17 Aug 78

ATP 4;

Conflicting Bindings and Generalized Substitutions

Mabry Tyson
W. W. Bledsoe

Problems arise combining conjunctive subgoals whose solutions require conflicting bindings. Using a generalization of substitution, a method is given that allows the combination of the solutions.

One of the most productive methods of problem solving is problem reduction. If a problem can be split into two independent parts each of which may be solved separately, finding solutions to the smaller problems is a much simpler task (1). While problem reduction is very basic to human problem solving, it is perhaps more important to problem solving by machines (2,3). The recursion of algorithms that solve problems by reducing them to subproblems (to which the algorithm is reapplied) contributes to the clarity and conciseness of the algorithm. Also, present computer programs are not as adept at separating out the chaff as humans are and are therefore more susceptible to combinatorial explosions.

Solutions to independent problems are orthogonal and can be combined without interference. This is not so if the problems are not completely independent. If two parts of a problem are somewhat interdependent, it is necessary to confirm that their two solutions can be combined to return a single solution for the whole problem. Actually it is not the two parts that must be independent, it is their solutions. Since any problem may have a number of different solutions, for two problems, one pair of solutions may be independent while another pair may be mutually exclusive. (Example - If you need a number that is both odd and prime you would lose if you picked the first odd number for one subgoal and the first prime number for the other subgoal.)

The divide and conquer methodology is central to the method of theorem proving often referred to as "natural deduction" (2,4,5,6). One of its principal proof techniques is splitting a single goal into multiple goals and later combining the results. Thus to prove

$$H \Rightarrow A \,\&\, B$$

the two subgoals

$$H \Rightarrow A \quad \text{and} \quad H \Rightarrow B$$

are proved. Unfortunately it is not quite as simple as this due to the presence of variables that may occur in H, A, B, or even in higher subgoals. This paper will examine the problems of independence of subgoals and their solutions within the context of a particular methodology of theorem proving. We will also show how the solutions must be combined to provide a solution of the higher level goal. Most of the time two subgoals may be combined rather simply. We present a theorem that defines the necessary conditions for this. We also present a theorem that covers the situations where the two solutions interfere with each other. In order to do this we will generalize the concept of substitutions.

The UT interactive theorem prover is a natural deduction system developed by Bledsoe's group at the University of Texas which has been used over a number of years to prove theorems in such areas as set theory, topology, program verification, and limit theorems of calculus and analysis. The following discussion will be presented in terms of this implementation although the ideas and techniques extend much further. For a complete discussion of this prover see (7).

When a closed formula, E, is given to the prover, it skolemizes the formula into an open formula, S. By the nature of the skolemization, if there exists some substitution, θ, such that $S\theta$ is ground and true then the original formula E is true. Likewise if there is some set of substitutions $\theta_1, \theta_2, \ldots, \theta_m$ such that

$$S\theta_1 \vee S\theta_2 \vee \ldots \vee S\theta_m$$

is both ground and true then E is true.

Although an open formula which is true for every interpretation of its free variables could be called true, we will call such a formula ground-true. The name derives from the fact that a proof of that formula would treat the variables as ground terms. An example would be the tautology

$$X \vee \neg X.$$

The similar closed formula

$$\forall X\ (X \vee \neg X)$$

is true while the previous open formula is only ground-true. A formula that is true will also be considered to be ground-true. The reason for this distinction is that in the prover a goal will typically contain free variables that may be bound in the proof. We will refer to such a goal as provable if there is some substitution that will make the goal ground-true.

Consider what happens when the prover is given the formula

$$(\forall X\ P(X)) \rightarrow P(A) \ \&\ P(B)).$$

The skolemized form is

$$P(X) \rightarrow P(A) \ \&\ P(B)$$

which is given to the routine IMPLY. IMPLY is the recursive routine that has the task of determining a substitution (which IMPLY returns as its value) which makes its input ground-true. For this example IMPLY will recur on the two subgoals

$$P(X) \rightarrow P(A) \quad \text{and} \quad P(X) \rightarrow P(B).$$

The two subgoals are proved with a substitution {A/X} for the first and {B/X} for the second.

But is this enough for IMPLY (as opposed to the full prover) to report it has proved its original input? The two subgoals were not independent as they both contained the variable X so their solutions cannot be easily combined. In fact, no ordinary substitution exists which makes the original input to IMPLY ground-true. As we shall see, in this particular case the conflict in the substitutions for X does not lead to problems that prevent the original input from being proved true but there are non-theorems whose downfall is due to a similar step in attempted proofs.

Remember that interdependent goals require some special manipulations in order for their separate solutions to be combined to provide a solution for the combined goal. The prover does check to see that solutions returned from subgoals are such that they may be combined.

In doing an AND-SPLIT such as above, the prover does not actually prove the two subgoals independently. Instead, it waits until the first subgoal succeeds and then uses the result of that in setting up the second subgoal. This significantly reduces the chance that the second subgoal would be proved inconsistently with the first. If IMPLY is given the goal

$$H \Rightarrow (A \,\&\, B)$$

it will first prove

$$H \Rightarrow A$$

using some substitution θ. Then it will form the second subgoal

$$H \Rightarrow (B\theta)$$

and attempt to prove it. Consider the proof of

$$\exists X\ (P(A) \,\&\, Q(B) \,\&\, Q(A) \Rightarrow (P(X) \,\&\, Q(X)))$$

and it should be clear why the θ needs to be applied to the conclusion of the second subgoal.

When the second subgoal returns a substitution of λ, the prover will return the composition of the two substitutions, $\theta\lambda$, as the substitution that proves

$$H \Rightarrow (A \,\&\, B).$$

To prove that our proof method is sound we need to prove that if IMPLY returns a substitution then that substitution will make the input formula ground-true. We will only prove this for the

AND-SPLIT rule but the other rules used by IMPLY are proved similarly. In order to prove it we need to show that if

$$(H \Rightarrow A)\theta$$

and

$$(H \Rightarrow B\theta)\lambda$$

are both ground-true (inductive hypothesis) then

$$(H \Rightarrow A \ \& \ B)\theta\lambda$$

is also ground-true.

Difficulties arise when the two substitutions contain a conflict such as the ones above ({A/X} and {B/X}). Two substitutions are in conflict if they substitute different terms for the same variable. Problems are also encountered if a substitution is such that some element of its domain occurs in its range, that is, if the composition of the substitution with itself differs from the original substitution. So we will put conditions on the solutions returned by the subgoals in order to make the theorem provable.

Definition. A substitution θ is called normal if the composition of θ with itself is θ again.

Definition. Two substitutions θ and λ are said to conflict if their domains are not disjoint.

Theorem: If θ, λ, and $\theta\lambda$ are all normal and θ and λ do not conflict, then if

$$(H \rightarrow A)\theta$$

is ground-true and

$$(H \rightarrow B\theta)\lambda$$

is ground-true, then

$$(H \rightarrow A \ \& \ B)\theta\lambda$$

is also ground-true.

Previously the prover would note conflicts in substitutions and would halt a proof if the conflicts might give rise to problems. The above example with the conflicting substitutions was handled by returning the self-conflicting substitution {A/X, B/X}. The prover would not allow this substitution to be applied to any formula containing an X. By using this method the prover handles most cases. However, the simplest example of a theorem the prover would halt on is

$$Q(A) \;\&\; Q(B) \Rightarrow \exists X(\; (P(X) \rightarrow P(A) \;\&\; P(B)) \;\&\; Q(X)).$$

The proof of the first half of this conclusion proceeds as in the previous example but the prover halts when trying to substitute {A/X, B/X} into Q(X). However, we have developed a theory that allows the prover to adequately handle these conflicts and allow proofs to proceed.

The central idea of the theory is the notion of generalized substitutions. Basically a generalized substitution contains both substitutions and information about the relationship of these substitutions.

Definition. θ is a generalized substitution if
(1) θ is an ordinary substitution, or
(2) θ has the one of the forms

$$(\theta 1 \vee \theta 2), \quad \text{or} \quad (\theta 1 \;\&\; \theta 2)$$

where $\theta 1$ and $\theta 2$ are generalized substitutions.

Definition. If θ is a generalized substitution, then we define θ' by
(1) $\theta' = \theta$ if θ is an ordinary substitution,
(2) $(\theta 1 \vee \theta 2)' = (\theta 1' \;\&\; \theta 2')$,
(3) $(\theta 1 \;\&\; \theta 2)' = (\theta 1' \vee \theta 2')$.

Definition. A generalized substitution is said to be a pure disjunction (conjunction) if it contains no & symbols (v symbols).

Ordinary substitutions are both pure disjunctions and pure conjunctions.

Definition. If A is a formula and θ is a generalized substitution, then $A\theta$ is the formula gotten by applying θ from left to right, ie,
(1) $A\theta$ is the usual result if θ is an ordinary substitution,
(2) $A(\theta 1 \vee \theta 2) = A\theta 1 \vee A\theta 2$,
(3) $A(\theta 1 \;\&\; \theta 2) = A\theta 1 \;\&\; A\theta 2$.

Properties: If θ and λ are generalized substitutions, λ is a pure disjunction, and A and B are formulas then λ' is a pure conjunction and
(1) $(\theta')' = \theta$
(2) $-(A\theta) = (-A)\theta'$
(3) $(A \vee B)\lambda = A\lambda \vee B\lambda$
(4) $(A \;\&\; B)\lambda = A\lambda' \;\&\; B\lambda'$
(5) $(A \rightarrow B)\lambda = (A\lambda' \rightarrow B\lambda)$

The following theorem justifies an AND-SPLIT that makes use of generalized substitutions.

Theorem: If θ and λ are pure disjunctive generalized substitutions then if θ is such that

$$(H \text{ -> } A)\theta$$

is ground true and λ is such that

$$(H \text{ -> } B\theta')\lambda$$

is ground true then

$$(H \text{ -> } A \text{ \& } B) \ (\theta\lambda \text{ v } \lambda)$$

is ground true.

If IMPLY is given the goal of

$$(H \text{ => } A \text{ \& } B)$$

it will first form the subgoal

$$(H \text{ => } A).$$

If this subgoal now returns the substitution θ, IMPLY will form the second subgoal

$$(H \text{ => } B\theta').$$

If this subgoal returns the substitution λ, IMPLY will return the substitution of

$$(\theta\lambda \text{ v } \lambda)$$

for the original goal.

Thus the two conflicting solutions of the subgoals generated by the earlier example

$$P(X) \text{ => } P(A) \text{ \& } P(B)$$

can be combined into the single generalized substitution

$$(\{A/X\} \text{ v } \{B/X\}).$$

When this method is applied to the example

$$Q(A) \text{ \& } Q(B) \text{ => } \exists X(\ (P(X) \text{ -> } P(A) \text{ \& } P(B)) \text{ \& } Q(X))$$

the second subgoal becomes

$$Q(A) \text{ \& } Q(B) \text{ => } Q(X)(\{A/X\} \text{ v } \{B/X\})'$$

which is just

$$Q(A) \ \& \ Q(B) \Rightarrow Q(X)\{A/X\} \ \& \ Q(X)\{B/X\}$$

and is proved.

At first glance it appears that the use of generalized substitutions increases the amount of work in the simpler cases one level up. If θ is $\{A/X\}$ and λ is $\{B/Y\}$ then the returned substitution is

$$(\{A/X, B/Y\} \ v \ \{B/Y\}).$$

If this goal were the first subgoal of the higher up goal

$$H \Rightarrow ((A \ \& \ B) \ \& \ P2)$$

then we would need to prove

$$H \Rightarrow P2(\{A/X, B/Y\} \ v \ \{B/Y\})'$$

which is

$$H \Rightarrow P2\{A/X, B/Y\} \ \& \ P2\{B/Y\}.$$

It is easy to see that this is equivalent to

$$H \Rightarrow P2\{A/X, B/Y\}$$

which is what would have to be proved using the procedure that does not allow conflicts. The prover can detect this rather simply.

References

1. Nils J. Nilsson. Problem-Solving Methods in Artificial Intelligence, McGraw-Hill, 1971.

2. A. Newell, J.C. Shaw and H.A. Simon. Empirical explorations of the logic theory machine: a case study in heuristics. RAND Corp. Memo P-951, Feb. 28, 1957. Proc. Western Joint Computer Conf. 1956, 218-239. Computers and Thought, Feigenbaum and Feldman (Eds.), 134-152.

3. James R. Slagle. A heuristic program that solves symbolic integration problems in freshman calculus, JACM, Vol 10, 1963, 507-520.

4. H. Gelernter. Realization of a geometry theorem-proving machine. Proc. Int'l Conf. Information Processing, 1959, Paris UNESCO House, 273-282.

5. Raymond Reiter. A semantically guided deductive system for automatic theorem proving. Proc. Third IJCAI, 1973, 41-46.

6. Richard Fikes and Gary Hendrix. A network-based knowledge representation and its natural deduction system, Proc. Fifth IJCAI, 1977, 235-246

7. W.W. Bledsoe and Mabry Tyson. The UT interactive theorem prover. The Univ. of Texas at Austin Math. Dept. Memo ATP-17a, May 1975.