ARP 369

Printing Requisition/Graphic Services

Dept. No. 67308

1. Please complete unshaded areas on form as applicable. (4 part no carbon required).
2. Distribute copies as follows: White, Canary and Pink—Printing, Arts Library or applicable Copy Centre Goldenrod—**Retain.**
3. On completion of order, pink copy will be returned with printed material. Canary copy will be costed and returned to requisitioner, **Retain as a record of your charges.**
4. Please direct enquiries, quoting requisition number, to Printing/Graphic Services, Extension 3451.

Title or Description
Mechanical Hypothesis Formation - CS-78-33

Date	**Date Required**	**Account**
Feb. 19/80	ASAP - Please	126-4100-01

Signature	**Signing Authority**
J. Macpherson	T. Pietrzykowski

Department	**Room**	**Phone**	**Delivery**
Computer Science	5181	3112	☐ Mail ☐ Pick-up ☐ Via Stores ☐ Other

Reproduction Requirements ☑ Offset ☐ Signs/Repro's ☐ Xerox
Number of Pages 16 **Number of Copies** 50

Type of Paper Stock ☑ Bond ☐ Book ☑ Cover ☐ Bristol ☐ Supplied — Supplied Fronts & Backs

Paper Size ☑ 8½ x 11 ☐ 8½ x 14 ☐ 11 x 17

Paper Colour ☑ White ☐ Other **Ink** ☐ Black

Printing ☑ 1 Side ☐ 2 Sides **Numbering** to

Binding/Finishing Operations ☑ Collating ☐ Corner Stitching ☑ ~~Stapling~~ Stitched ☐ Tape ☐ Plastic Ring ☐ Perforating

Folding Finished Size 3 STAPLES **Cutting Finished Size**

Special Instructions
Fronts & backs supplied

Cost: Time/Materials	Fun.	Prod.Un.	Prod. Cl.	Opr. No.	Mins	Total
Signs/Repro's	1					
Camera	2					
Correcting & Masking Negatives	3					
Platemaking	4					
Printing	5					
Bindery	6		93-15			
7 13						
Sub. Total Time						
Sub. Total Materials						
Prov. Tax						
Total						

Film Qty Size
Plates Qty Size & Type
Paper Qty Size
Plastic Rings Qty Size
Outside Services

Graphic Services May 68 482-2

Mechanical Hypothesis Formation*

by

T. Pietrzykowski
Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada

Research Report CS-78-33

December 1979

* This paper was supported by NRC Grant A5267.

Abstract

In this paper we deal with the problem of the mechanization of establishing a cause for given events. A definition of mechanical cause in terms of predicate logic is provided. Various aspects of this definition are discussed, and a number of restricting conditions are suggested (in order to satisfy some essential intuitions behind the concept of a cause). An extensive example of the application of our concepts to the problem of pattern recognition is also presented.

1. Introduction

We may view the area of problem solving as answering various kinds of questions. The first type of question is: "Is it true that ..." and we are on the ground of mechanical theorem proving. The second group of questions begins with the key words: "Who", "Where", "What", "When" or "How many (or much)". They lead us towards the question-answering systems where an individual answer is a constant representing a person, place, thing, time or number respectively. The third group involves the question: "How to" (specification of the goal) and the answer should be a description of a plan (or program). Obviously this is the domain of mechanical plan formation (or if one prefers: automatic programming). In all the above-mentioned situations, mechanical deduction is employed and there are numerous, well known publications investigating this, so it does not seem to be necessary to go into any further details.

However, there is another type of questions, which we encounter very frequently in every walk of life, starting with the word "Why". Apparently such questions ask for a reason, cause, or, in other words, a hypothesis which can explain some events which have been observed. Such hypothetical cause combined with our knowledge of the subject matter should make it possible to prove the existence of above events by purely logical reasoning. Interestingly this situation is different from the three cases discussed above: there the answer had a flavour of certainty while here we are making only a hypothesis and we are never able to tell whether this hypothesis is a real cause. But in spite of such philosophical difficulties we do make such hypotheses in nearly every moment of our conscious life where the range of subjects extends from the kitchen sink to quantum mechanics. Such overwhelming popularity of hypothesis formation seems to provide sufficient motivation to try to mechanize some aspects of these processes. Surprisingly however, with the exception of the research of Shortliffe [9] which will be discussed later, not much has been published (for these and related topics, see [1], [3], [4], [7] and [10]). In Chapter 2 of this paper a formal definition of mechanical cause is given. This chapter also contains a discussion of conditions necessary to restrict this definition in order to remove trivialities and capture the essential features of many possible mechanical causes. There is presented an open problem.

Chapter 3 provides an extensive description of an example of the application of ideas of mechanical causation to the recognition problem. On the basis of this example, some more general results about the relations between the structure of knowledge, events and causes are discussed.

2. Mechanical causation

2.1 Preliminaries

A wff with matrix being a conjunction of literals will be called conjunct. If B is a set of wffs then $\bigwedge B$ denotes the conjunction of all elements of B. If a,b are wffs and K is a set of wffs then $a \vdash_K b$ means that $(a \wedge \bigwedge K) \supset b$ is valid. We assume existence of special wffs: TRUE and FALSE with usual properties.

2.2 Mechanical cause

2.2.1 Definition: Let K be a set of wff called Knowledge base whose elements are called descriptors and e be a conjunct called event.

A conjunct c is called a cause of e under K (cause of e or cause) iff $c \vdash_K e$.

c is minimal iff for all causes x of e, $\vdash c \supset x$ implies $\vdash c \equiv x$.

c is basic iff for all causes x of c, $\vdash x \supset c$.

c is relatively nontrivial

(i) $\vdash c \supset e$ does not hold

(ii) there is a set $K' \subset \{x' \mid x'$ is an instance of $x \in K\}$ such that c is a cause of e under K' and $\bigwedge K' \wedge c$ is satisfiable.

c is absolutely nontrivial if the condition (ii) holds for $K' = K$.

2.2.2 Examples and comments

Minimality Let $K = \{\exists x P(x) \supset q\}$ and $e = q$. Then $\forall x P(x)$, $\exists x P(x) \wedge r$ and $\exists x P(x)$ are causes of e but only the last one is minimal.

Basicness Let $K = \{p \supset q, q \supset r\}$ and $e = r$. Then both q and p are causes of e but only p is basic.

Triviality Let $K = \{p\}$ and $e = q$. Then any cause of e is trivial. But if $K = \forall x(P(x) \supset P(f(x))), -P(a), -P(ff(a)) \supset q\}$ and $e = P(f(a)) \wedge q$. Clearly there is no absolutely nontrivial cause but $P(a) \wedge -P(ff(a))$ is relatively nontrivial with the set $K' = \{P(a) \supset P(f(a)), -Pff(a)) \supset q\}$. Note that the set K' could not be a subset of K.

2.3 Finiteness of the set of causes which are minimal and basic

We shall identify wffs which are mutually equivalent. For $K = \{\forall x(R(x) \supset P(x)), \forall y(P(f(y)) \supset P(y))\}$, $e = P(a)$. The set minimal and basic causes: $\{R(f(a)),\ldots,R(f^{(i)}(a)),\ldots\}$ is clearly infinite. However, it seems to be important that f is a function symbol which is not a Skolem one because if we define $K' = \{\forall x(R(X) \supset P(x)), \forall y(\exists z P(z) \supset P(y))\}$ the only cause for e is: $\forall x R(x)$. Interestingly note that if we Skolemize K' it looks very similar to K.

Open problem Let K be a set wffs which do not contain any function symbols. Is the set of minimal and basic causes under K always finite for any event e?

3. Application to the identification problem

In the previous chapter we provided the basic definitions of mechanical hypothesis formation. However the abstract character of these definitions, perhaps makes it difficult for the reader to see how to apply them in practical situations. Therefore we present here an example using our concepts of mechanical causation to solve a problem of identification. The example used will be rather concrete, however the techniques applied have more general character.

3.1 Description of the problem

Let us assume that we are facing the problem of mechanical recognition of the letters of the alphabet. We have at our disposal a device called the <u>analyzer</u> which accepts as input a figure and outputs certain elementary logical information. In this case let us assume that the analyzer is able to recognize the following <u>types</u> of characteristic points of an input picture:  : <u>end</u>,  : <u>angle</u>,  : <u>triple</u> and  : <u>quadruple</u>. It should be noted, for the sake of general interest that the types of points, with the exception of angle, are invariants of topological transformation. Angle, however, is an invariant of differentiable transformations. The output information of the analyzer could be divided into the following categories:

(α) discovery of at least one point of a given type:

$$\exists x\ P(x)$$

where P is called <u>characteristic predicate</u> and can be one of the following: <u>end</u>, <u>ang</u>, <u>tri</u> or <u>qua</u> which correspond respectively to the type of a characteristic point.

(β) discovery of certain number of distinct points of a given type.

For two points it will be:

$$\exists x\ \exists y(P(x) \wedge P(y) \wedge x \neq y)$$

where '=' is the equality predicate, for three points:

$$\exists x\ \exists y\ \exists z(P(x) \wedge P(y) \wedge P(z) \wedge x{\neq}y \wedge x{\neq}z \wedge y{\neq}z)$$

The reader can easily generalize this type of message for any number of points.

(γ) statement that the given input contains <u>only</u> certain types of points:

$$\forall x(P_1(x) \vee P_2(x) \ \dots \vee P_k(x))$$

where P_i $(1 \leq i \leq k)$ are characteristic predicates. For example, the statement

$$\forall x(\mathrm{end}(x) \vee \mathrm{tri}(x))$$

says that in the picture there are only end points and triple points.

The reader should note that the above classification of outputs applies to a wider range of problems than recognizing pictures through their characteristic points.

The final output of the analyzer, the <u>event</u>, is a conjunction of statements of the kind described above.

The events are used to recognize what kind of picture has been analyzed. For the sake of simplicity we shall assume that there are only five possible kinds of pictures:

A, B, C, D, E

The purpose of the following discussion is to find out what events and what kind of knowledge base are needed in order to uniquely recognize given letters.

3.2. Existential descriptors

In this section we shall introduce the simplest of elements of the knowledge base, namely existential descriptors. The general format of such a descriptor is:

$$LP \supset \exists x_1 \ldots \exists x_k \; (P_1(x_1) \wedge \ldots \wedge P_k(x_k))$$

where LP is a letter predicate and $P_i (1 \leq i \leq k)$ are characteristic predicates. The letter predicates are 0-argument predicates: a, b, c, d or e, corresponding in obvious ways to the respective pictures. We shall assume that our knowledge base contains the following existential descriptors:

(1) $a \supset \exists x \; \exists y \; \exists z \; (\mathrm{ang}(x) \wedge \mathrm{end}(y) \wedge \mathrm{tri}(z))$

(2) $b \supset \exists x \; \exists y \; (\mathrm{ang}(x) \wedge \mathrm{qua}(y))$

(3) $c \supset \exists x \; \mathrm{end}(x)$

(4) $d \supset \exists x \; \mathrm{ang}(x)$

(5) $e \supset \exists x \; \exists y \; \exists z \; (\mathrm{ang}(x) \wedge \mathrm{end}(y) \wedge \mathrm{tri}(z))$

The meaning of the above descriptors is rather obvious: for example, (1) says that the picture A contains at least one angle, one end and one triple point.

Let us examine how our knowledge base works on example of some events. Let

$$E = \exists x \exists y (\mathrm{end}(x) \wedge \mathrm{ang}(y)).$$

In this case the set $\tilde{S}$ of non-trivial, minimal and basic causes (to which we shall refer in future as n-m-b causes) is as follows:

$$\tilde{S} = \{a,\ b \wedge c,\ c \wedge d,\ e\}$$

clearly there are four possible causes: the picture could be A or E or B and C together or C and D together. In this situation the simultaneous appearance of two letters is impossible but there are situations, for example, identifying illness (cause) on the basis of symptoms (events), when such conjunction is legitimate: like two illnesses occuring together. So we do not yet reject any of the above possibilities. In order to obtain more resolving power we have to make our knowledge base more sophisticated, which is described in the following section.

3.3. Negative descriptors

The role of these descriptors is to specify types of characteristic points which a given picture does not contain. The general form of negative descriptors is

$$LP \supset \forall x(\neg P_1(x) \vee \ldots \vee \neg P_k(x)) \ .$$

The meaning of the symbols is the same as in 3.2. We shall assume that our knowledge base has the following negative descriptors:

(6) $a \supset \forall x(\neg qua(x))$

(7) $b \supset \forall x(\neg end(x) \wedge \neg tri(x))$

(8) $c \supset \forall x(\neg ang(x) \wedge \neg tri(x) \wedge \neg qua(x))$

(9) $d \supset \forall x(\neg end(x) \wedge \neg tri(x) \wedge \neg qua(x))$

(10) $e \supset \forall x(\neg qua(x))$

Now for the same event E as in 3.2 the set $\tilde{S}$ of n-m-b causes is reduced to:

$$\tilde{S} = \{a, e\}$$

It is interesting to notice that it is the condition (ii), nontriviality that eliminates the conjuncts $b \wedge c$ and $c \wedge d$ from the n-m-b causes. Let us analyze it in the case of $b \wedge c$. Assuming $C = b \wedge c$, after the skolemization of (7) and (3) we find the following clauses are a part of $C \wedge \Lambda K$:

$$\neg b \vee \neg \text{end}(x) \qquad \text{from (7)}$$

$$\neg c \vee \text{end}(\alpha) \qquad \text{(3)}$$

$$\left.\begin{matrix} b \\ c \end{matrix}\right\} C$$

where α is a Skolem constant. Obviously the above set is unsatisfiable which makes $C \wedge \Lambda K$ unsatisfiable and contradicts the nontriviality condition. In similar manner $c \wedge d$ is eliminated. It is clear that the negative descriptors are a very useful tool to improve the selectivity of the knowledge base. However, they have an important drawback: in a situation where we have a large number of characteristic features (in our case they are the characteristic points), the negative descriptors may become very large (for example, when we deal with hundreds of symptoms in medical diagnosis). In order to handle such a situation we have to choose, among all the features which a given object does not have, only a few, and record these as a negative descriptor. It motivates us to introduce another type of descriptors which will play a similar role.

3.4. Excluding descriptors

Here we present a type of descriptors which state that a given picture possesses no other characteristic points than those specified. We achieve it by the following formula:

$$PL \supset \forall x (P_1(x) \vee \ldots \vee P_k(x))$$

where the meaning of the symbols used is the same as in 3.2.

In the case of our example we shall add to K the following:

(11) $a \supset \forall x\ (\mathrm{end}(x) \vee \mathrm{ang}(x) \vee \mathrm{tri}(x))$

(12) $b \supset \forall x\ (\mathrm{ang}(x) \vee \mathrm{qua}(x))$

(13) $c \supset \forall x\ \mathrm{end}(x)$

(14) $d \supset \forall x\ \mathrm{ang}(x)$

(15) $e \supset \forall x\ (\mathrm{end}(x) \vee \mathrm{ang}(x) \vee \mathrm{tri}(x))$

We also add to K the following statements which show that different characteristic points are indeed distinct:

(16) $\forall x\ (\mathrm{end}(x) \supset (\neg \mathrm{ang}(x) \wedge \neg \mathrm{tri}(x) \wedge \neg \mathrm{qua}(x)))$

(17) $\forall x\ (\mathrm{ang}(x) \supset (\neg \mathrm{tri}(x) \wedge \neg \mathrm{qua}(x))$

(18) $\forall x\ (\mathrm{tri}(x) \supset \mathrm{qua}(x))$

The general form of such statements is:

$$\forall x\ (P_i(x) \supset (\neg P_{i+1}(x) \wedge \ldots \wedge \neg P_n(x)) \qquad 1 \le i \le n-1$$

where n is the number of <u>all</u> characteristic features (in our case $n = 4$).

We shall now show that (12) and (16) combined will play an analogous role as (7) in excluding $b \wedge c$ as a trivial cause. If $C = b \wedge c$ then the following is a part of $C \wedge \Lambda K$:

$\neg b \vee \text{ang}(x) \vee \text{qua}(x)$ (12)

$\neg \text{end}(x) \vee \neg \text{ang}(x)$ from (16)

$\neg \text{end}(x) \vee \neg \text{qua}(x)$

$\neg c \vee \text{end}(\alpha)$ (3)

b

c } C

Again it is easy to notice that the above set is unsatisfiable. Instead of (3) we could use (13) which is stronger, but we wanted to show that $b \wedge c$ would be excluded even in the absence of (13). In order to eliminate $c \wedge d$ we could use: (3), (14) and (16).

3.5. Quantitive descriptors

In the previous sections we discussed ways of eliminating from among n-m-b causes such elements as $b \wedge c$ or $c \wedge d$. However, the set of n-m-b causes still consists of two elements, a and e so we do not know if it is a letter A or E which the analyzer tested. Problems here are twofold: first in our knowledge base, all descriptors referring to A and E, that is: (1) and (5), (6) and (10) as well as (11) and (14), present the same properties. Secondly, the data given in the event $E = \exists x \ \exists y(\text{ang}(x) \wedge \text{end}(y))$ does not provide enough information to distinguish between letters A and E even under the most sophisticated knowledge base. Therefore we have to both extend the knowledge base and make the events more informative.

In order to achieve the first, we shall introduce quantitive descriptors. Essentially they are of two types: one says that a given type of picture contains no less than and another, a stronger version, that it contains exactly a certain number of characteristic points. The general form for the first type is:

$$PL \supset \exists x_1 \ldots \exists x_k (P_\ell(x_1) \wedge \ldots \wedge P_\ell(x_k) \wedge \bigwedge_{\substack{1 \le j \le k \\ i < j \le k}} (x_i \ne x_j))$$

and it says that the picture PL contains no less than k points of P_ℓ kind.

For the second type we have:

$$PL \supset \exists x_1 \ldots \exists x_k [P_\ell(x_1) \wedge \ldots \wedge P_\ell(x_k) \wedge \bigwedge_{\substack{1 \le i \le k \\ i < j \le k}} (x_i \ne x_j) \wedge \forall y (P_\ell(y) \supset$$

$$\supset (y = x_1 \vee \ldots \vee y = x_k))]$$

which states that PL contains exactly k points of the P_ℓ type.

We augment our knowledge base with two such descriptors; their meaning seems to be self evident:

(19) $a \supset \exists x\, \exists y\, (tri(x) \wedge tri(y) \wedge x \ne y)$ (first type)

(20) $e \supset \exists x\, (tri(x) \wedge \forall y\, (tri(y) \supset x = y))$ (second type)

Obviously we could introduce many more but for the sake of brevity in the presentation we will use only first of the above ones.

As we mentioned before, in order to utilize the power of such descriptors we need more information from events. So far we have used only events consisting of outputs of category (α) (see 3.1); and now we will add information of types (β) and (γ). Let us consider the event

$E = \exists x\, \exists y\, (tri(x) \wedge tri(y) \wedge x \ne y) \wedge \exists u\, ang(u) \wedge end(v)$

which consists of the outputs of the type (β) and (α). (We abandoned here putting all quantifiers to the left for the sake of clarity.) The set of minimal and basic causes for E is: $\{a, e \wedge \forall x\, \exists y$ ~~(tri(y) ∧ x ≠ y)}.~~

The first cause is apparently non-trivial. However, if we skolemize the second one and combine it with description (20) we obtain:

$$\left.\begin{array}{l} e \\ \mathrm{tri}(\varphi(x)) \\ x \neq \varphi(x) \end{array}\right\} C$$

$$\neg e \vee \neg \mathrm{tri}(z) \vee \psi = z \qquad \text{(from 20)}$$

where φ and ψ are skolem symbols, which is clearly unsatisfiable and contradicts the nontriviality condition.

Now we shall discuss a problem of distinguishing between pictures such that the set of characteristic points of one of them is a subset of that of another: for example, distinguishing D from E. Let us suppose that it is the letter D which is analyzed. Using outputs of the type (α) and (β) the most informative event will be:

$$E = \exists x\, \exists y\, (\mathrm{ang}(x) \wedge \mathrm{ang}(y) \wedge x \neq y)\ .$$

Even if we add to K the following quantitive descriptors:

(21) $d \supset \exists x\, \exists y\, (\mathrm{ang}(x) \wedge \mathrm{ang}(y) \wedge x \neq y)$

(22) $e \supset \exists x\, \exists y\, (\mathrm{ang}(x) \wedge \mathrm{ang}(y) \wedge x \neq y)$

still the set of n-m-b causes will consist of both literals: d and e. In order to resolve such ambiguity we have to resort to outputs of the type (γ) combined with (β). In the case of the letter D they will generate the event:

$$E = \forall x\, \mathrm{ang}(x) \wedge \exists x\, \exists y\, (\mathrm{ang}(x) \wedge \mathrm{ang}(y) \wedge x \neq y).$$

Clearly in this case the only n-m-b cause is the formula consisting of the literal d . This concludes the discussion of our example.

Bibliography

[1] Buchnan, B., G. Sutherland, and E.A. Feigenbaum;
Heuristic Dendral, Machine Intelligence 4, 1969, pp. 209-254.

[2] Chang, C.L., and R.C.T. Lee;
Symbolic Logic and Mechanical Theorem Proving,
Academic Press, 1973.

[3] Cohen, B.L.,
A powerful and efficient structural pattern recognition system;
Artificial Intelligence, Vol. 9, No. 3., Dec. 1977, pp. 223-255.

[4] Lenat, D.B.,
Automated theory formation in mathematics;
Fifth IJCAI, Cambridge, 1977, pp. 833-842.

[5] Mendelson, E.,
Introduction to Mathematical Logic, D. Van Nostrand, 1964.

[6] Nilsson, N.J.,
Problem solving methods in artificial intelligence, McGraw Hill, 1971.

[7] Report of First AIM (AI in Medicine) Workshop, Rutgers University, 1975.

[8] Robinson, J.A.,
A machine oriented logic based on the resolution principle,
J ACM 12, 1(Jan. 1965), pp. 23-41.

[9] Shortliffe, E.H.,
Computer-Based Medical Consultations: MYCIN,
Elsevier Computer Science Library, 1976.

[10] Verne, S.A.,
Induction of concepts in the predicate calculus,
IJCAI-75, pp. 281-287.