applications. Certainly, the procedure should be extended to include a mechanization of induction as well as the other two forms of inference: reasoning by analogy, according to Peirce, consists of an induction and an abduction followed by a deduction. Additional applications that may shed further light on the processes involved, such as procedures for strategic planning in theorem proving, should be investigated.

While we have in no way begun to exhaust the questions and issues that this new methodology raises, results from our initial forays *into* the field have encouraged us to continue with the task at hand.

<u>Bibliography</u>

1. Sloman, A., Interactions between philosophy and artificial intelligence: the role of intuition and non-logical reasoning in intelligence. Artificial Intelligence (2), (1971), pp. 209-225.

2. Anderson, B. and Hayes, P., An Arraignment of Theorem Proving or The Logicians' Folly, Dept. of Computational Logic, University of Edinburgh. DCL Memo #50, 1972.

3. Dr-eyfus, H.L., <u>What Computers Can't Do</u>. Harper &. Row, 1972.

4. Peirce, C.S., <u>Collected Papers of Charles Sanders Peirce</u>. C. Hartshorne, P. Weiss, and A. Burks (eds.) 8 vols., Cambridge, Mass., 1931-1958, especially Vol. II, p. 272-607.

5. McCulloch, W.S., "What's in the brain that ink may character?" Presented at the International Congress for Logic, Methodology, and Philosophy of Science, 1964. Reprinted in: W.S. McCulloch, <u>Embodiments of Mind</u>, The MIT Press, 1962.

6. Goudge, T.A., <u>The Thought of C.S. Peirce,</u> Dover, 1969.

7. Plotkin, G.D., A further note on inductive generalization. In: Mach. Int. 6, Meltzer, D. and Michie, B. (eds.), Amer. Elsevier, 1971,

8. Popplestone. R.J., An experiment in automatic induction. In; Mach. Int. 5, Meltzer, B. and Michie, D. (eds.), Amer. Elsevier, 1970.

9. Morgan, C.G., Hypothesis Generation by Machine. Artificial Intelligence 2 (1971) pp. 179-187.

10. Pople, H.E., Jr., A goal oriented language for the computer, in _Representation and meaning - Experiments with information processing systems_. B. Simon and L. Siklossy (eds.), Prentice-Hall, 1972.

11. Loveland, D.W., A unifying view of some linear Herbrand Procedures. JACM, Vol. 19, 2, April, 1972.

12. Nilsson, N.J,, _Problem Solving Methods in artificial Intelligence_. McGraw-Hill, New York, New York 1971.

13. Davis, M., Eliminating the irrelevant from mechanical proofs, Proc. Symposia in Applied Maths, Vol. XV, AMS. (Ed. Metropolis, et. al.) 1963, pp. 15-30.

14. Cooper, D.C., Theorem Proving in Computers, in _Advances in Programming and Non-Numerical Computation_. L. Fox (ed.), Pergamon Press, 1966. pp. 155-182.

15. Maslov, S.Ju., Proof-search strategies for methods of the resolution type. In: Mach, Int. 6, Meltzer, B. and Michie, D. (eds.) Amer. Elsevier, 1971.

16. Robinson, J.A., A machine-oriented logic based on the resolution principle. J. Ass. Comput. Mach., 12, pp. 23-41.

17. Isner, D.W., An inferential processor for interacting with biomedical data using restricted natural language. Proc. SJCC, 1972.

18. Pople, H. and Werner, G., An information processing approach to theory formation in biomedical research. Proc. SJCC, 1972.

19. Smith, G., Jr. Mechanized procedures for strategic planning in synthetic organic chemistry. Doctoral Dissertation, Dept. of Chemistry, Univ. of Pittsburgh, 1973.