FUNCTION SYMBOL ELIMINATION

W. W. Bledsoe

ATP 68 June 1982

FUNCTION SYMBOL ELIMINATION

W. W. Bledsoe

Abstract. Function symbol elimination (FSE) is a lot like variable elimination (VE). In VE, a variable symbol x, which is eligible in an inequality clause, can be removed, leaving an equivalent, more useful, clause. In this paper we present a similar rule for (partially) removing function symbols f (where f represents a continuous function on the reals), and show its use in some proofs.

1. Introduction

Variable Elimination (VE) has proved to be a useful tool in the automatic proofs of inequality theorems, especially when used in conjunction with shielding-term removal [1,2].

A variable symbol x is said to be eligible in a clause C, if it does not appear within the arguments of an uninterpreted function symbol [1]. For example, x is eligible in $(a \leq x \vee x \leq f(y))$ but not in $(a \leq x \vee x \leq f(x))$.

Variable Elimination Rule:

If x is eligible in a clause C and x occurs in C only in the literals

(1) $a_i \not\leq x;\ i = 1,n \quad ; \quad x \not\leq b_j;\ j = 1,n$

then C is replaced by its VE-Resolvent C' which is gotten by removing the literals (1) from C and replacing them by literals

(2) $a_i \not\leq b_j\ ,\ i = 1,n;\ j = 1,n.$

It should be noted that if either n or m is zero, then no literal is added to replace those deleted. The rule is extended appropriately to include the symbol '<'.

Examples

Ex. 1. $C = (a \not\leq x \vee x \not< b)\ ;\ C' = (a \not< b)$

Ex. 2. $C = (x < a \vee x \leq b \vee \ell)\ ;\ C' = \ell$

Ex. 3. $C = a \leq x \vee x \not\leq b \vee f(x) \leq c$

(x is not eligible in C so it cannot be eliminated).

It is the purpose of this paper to define a similar rule for (partially) eliminating function symbols f, where f represents a continuous function on the reals. We prove its soundness in Section 4.

First let us give some motivation. A clause C of the form $(a \not\leq x \vee x \not\leq b)$, where x does not occur in a or b, represents the (unnegated) theorem.

(3) $\exists x(a < x < b)$.

and (3) is true if and only if $(a < b)$. Thus C can be replaced by $C' = (a \not\leq b)$.

The existence of such an x, between a and b, in (2) is guaranteed by the denseness of the real numbers.

A similar situation holds for continuous functions on the reals.

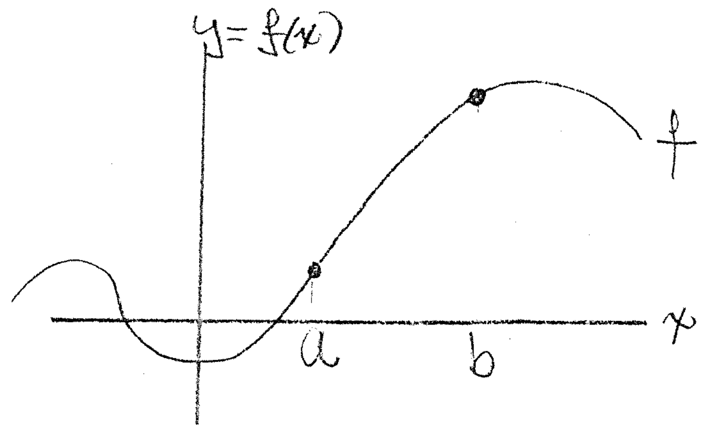


Figure 1

For example, in Figure 1, we know that there is a number c between a and b for which $f(a) < f(c) < f(b)$, (provided that f is continuous), because the values of f(x) are dense between f(a) and f(b). This property of continuous functions is usually expressed as the <u>intermediate value theorem</u>

2. Function Symbol Elimination

In all that follows we will assume that f is a continuous function of one real variable. This assumption could be indicated by simply carrying along the additional literal

$$\sim \text{continuous } f$$

but we will suppress that for brevity.

FSE-Rule. (Tentative)

If f represents a continuous function of one real variable, and C is a clause of the form

$$(4) \qquad f(x) \le u \vee w \le f(x) \vee \bigvee_i (x \le c_i) \vee \bigvee_j (d_j \le x) \vee E,$$

where x is a variable symbol which does not occur in c_i, d_j, or E, then we may add its FSE-resolvent C'

$$(5) \qquad \begin{array}{l} u < f(y) \vee f(z) < w \vee w \le u \vee E \\ \vee \bigvee_i (y < c_i \vee z < c_i) \vee \bigvee_j (d_i < y \vee d_j < z), \end{array}$$

where y and z are new variable symbols.

It should be noted that the FSE-rule does not actually eliminate the function symbol f. In fact there is no net reduction in f's! But it does introduce new variables y and z, which can be instantiated separately. This will be further discussed below.

Note that the second line of (5) contains only $<$'s (no $\leq$'s). (This will be modified in the general FSE-rule stated below). The rule will now be extended appropriately to include the symbol '$<$' .

In order to include an option for either $\leq$ or $<$ in the rule, we will use symbols such as $\underset{\cdot}{<}$, $\underset{\cdot\cdot}{<}$, $\underset{\cdot\cdot\cdot}{<}$, $\underset{i}{<}$, $\underset{j}{<}$. In any particular use, each of these symbols will represent $\leq$ or $<$. For example

$$(a \underset{\cdot}{<} b \longrightarrow a \underset{\cdot}{<} b)$$

is a theorem no matter whether $\underset{\cdot}{<}$ is $<$ or $\leq$ (but it must have the same value for both of its occurrences).

Furthermore, we will use the notation $\underset{\cdot}{<}'$ to mean $\leq$ if $\underset{\cdot}{<}$ is $<$, and $<$ otherwise. Similarly for $\underset{\cdot\cdot}{<}'$, $\underset{i}{<}'$, etc., and the notation $(\underset{\cdot}{<} , \underset{\cdot\cdot}{<})$ will be equal to $\leq$ if either of $\underset{\cdot}{<}$ or $\underset{\cdot\cdot}{<}$ is $\leq$, else it is $<$.

FSE-Rule.

If f represents a continuous function of one real variable, and C is a clause of the form

$$(4') \quad f(x) \underset{\cdot}{<} u \vee w \underset{\cdot\cdot}{<} f(x) \vee \bigvee_i (x \underset{i}{<} c_i) \vee \bigvee_j (d_j \underset{j}{<} x) \vee E$$

where x is a variable symbol which does not occur in c_i, d_j, or E, then we may add its FSE-resolvent C',

$$(5') \quad \begin{aligned} & u < f(y) \vee f(z) < w \vee w \underset{\cdot\cdot\cdot}{<} u \vee E \\ & \vee \bigvee_i (y \underset{i}{<} c_i \vee z \underset{i}{<} c_i) \vee \bigvee_j (d_j \underset{j}{<} y \vee d_j \underset{j}{<} z) , \end{aligned}$$

where y and z are new variable symbols, and $\underset{\cdots}{<}$ is $(\underset{\cdot}{<}, \underset{\cdot\cdot}{<})$. If $\underset{\cdot}{<}$ or $\underset{\cdot\cdot}{<}$ is $\leq$, then the $\underset{i}{<}$ and $\underset{j}{<}$ in (5') can all be made $<$ (which results in a stronger more useful resolvent).

Examples.

Ex. 4.

$$C = (f(x) \leq c \vee d \leq f(x) \vee a \leq x$$

$$C' = (c \leq f(y) \vee f(z) \leq d \vee d \leq c \vee a < y \vee a < z)$$

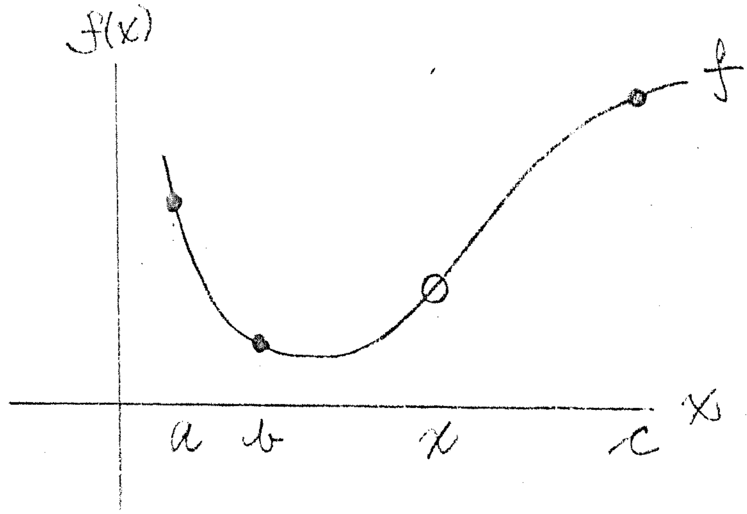
Ex. 5.

$$C = (f(x) < c \vee d < f(x) \vee a \leq x)$$

$$C' = (c < f(y) \vee f(z) < d \vee d < c \vee a \leq y \vee a \leq z)$$

3. Example Proofs using FSE

Theorem LS1. Continuous $f \wedge a < b < c \wedge f(b) < f(a) < f(c) \longrightarrow \exists x(b < x \wedge f(b) < f(x) < f(a))$.



Clauses.

0. continuous f
1. $a < b$ (not needed)
2. $b < c$
3. $f(b) < f(a)$
4. $f(a) < f(c)$
5. $x \leq b \vee f(x) \leq f(b) \vee f(a) \leq f(x)$
6. $f(b) < f(y) \vee f(z) < f(a) \vee f(a) \leq f(b)$
   $\vee\ y < b \quad z < b$ — 5, FSE, f(x)
7. $f(z) < f(a) \vee f(a) \leq f(b) \vee z < b$ — 6, b/y
8. $f(a) \leq f(b) \vee c < b$ — 7, 4, c/z
9. $c < b$ — 8, 3
10. $\square$ — 9,2.

<u>Theorem LS2</u>. Continuous f $\wedge$ a < b < c $\wedge$ f(b) < f(a) < f(c)

$\longrightarrow$ $\exists x \exists y \exists z\ (z < y < x \wedge f(y) < f(x) < f(z))$

<u>Clauses</u>.

1. $a < b$
2. $b < c$
3. $f(b) < f(a)$
4. $f(a) < f(c)$
5. $y \leq z \vee x \leq y \vee f(x) \leq f(y) \vee f(z) \leq f(x)$

---

6. $f(y) < f(y') \vee f(z') < f(z) \vee f(z) \leq f(y)$
   $y \leq z \vee y' < y \vee z' < y$ — 5, FSE, f(x)
7. $f(z') < f(z) \vee f(z) \leq f(y) \vee y \leq z \vee z' < y$ — 6, y/y'
8. $f(z') < f(a) \vee b \leq a \vee z' < b$ — 7, 3, a/z, b/y
9. $b \leq a \vee c < b$ — 8, 4, c/z'
10. □ — 9, 1, 2

<u>Theorem IMV1</u>. Continuous f and $a < b \wedge f(a) < f(b)$

$\longrightarrow \exists x(a < x < b \wedge f(a) < f(x) < f(b))$

<u>Clauses</u>.

1. $a < b$
2. $f(a) < f(b)$
3. $b \leq x \vee x \leq a \vee f(x) \leq f(a) \vee f(b) \leq f(x)$

---

4. $f(a) < f(y) \vee f(z) < f(b) \vee f(b) \leq f(a)$
   $\vee\ b < y \vee b < z \vee y < a \vee z < a$ — 3, FSE, f(x)
5. $f(z) < f(b) \vee f(b) \leq f(a) \vee b < a \vee b < z \vee z < a$ — 4, a/y
6. $f(b) \leq f(a) \vee b < a$ — 5, b/z
7. $\square$ — 6, 1, 2

Theorem IMV2. Continuous $f \wedge a \leq b \wedge f(a) \leq c \leq f(b)$

$\rightarrow \exists x(a \leq x \leq b \wedge f(x) = c)$.

Clauses.

1. $a \leq b$
2. $f(a) \leq c$
3. $c \leq f(b)$
4. $x < a \vee b < x \vee f(x) < c \vee c < f(x)$

(Note: $f(x) = c$ has been replaced by $c \leq f(x) \leq c$.)

5.	$c < f(y) \vee f(z) < c \vee y < a \vee z < a \vee b < y \vee b < z$	4, FSE, f(x)
6.	$f(z) < c \vee a < a \vee z < a \vee b < a \vee b < z$	5, 2, a/y
7.	$b < a$	6, 3, b/z
8.	$\square$	7,1.

4. Soundness of FSE

Df. $[\![x,y]\!] = [x,y] \cup [y,x]$

$(\!(x,y)\!) = (x,y) \cup (y,x)$

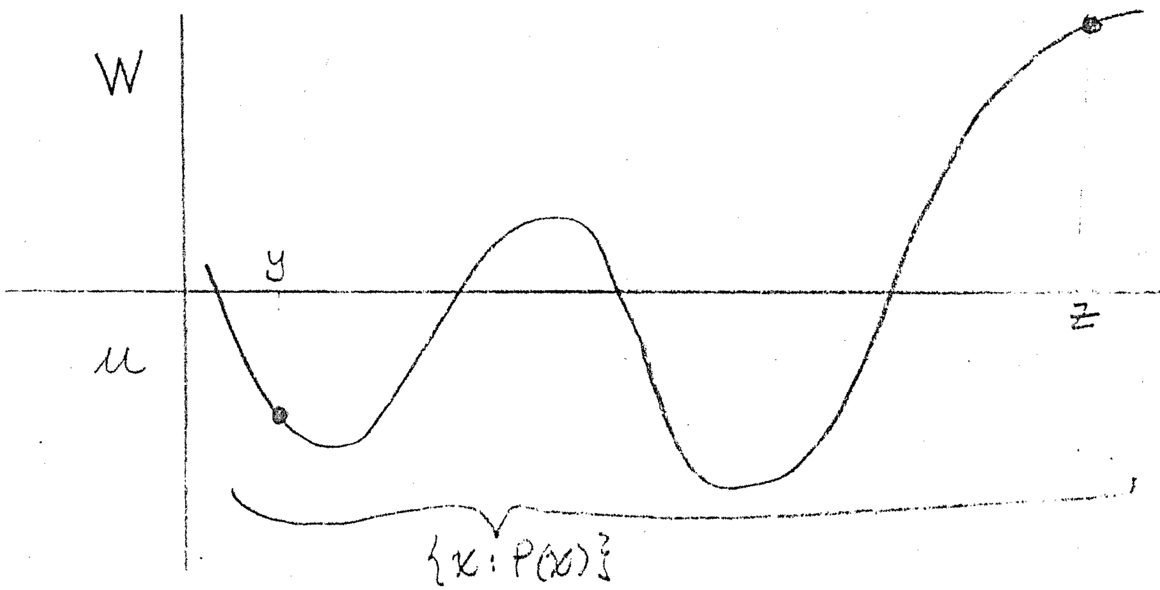
(Similarly for $[\![x,y)\!)$ and $(\!(x,y]\!]$.)

Thus $[\![2,1]\!] = [1,2]$, $(\!(1,3]\!] = (1,3]$, etc.

Theorem 1. If f is continuous then

$$\forall u \forall w [\exists y \exists z (f(y) \leq u < w \leq f(z) \wedge \forall x \varepsilon (\!(y,z)\!) P(x))$$

$$\longrightarrow \exists x (u \underset{\bullet}{<} f(x) \underset{\bullet\bullet}{<} w \wedge P(x))$$



Proof. By the intermediate value theorem there is an x in $(\!(y,z)\!)$ for which

$$u < f(x) < w.$$

But this implies $u \underset{\bullet}{<} f(x) \underset{\bullet\bullet}{<} w$, and since $P(x)$ holds for all x's in $(\!(y,z)\!)$, the proof is complete.

Theorem 2. If f is continuous then

$$\forall u \forall w [\exists y \exists z (f(y) \leq u \leq w \leq f(z) \wedge \forall x \in [\![y,z]\!]\ P(x))$$

$$\longrightarrow \exists x (u \leq f(x) \leq w \wedge P(x))$$

Proof. If $u < w$, use Theorem 1, otherwise use the intermediate value theorem to choose an x in $[\![y,z]\!]$ for which $f(x) = u = w$. Then by hypothesis, $P(x)$ holds.

Theorem 3. If $a \neq b$, then

$$\forall x \in (\![a,b]\!) [\bigwedge_i (c_i \underset{i}{<} x) \wedge \bigwedge_j (x \underset{j}{<} d_j)]$$

$$\equiv \bigwedge_i (c_i \leq a \wedge c_i \leq b) \wedge \bigwedge_j (a \leq d_j \wedge b \leq d_j).$$

Proof. Let $a \neq b$. We complete the proof in 7 steps. Steps 1 and 2 are obvious; step 3 follows from steps 1 and 2; step 5 follows from step 4; the proof of step 6 is similar to that of step 5; step 7 follows from steps 5 and 6.

<u>Step 1</u>. $a < b \rightarrow [\forall x \varepsilon (a,b)(c \leq x) \equiv (c \leq a)]$.

<u>Step 2</u>. $a < b \rightarrow [\forall x \varepsilon (a,b)(c < x) \equiv (c \leq a)]$.

<u>Step 3</u>. $a < b \rightarrow [\forall x \varepsilon (a,b)(c \underset{\cdot}{<} x) \equiv (c \leq a)]$.

<u>Step 4</u>. $\forall x \varepsilon \llparenthesis a,b \rrparenthesis (c \underset{\cdot}{<} x) \equiv (c \leq a \wedge c \leq b)$.

<u>Proof</u>. If $a < b$, then $\llparenthesis a,b \rrparenthesis = (a,b)$, and by Step 3

$$\forall x \varepsilon \llparenthesis a,b \rrparenthesis (c \underset{\cdot}{<} x) \equiv (c \leq a < b)$$

$$\rightarrow (c \leq a \wedge c \leq b),$$

and

$$(c \leq a \wedge c \leq b) \rightarrow (c \leq a)$$

$$\rightarrow \forall x \varepsilon (a,b)(c \underset{\cdot}{<} x).$$

Thus the result holds for $a < b$. Similarly it holds for $b < a$.

<u>Step 5</u>. $\forall x \varepsilon \llparenthesis a,b \rrparenthesis \bigwedge_i (c_i \underset{i}{\underset{\cdot}{<}} x) \equiv \bigwedge_i (c_i \leq a \wedge c_i \leq b)$

<u>Step 6</u>. $\forall x \varepsilon \llparenthesis a,b \rrparenthesis \bigwedge_j (x \underset{i}{\underset{\cdot}{<}} d_j) \equiv \bigwedge_j (a \leq d_j \wedge b \leq d_j)$

<u>Step 7</u>. $\forall x \varepsilon (a,b) [\bigwedge_i (c_i \underset{i}{\underset{\cdot}{<}} x) \wedge \bigwedge_j (x \underset{j}{\underset{\cdot}{<}} d_j)]$

$$\equiv \bigwedge_i (c_i \leq a \wedge c_i \leq b) \wedge \bigwedge_j (a \leq d_j \wedge b \leq d_j).$$

<u>QED</u>.

<u>Theorem 4</u>. $\forall x \varepsilon [\![a,b]\!] \; [\bigwedge_i (c_i \underset{i}{\leq} x) \wedge \bigwedge_j (x \underset{j}{\leq} d_j)]$

$$\equiv \bigwedge_i (c_i \underset{i}{\leq} a \wedge c_i \underset{i}{\leq} b) \wedge \bigwedge_j (a \underset{j}{\leq} d_j \wedge b \underset{j}{\leq} d_j)$$

<u>Proof</u>. The proof is modeled after that of Theorem 3 except that we have $[\![a,b]\!]$ instead of $(\!(a,b)\!)$, and in Steps 1-3, we have the hypothesis, $a \leq b$, instead of $a < b$, and the right sides of Steps 3-7 employ $\underset{i}{\leq}$ and $\underset{j}{\leq}$ instead of $\leq$. Only Steps 1-4 are given.

<u>Step 1</u>. $a \leq b \longrightarrow [\forall x \varepsilon [a,b] (c \leq x) \equiv (c \leq a)]$

<u>Step 2</u>. $a \leq b \longrightarrow [\forall x \varepsilon [a,b] (c < x) \equiv (c < a)]$

<u>Step 3</u>. $a \_ b \longrightarrow [\forall x \varepsilon [a,b] (c \underset{\bullet}{<} x) \equiv (c \underset{\bullet}{<} x)]$

<u>Step 4</u>. $\forall x \varepsilon [\![a,b]\!] . (c \underset{\bullet}{<} x) \equiv (c \underset{\bullet}{<} a \wedge c \underset{\bullet}{<} b)$

<u>Theorem 5</u>. If f is continuous then

$$\forall u \forall w [\exists y \exists z (f(y) \leq u < w \leq f(z)$$

$$\wedge \bigwedge_i (c_i \leq y \wedge c_i \leq z) \wedge \bigwedge_j (y \leq d_j \wedge z \leq d_j))$$

$$\longrightarrow \exists x (u \underset{\cdot}{<} f(x) \underset{\cdot\cdot}{<} w \wedge \bigwedge_i (c_i \underset{i}{<} x) \wedge \bigwedge_j (x \underset{j}{<} d_j))]$$

and

$$\forall u \forall w [\exists y \exists z (f(y) \leq u \leq w \leq f(z)$$

$$\wedge \bigwedge_i (c_i \underset{i}{<} y \wedge c_i \underset{i}{<} z) \wedge \bigwedge_j (y \underset{j}{<} d_j \wedge z \underset{j}{<} d_j))$$

$$\longrightarrow \exists x (u \leq f(x) \leq w \wedge \bigwedge_i (c_i \underset{i}{<} x) \wedge \bigwedge_j (x \underset{j}{<} d_j))]$$

<u>Proof</u>. Let $P(x) = \bigwedge_i (c_i \underset{i}{<} x) \wedge \bigwedge_j (x \underset{j}{<} d_j)$. Then the first part of Theorem 5 follows from Theorems 1 and 3, while the second part follows from Theorems 2 and 4.

<u>QED</u>.

<u>Theorem 6</u>. If f is continuous then

$$\exists u \exists w [\forall x (f(x) \underset{\cdot}{<} u \vee w \underset{\cdot\cdot}{<} f(x)$$

(6.1) $$\vee \bigvee_i (x \underset{i}{<} c_i) \vee \bigvee_j (d_j \underset{j}{<} x))$$

$$\longrightarrow \forall y \forall z (u < f(y) \vee f(z) < w \vee w \leq u$$

(6.2) $$\vee \bigvee_i (y < c_i \vee z < c_i) \vee \bigvee_j (d_j < y \vee d_j < z))]$$

and

$$\exists u \exists w [\forall x (f(x) < u \vee w < f(x)$$

(6.3) $$\vee \bigvee_i (x \underset{i}{<} c_i) \vee \bigvee_j (d_j \underset{j}{<} x))$$

$$\longrightarrow \forall y \forall z (u < f(y) \vee f(z) < w \vee w < u$$

(6.4) $$\vee \bigvee_i (y \underset{i}{<} c_i \vee z \underset{i}{<} c_i) \vee \bigvee_j (d_j \underset{j}{<} y \vee d_j \underset{j}{<} z))]$$

<u>Proof</u>. Negate the parts of Theorem 5. <u>QED</u>.

The soundness of FSE follows from Theorem 6, because the clauses of the FSE Rule are the clausal forms of parts of Theorem 6. (Except for the common part E in (4') and (5')).

For example, if both $\underset{\cdot}{\leq}$ and $\underset{\cdot\cdot}{\leq}$ in (4') are <, then $\underset{\cdot\cdot\cdot}{\leq}$ is <, and (except for E) (4') and (5') are the clausal forms of (6.3) and (6.4), and if one of $\underset{\cdot}{\leq}$ and $\underset{\cdot\cdot}{\leq}$ in (4') is $\leq$, then $\underset{\cdot\cdot\cdot}{\leq}$ is $\leq$, and (except for E) (4') is the clausal form of (6.1) and (5'), with its $\underset{i}{\leq}$ and $\underset{j}{\leq}$ replaced by < (see remark at the end of the FSE rule), is the clausal form of (6.2).

<u>Completeness</u>. The FSE-elimination rule is obviously complete since the FSE-resolvent C' is does not replace C, but is added as another clause.

Replacement of C by its FSE-resolvent.

We cannot replace C by C' in the FSE rule, becasue the implication in Theorem 1 in this Section cannot be reversed. For example, if

$$f(x) = \sin x$$

$$u = -2, \quad w = +2,$$

we see that

$$\exists x(u < f(x) < w)$$

but not

$$\exists y \exists z(f(y) \leq u < w < f(z)).$$

This is unfortunate because a replacement rule (such as the VE-Rule [1]) is much more powerful than an addition rule, because the prover can cease to generate resolvents from the clause that is replaced. In the case of FSE, we can only require that the program "prefer" the resolvent C' over C but not discard C altogether. Such "preferring" can be implemented by an agenda mechanism such as that used in [3].

Intermediate Value Theorem.

Since FSE is based on IMV (the intermediate value theorem), it is clear that one can get the same effect as FSE by simply adding IMV as an addition hypothesis. But this seems to greatly increase the proof search, at least for the examples given in this paper. More experience, in the proofs of harder theorems, is needed before the effectiveness of FSE can be fairly evaluated.

For example, the following is a proof of Theorem LS2, from Section 3, using IMV.

1. $a < b$
2. $b < c$
3. $f(b) < f(a)$
4. $f(a) < f(c)$
5. $y \le z \vee x \le y \vee f(z) \le f(y) \vee f(z) \le f(x)$

6. $t \le s \vee u \le f(s) \vee f(t) \le u \vee f(z_{stu}) = u$
7. " " " $\vee\ s < z_{stu}$
8. " " " $\vee\ z_{stu} < t$

---

9. $t \le b \vee u < f(b) \vee f(t) \le u \vee f(z_{btu}) = u$ 6, 3, b/s
10. $c \le b \vee u < f(b) \vee f(a) < u \vee f(z_{bcu}) = u$ 9, 4, c/t
11. $c \le b \vee f(a) < f(b) \vee f(z_0) = f(a)$ 10, f(a)/u
where $z_0 = z_{bcf(a)}$
12. $f(z_0) = f(a)$ 11, 2, 3

13.	$b < z_0$ (similar to Steps 9-12)	7, 3, 4, 2, 3
14.	$z_0 < c$ (similar to Steps 9-12)	8, 3, 4, 2, 3
15.	$f(b) < f(z_0)$	12, 3, sub =
16.	$t \leq s \vee u \leq f(s) \vee f(t) \leq u \vee$ $y \leq z \vee z_{stu} \leq y \vee u \leq f(y) \vee f(z) \leq u$	5, 6, z_{stu}/x, sub =
17.	$t \leq s \vee u \leq f(s) \vee f(t) \leq u \vee$ $y \leq z \vee s < y \vee u \leq f(y) \vee f(z) \leq u$	16, 7
	(Now u is eligible in 16 and can be eliminated)	
18.	$t \leq s \vee y \leq z \vee s < y \vee$ $f(t) \leq f(s) \vee f(t) \leq f(y) \vee f(z) \leq f(s) \vee f(z) \leq f(y)$	17, VE, u
19.	$z_0 \leq b \vee b \leq a \vee b < b$	18, 15, 3, z_0/t, b/s, b/y, a/z
20.	$\square$	19, 13, 1

Even the proof of IMV2 is easier using the FSE rule.

Eligibility.

In VE it is required that the variable x being eliminated from a clause be eligible in that clause. (See definition above). While no such requirement is explicitly given for FSE, an implied one is there, in that we require that x (the argument in f(x)) not occur in the c_i, d_j, or E. But the symbol f itself can indeed occur in other parts of C with other arguments.

References.

1. W. W. Bledsoe and Larry M. Hines. Variable Elimination and Chaining in a Resolution-based Prover for Inequalities. University of Texas Math Department Memo ATP-56A, April 1980. Proc. 5th. Conference on Automatic Deduction, Les Arcs, France, July 1980, Springer-Verlag.

2. W. W. Bledsoe, Kenneth Kunen, and Robert Shostak. Completeness Results for Inequality Provers. University of Texas Math Department Memo ATP-65. In preparation.

3. Mabry Tyson. APRVR: A Priority-ordered Agenda Theorem Prover. Ph.D. Dissertation, Department of Computer Science, University of Texas, August 1981.