Mechanical Generalization and Semantic Distance

by

T. Pietrzykowski

Research Report CS-75-25

Department of Computer Science

University of Waterloo
Waterloo, Ontario, Canada

October 1975

ABSTRACT

Let S (specific statements) and K (knowledge base) be conjunctions of clauses. The conjunction of clauses G is called a generalization of S over K if $(K \wedge G) \supset S$ and G satisfies a number of constraints which prevent pathologies and trivialities. We shall denote it $G \underset{K}{\geq} S$. The relation "$\underset{K}{\geq}$" is used to define atomic generalization "$\underset{K}{\triangleright}$" as follows: $G \underset{K}{\triangleright} S$ iff $G \underset{K}{>} S$ and for no L $G \underset{K}{>} L \underset{K}{>} S$. The degree of generalization $\delta(G,S)$ is defined as the length of the shortest chain of atomic generalizations necessary to reach S from G. This in turn is used to define semantic distance $D(S_1,S_2) = \min_G \{\phi(\delta(G,S_1)) + \phi(\delta(G,S_2))\}$ where ϕ is a weighting function. There is an algorithm provided to produce generalizations, which is based on generating singelton resolvents of the set $K \wedge \neg S$ (in the case where S is a one clause set), and replacing some of the terms occurring in them by variables.

0 INTRODUCTION

The purpose of this paper is to investigate mechanical aspects of creating generalizations. Also we present an attempt to use our definition of generalization to formulate the concept of semantic distance. In his paper [1] Plotkin defined the notion of generalization in the following way: let us have a set of clauses S. We shall call a clause G a generalization of S if all the clauses of S collapse into G by replacing some of the terms occurring in them by variables. This approach provides a useful tool, but it cannot be applied to more complex situations when our generalization depends not only on S but also on some accompanying knowledge. Consequently, in such situations the relations between generalization, knowledge and specific statements become more complex then in Plotkins case.

In §1 we provide the basic intuitive justification for the definition which is presented in §2. We show there also that generalizations can be "partially ordered," which in turn is used to define a concept of semantic distance. The last paragraph provides a rather rough description of how to mechanize the generation of generalizations.

1. Intuitive Definition of generalization

Before we proceed with the formal definition, we shall once more discuss the meaning of the term: generalization. This approach seems to be justified by rich, negative experience of people who have introduced formalizations which eventually had little to do with the original intuitions giving rise to the formalized ideas. Let us start with some dictionary definitions. Webster's 3rd New International Dictionary (1971) says: 'generalize' = 'to derive or induce (a general conception or principle) from particulars' or 'to make general (as by existential or universal quantification): render applicable to wider class'. The idea of 'general' seems to be better presented in the Oxford English Dictionary (1971): 'of a rule, law, principle, formula, description: Applicable to a variety of cases, true or purporting to be true for all or most of the cases which come under its terms' . From these definitions it explicitly follows that: we have a set of specific (particular) statements S, and from these statements we infer a general one of which the specific statements are cases. Here we come to a very interesting question: how do we define 'general' and how do we reason towards it? First let us note the implicit existence of some knowledge which will make it possible to verify that our general statement really yields the specific ones. This we shall call the knowledge base and denote K . Now we are ready to define the generalization G as a statement from which we can deduce each statement of S, with the help of the knowledge base. The above relation can be written as follows

$$(K \wedge G) \supset S.$$

This formalization captures the essence of our intuitions, however it is too liberal and leaves room for pathologies and trivialities (see the next paragraph). Therefore we have to approach this definition in a more technical fashion.

2. Formal definition of generalization

In the following we shall use a clausal formalism. The reason for it is twofold: the widespread familiarity with such notation and the simplicity of providing conditions for removing trivial and pathological cases from our definition. A wff is a conjunction

$$C_1 \wedge \ldots \wedge C_n \qquad 1 \leq n$$

of clauses C_i $1 \leq i \leq n$, where a clause is a disjunction

$$L_1 \vee \ldots \vee L_m \qquad 0 \leq m$$

of literals L_i $(1 \leq i < m)$ which are atomic formulae $P(t_1,\ldots,t_k), 1 \leq k$ or their negations, where P is a k-ary predicate symbol and $t_1,\ldots,t_k$ are terms. A term is a variable, a constant, or an expression $f(s_1,\ldots,s_\ell)$ where f is a ℓ-ary function symbol and $s_1,\ldots,s_\ell$ are terms. In the future we shall often refer to wffs as sets of clauses, and to clauses as sets of literals. We shall say that a clause C' is a variant of a clause C iff C' is obtained from C by renaming some of the variables occurring in C. Let K be a wff which represents the knowledge base. We shall assume that K is satisfiable (it may, however, be a part of a larger, contradictory knowledge), Def. 2.1. Let us assume that S is a wff consisting of a single clause. We shall say that a clause G is a generalization of S over K and denote this $G \underset{K}{\geq} S$ iff

$G \underset{K}{\geq} S$ iff

(1) $(K \wedge G) \supset S$,

(2) $K \wedge G$ is satisfiable,

(3) Let G' and S' denote clauses which result when a literal is removed from each of G and S, respectively. Let $\tilde{G}$ denote a clause resulting from

G by identifying a pair of occurrences of distinct variables.

(3.1) for any $\tilde{G}$ $(K \wedge \tilde{G}) \supset S$ does not hold,

(3.2) for any S' $(K \wedge G) \supset S'$ does not hold,

(3.3) for any G' there exists S' such that $(K \wedge G') \supset S'$ holds.

We shall say that G is <u>trivial</u> if it has a variant G' such that G'=S. When G is not a trivial generalization of S then we shall denote it $G \underset{K}{>} S$. We shall provide a motivation for introducing conditions (1) - (3) after extending our definition to unrestricted S.

Def.2.2. Let $S = S_1 \wedge \ldots \wedge S_m$, $1 \le m$. We shall say that $G = G_1 \wedge \ldots \wedge G_n$ is a <u>generalization</u> of S over K and denote it $G \underset{K}{\geq} S$ iff

(4) for each i $(1 \le i \le m)$ exists j such that $G_j \geq S_i$,

(5) G is minimal in the sense that if we remove any G_j $(1 \le j \le n)$ then the condition (4) does not hold.

We shall say that G is not trivial and denote this $G \underset{K}{>} S$ iff there exists a pair i,j such that $G_j \underset{K}{>} S_i$.

In the future, if there is no danger of confusion, we shall denote $G \underset{K}{>} S$ simply by $G > S$. We shall call a generalization G <u>unconditional</u> iff $G \supset S$. In this terminology, Plotkins [1] definition of generalization is a special case of an unconditional one.

<u>Corollary 2.1</u>. Referring to the Def. 2.2 we have the following:

$n \le m$

and

$(K \wedge G) \supset S.$

Now we shall explain reasons for introducing conditions (1) to (5).

First of all, the reason for splitting the definition into two parts, is simply for ease of comprehension.

The condition (1) represents the heart of the whole definition. It becomes a theorem in the multiclausal case (see Cor. 2.1).

The condition (2) guards against pathological cases where G is totally unrelated to S but satisfies (1).

The conditions (3) are, perhaps more questionable, but I hope that the following examples will justify their introduction. First I want to justify (3.1). Let K = "x is less than 1/2 a year old implies x is less than 1 meter high", S = 'x is less than 1/2 a year old implies x does not walk'. A generalization G = 'x is less than 1 meter high implies x does not walk' may be rather untrue (midgets can walk) but reasonable. However a statement G' = 'x is less than 1 meter high implies y doesn't walk' is utterly nonsensical: what does the fact that some x is less than 1 meter high have to do with the fact that some unrelated y does not walk! Condition (3.2) prevents situations such as that which arises when for the same knowledge base as in the above example, we could have a generalization as follows: 'x does not walk' which is hardly acceptable.

Condition (3.3) guards against redundancy of G. For example with K = $\{p \supset q,\ q \supset r\}$ and S = $\{s \supset r\}$ we have interesting generalizations like $s \supset q$ or $s \supset p$. What we want to avoid is $s \supset (q \vee p)$ which would be admissible without (3.3).

The condition (4) really provides the basic idea of our definition. It is important to notice that according to it, each clause of S is generalized individually which is stronger than just saying that $(K \wedge G) \supset S$. The condition (5) guards against irrelevant clauses which otherwise could always be added to G.

Now let us illustrate our definition on the following examples

Example 2.1 Let

$K = \{Saba(x) \supset Cat(x),\ Salomon(x) \supset Cat(x),\ Eat\ mice(x) \supset$
$Eat\ rodents(x),\ Eat\ mice(x) \supset Eat\ small\ animals(x)\}$

$S = \{Saba(x) \supset Eat\ rodents(x),\ Salomon(x) \supset Eat\ rodents(x)\}$

The following will be non-trivial generalizations:

$G_1 = \{cat(x) \supset eat\ rodents(x)\}$

$G_2 = \{Saba(x) \supset eat\ mice(x),\ Salomon(x) \supset eat\ rodents(x)\}$

$G_3 = \{cat(x) \supset eat\ mice(x)\}$

G_1 seems to be the most interesting one since there was no real need to introduce the predicate Eatmice(x) into the picture.

Example 2.2.

Let us choose K to be the same as in Example 2.1.

If $S = \{Saba(x) \supset Eat\ small\ animals(x),$
$Saba(x) \supset Eat\ rodents(x)\}$

then $G = \{Saba(x) \supset Eat\ mice(x)\}$

is a non-trivial generalization.

Example 2.1 seems to agree with our intuition, since in G_1 we inferred from the fact that two particular cats eat rodents, that cats in general eat rodents. But Example 2.2 may appear different: to eat rodents or small animals seems to be more general than to eat mice. However, the following view can help: the more general a statement, the more likely it is to be false in specific cases. In this case, Saba may not eat mice as inferred, but chipmunks, which are both small animals and rodents.

Example 1.3   Let us again choose K as in Example 2.1.

If   S = {Saba(x) ⊃ Eat small animals(x), Salomon(x) ⊃ Eat rodents(x)}

then the only single clause generalization would be

G = {Cat(x) ⊃ Eat mice(x)}.

"This is also a generalization for both previous examples, although the single clause generalizations from the previous examples are not generalizations in this case".

Example 2.4. Consider the following knowledge base

K = {Canary(x) ⊃ Bird(x),
Sparrow(x) ⊃ Bird(x),
Ostrich(x) ⊃ Bird(x)}

Now if

S = {Canary(x) ⊃ Fly(x),
Sparrow(x) ⊃ Fly(x)}

then the clause G = {Bird(x) ⊃ Fly(x)} will be a proper generalization. However if we add to K the following clauses: {Ostrich(x) ⊃ ~ Fly(x), ∃x Ostrich(x)} obtaining K', G will not be a generalization of S over K' because K' ∧ G unsatisfiable. The last example illustrates the situation which happens often in so called real life, when we make generalizations which work locally but as a rule contradict some other facts. We often treat these facts as exceptions to the rule. (See the Oxford definition in the Introduction). Now we shall investigate some properties of generalization.

Theorem 2.1 Let K, P_1,P_2 and P_3 be wffs. If $P_1 \geq P_2$ and $P_2 \geq P_3$ then $P_1 \geq P_3$. Unfortunately, this is not true if '≥' is replaced by '>', as illustrated in the example below.

<u>Example 2.5</u> Let K = {P(x) ⊃ P(f(x)), P(f(x)) ⊃ P(x).} Then the following holds: P(a) $\underset{K}{>}$ P(f(a) and P(f(a)) $\underset{K}{>}$ P(a).

The property described in the example above may be a source of further troubles, therefore we shall introduce a special name for it in order to be able to ban it in some situations.

<u>Def. 2.3</u>. A wff K is called <u>circular</u> iff for some wffs P_1 and P_2, we have $P_1 \underset{K}{>} P_2$ and $P_2 \underset{K}{>} P_1$.

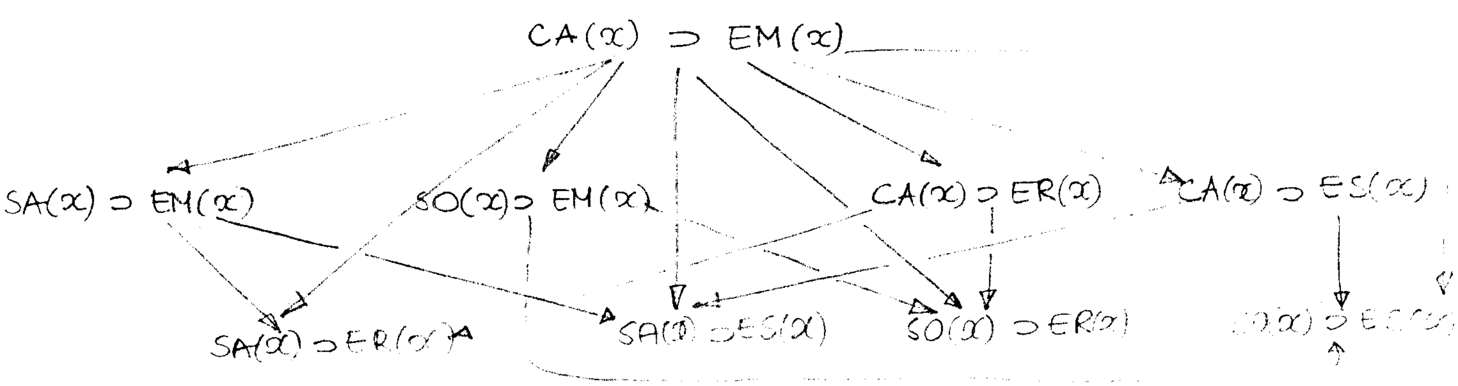
Now we can modify Theorem 2.1 as follows

<u>Theorem 2.2</u>. Let K be noncircular. Then for any wffs P_1, P_2, P_3 we have $P_1 > P_2$ and $P_2 > P_3$ implies $P_1 > P_3$.

Finally we shall provide an example which will illustrate the relation '>' on previously discussed material.

<u>Example 2.6</u>. Let K be the knowledge base described in Example 2.1. Then denoting the predicates: Cat, Saba, Salomon, eat rodents, eat small animals, eat mice, respectively by

CA, SA, SO, ER, ES, EM we have



where arrows stand for the sign '>'

3. Semantic distance

Now we enter a terrain where our intuitions are much less precise than those discussed in the previous paragraph. The notion of semantics is vague enough, and, combined with the notion of distance, it becomes arbitrary. Therefore we will try to rely only on the most simple intuitions and try to justify every step in our definition.

Let us examine the intuition behind the statement: 'these two objects closely resemble each other'. It usually means that we can discover a considerable number of features common to both. This obviously depends on our concept of equality of features, but we take it for granted. The more such features the objects have in common the more closely they are related. Now let us come back to our definition of "generalization": we extract common features of particular cases in order to formulate a rule which will encompass them. So let us consider the following situation if our objects are 2 texts then in order to capture their common features we shall attempt to provide a generalization of both. Now, as we know, it is usually possible to create many such generalizations, so in order to discover as many common features as possible, we shall try to obtain a generalization as close as possible to both texts, and measure the closeness by the minimal number of elementary generalizations necessary to reach a common generalization from both texts.

Let us now try to formalize the above intuitions.

Definition 3.1. We shall say that G is an atomic generalization of S and denote it $G \triangleright S$ iff $G > S$ and there is no R such that $G > R$ and $R > S$.

Example 3.1. Let us take K as in Example 2.1. Then

$\{cat(x) \supset eat\ mice(x)\} > \{Saba(x) \supset eat\ mice(x)\} > \{Saba(x) \supset eat\ rodents(x)\}$.

The notion of atomic generalization will be used as a quantitative tool to describe the notion of the degree of generalization.

Definition 3.2. Let K, S and G be wffs. We shall say that a sequence $R_0, R_1, \ldots, R_n$ $(1 \leq n)$ is a path from G to S iff

$$R_0 = G, \qquad R_n = S, \text{ and}$$

(1) $$R_i \triangleright R_{i+1} \qquad (0 \leq i \leq n-1).$$

n is called the length of the path. The length of the shortest path from G to S is called the degree of generalization of G to S over K and denoted

$$\delta_K(G,S).$$

The subscript K will often be omitted if there is no danger of ambiguities. It is interesting to notice that in spite of the fact that $G > S$ there might not be a path from G to S. For example consider K to be a knowledge base from Example 2.5. Then there is no path from P(a) to P(f(a) since there is no atomic generalizations there. (To see this more clearly let us notice that each generalization of P(a) is $P(f^n(a))$ where $n \geq 0$ and $f^n(\ )$ denotes $\overbrace{f(f(\ldots f(\ )\ldots))}^{n \text{ times}}$. Moreover for each $m,n \geq 0$ we have $P(f^m(a)) > P(f^n(a))$).

However we can ban this pathology and present the following theorem which specifies the area of practical application for Def. 3.2.

Theorem 3.1. If K is a noncircular knowledge base then for each G,S such that $G > S$ there exists a path from G to S and $\delta(G,S)$ is determined.

In the future all our discussions will deal with noncircular knowledge bases. It seems that this restriction is not too severe: circularity is quite a pathological property from the point of view of natural language and common sense knowledge.

To illustrate the idea of the path, let us examine the following example.

Example 3.2. Let $K = \{p \supset q, p \supset r, q \supset s, s \supset t, r \supset t\}$. If $G = \{p\}$ and $S = \{t\}$ then we have 2 possible paths from G to S: $p \triangleright q \triangleright s \triangleright t$ and $p \triangleright r \triangleright t$. Obviously $\delta(G,S) = 2$.

Let us now attack the key issue of this paragraph, namely semantic distance.

Definition 3.3. Let K, R and T be wffs. The semantic distance from R to T over K is denoted $D_K(R,T)$ and defined as:

(2) if there exists G such that $G > R$ and $G > T$ then

$$D(R,T) \overset{df}{=} \min_{\substack{G \\ G > R,T}} \{\phi(\delta(G,R)) + \phi(\delta(G,T))\}$$

or otherwise

(3) $$D(R,T) \overset{df}{=} +\infty .$$

The function ϕ is a weight function. It must be positive and strictly increasing. Now let us present an example which hopefully will convince the reader of the usefulness of the above definition. It will be applied to resolve ambiguities.

Example 3.3. Let us imagine a mechanical analysis of a picture. The purpose of this analysis is to identify various objects contained in this picture. The picture represents an office room. There are desks, a table, a telephone, waste baskets, books, etc. We shall assume the existence of a knowledge base describing relations among such objects. Specifically it contains information that desks and tables are furniture, that furniture and floor is made out of wood (a very unusual office!) We have a description of a telephone as a medium size, black object and the waste basket has the same attributes. We also have information that waste baskets usually stand on the floor and telephones on desks. Now let us imagine that in the process of analysis there has been discovered a medium size, black object residing on the table.

So our system is presented with a dilemma: is it a waste basket or a telephone. In order to resolve it we will try both possibilities and find out how they fit with the rest of our knowledge. In order to do so, we shall evaluate the semantic distance of each of the alternatives from the rest of the knowledge. The one which has a smaller distance will be the one which better fits the rest, and should be chosen as a more likely interpretation. Let us now present our problem in a formal way.

K = {

(1) on table(x) $\supset$ on furniture(x),

(2) on desk(x) $\supset$ on furniture(x),

(3) on furniture(x) $\supset$ on wood(x),

(4) on floor(x) $\supset$ on wood(x),

(5) on desk(x) $\supset$ telephone(x),

(6) on floor(x) $\supset$ waste basket(x)}.

One can object to (5) and (6) in that they are not too realistic but let us remember that our knowledge in this example is very local, and may be contradictory with some wider knowledge. Now our hypotheses concerning a black, mid size object on the table are as follows:

(α) on table(x) $\supset$ telephone(x)

(β) on table(x) $\supset$ waste basket(x)

In order to evaluate (α) let us choose from K something which seems to be closest (we shall formalize this later) namely (5), and for (β) it will be (6).

Now we shall find the distances D(α),(5)) and D((β),(6)). The closest common generalization for the first pair will be : on furniture(x) $\supset$ telephone(x) and for the second pair: on wood(x) $\supset$ waste basket(x). The structure of these is presented below:

$$\text{on furniture}(x) \supset \text{telephone}(x)$$

$$\swarrow \qquad\qquad \searrow$$

$$\text{on desk}(x) \supset \text{telephone}(x) \qquad\qquad \text{on table}(x) \supset \text{telephone}(x)$$

and

$$\text{on wood}(x) \supset \text{waste basket}(x)$$

$$\swarrow \qquad\qquad \searrow$$

$$\text{on floor}(x) \supset \text{waste basket}(x) \qquad\qquad \text{on furniture}(x) \supset \text{waste basket}(x)$$

$$\downarrow$$

$$\text{on table}(x) \supset \text{waste basket}(x)$$

Obviously the distance in the first case is $\phi(1) + \phi(1)$ while in the second $\phi(1) + \phi(2)$. Since ϕ is strictly growing, the first alternative is more promising and finally we decide that it is a telephone on the table, not a waste basket.

It was implicit in the above example that we have chosen from the knowledge base some fragments and measured their distances from the alternative statements representing the ambiguity. This really means that we try to find the distance from the whole knowledge to our alternatives. Therefore we shall introduce this concept in a formal way.

Definition 3.4 Let K and A be a wffs. We shall define a compatibility of K with A as follows

$$\text{comp}(K,A) \overset{\text{def}}{=} \min_{X \subseteq K} D_K(X,A).$$

4. Computing generalizations.

So far the concept of generalization and consequently of semantic distance, has been defined in purely existential terms. Now we shall deal with the practical aspects of computing. The procedure suggested below is sketched without details but seems to describe adequately, the basic ideas of the mechanical production of generalizations.

ALGORITHM. Let K be a wff and S a clause where $S = s_1 \vee \ldots \vee s_m$ $(m \geq 1)$.

(1) Let V be a set of variables occurring in S and let A be a set of constants such that $a \in A$ implies a does not occur in K or S. Suppose there is a 1-1 correspondence between V and A. Let $S' = s'_1 \vee \ldots \vee s'_m$ be a result of replacing in S occurrences of variables from V by corresponding constants from A.

(2) Let

$$P_{i_1 \ldots i_k} = \{s'_{i_1}, \ldots, s'_{i_k}\} \cup \{\bar{P} | P \text{ is a singleton resolvent of } K \wedge \bar{s}'_{i_1} \wedge \ldots \bar{s}'_{i_k} \ (k \geq 1) \text{ such that each } \bar{s}'_{i_j} \text{ is an ancestor of } P\}$$

where $\bar{P}$ is the complement of P.

(The sets $P_{i_1 \ldots i_k}$ are generally infinite)

(3) Let

$$\mathbf{H} \stackrel{df}{=} \{ C | C = c_1 \vee \ldots \vee c_p \text{ where } c_j \in P_{I_j} \text{ and } I_1, \ldots, I_p \text{ is a partition of } \{1, \ldots, m\}.$$

Let $\mathbf{T}(C)$ be a set of terms defined for any clause C as follows:

$$\mathbf{T}(c) \stackrel{df}{=} \{t_1, \ldots, t_k\} \mid t_i \text{ are distinct nonvariable terms occurring in C and for all } a \in A \text{ if a occurs in c then it occurs in some } t_i\}.$$

Let

$\mathbf{G}(K,S) \overset{df}{=} \{C|$ for some $C' \in \mathbf{H}$ and $\{t_1,\ldots,t_k\} \in \mathbf{T}(C')$

C results from C' by replacing each occurrence in C' of any $t \in \{t_1,\ldots,t_k\}$ by a variable, providing that these variables are distinct for different t's and different from variables occurring in C'$\}$.

The properties of the above algorithm are characterized by the following theorem

Theorem 2. Let K and S be wffs and $S = s_1 \wedge \ldots \wedge s_n$ where S_i is a clause. If K is satisfiable then a wff $G = G_1 \wedge \ldots \wedge G_m$ is a generalization of S over K iff for each S_i $(1 \leq i \leq m)$ exists j such that $G_j' \in \mathbf{G}(K,S_i)$, and G_j' is an alphabetic variant of G_j.

Since the set P and consequently $\mathbf{H}$ and $\mathbf{G}$ are generally infinite the algorithm above has to be equipped with some effective enumeration procedure, which will ensure the completeness of the process of production of generalizations. The termination of such a process is, however, undecidable.

We do not attempt here to provide such an enumeration procedure because the algorithm itself in its present form is rather unsuited to any practical use. It is provided only as a basis for possible future refinements by introducing search strategies.

Example 4.1. Let $K = \{p \supset t,\ r \supset q\}$ and $S = p \supset q$. Since there are no variables in S then $S = S' = -p \vee q$.

Denoting -p as s_1 and q as s_2 we have:

$P_1 = \{-p,-t\}$

$P_2 = \{q,r\}$

$P_{12} = \phi$

Therefore

$$\mathbf{H} = \{p \supset q,\ p \supset r,\ t \supset q,\ t \supset r\}$$

where $p \supset q$ is trivial.

Example 4.2. Let us now introduce variables and define $K = \{p(x) \supset t(x),$ $r(x) \supset q(x),\ -p(x) \vee q(y) \vee s(x)$ $S = -p(x) \vee q(y)$. Then $S' = -p(\alpha) \vee q(\beta)$. Denoting $-p(\alpha)$, $q(\beta)$ respectively as s_1, s_2 we have

$$P_1 = \{-p(\alpha),\ -t(\alpha)\}$$

$$P_2 = \{q(\beta),\ r(\beta)\}$$

$$P_{12} = \{-s(\alpha)\}$$

$$\mathbf{H} = -p(x) \vee q(y),\ -t(x) \vee q(y),\ -p(x) \vee r(y),\ -t(x) \vee r(y),\ - s(x)\}$$

We should note here the necessity of using distinct variables x and y in some of the elements of the set $\mathbf{H}$. Moreover the set P_{12} is not empty.

Example 4.3. Let $C = P(f(x),a)\ Q(f(a),b)$ where $a \in A$. The set $\mathbf{T}(C)$ is as follows:

$$\mathbf{T}(C) = \{a\},\ \{a,f(a)\},\ \{a,b\},\ \{a,f(a),b\},\ \{a,f(x)\}\ \{a,f(a),f(x)\}$$
$$\{a,b,f(x)\},\ \{a,f(a),b,f(x)\}$$

Now if $\mathbf{H} = \{C\}$ then

$$= \{P(f(x),u) \vee Q(f(u),b),\ P(f(x),u) \vee Q(v,b),$$

$$P(f(x),u) \vee Q(f(u),z),\ P(f(x),u) \vee Q(v,z),$$

$$P(w,u) \vee Q(f(u),b),\ P(w,u) \vee Q(v,b),$$

$$P(w,u) \vee Q(f(u),z),\ P(w,u) \vee Q(v,z)\}$$

5. Concluding remarks

It is clear that this report is a draft. I decided to present it in such form because in personal discussions, a number of people have expressed their interest in the idea of generalization as described above. However, I fully realize that both; the definition of generalizations, and the algorithm to generate them, may undergo considerable changes. But before such changes can be made, it is necessary to gain some practical experience in using the presented concepts. I hope that this report gives sufficient basis for applying these concepts, and that, as a result, I will get some feed back.

References

1. Plotkin, G.D. A note on inductive generalization, In Machine Intelligence 5, B. Meltzer and D. Michie (Eds.), American Elsevier, New York, 1970 pp. 153-164.

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