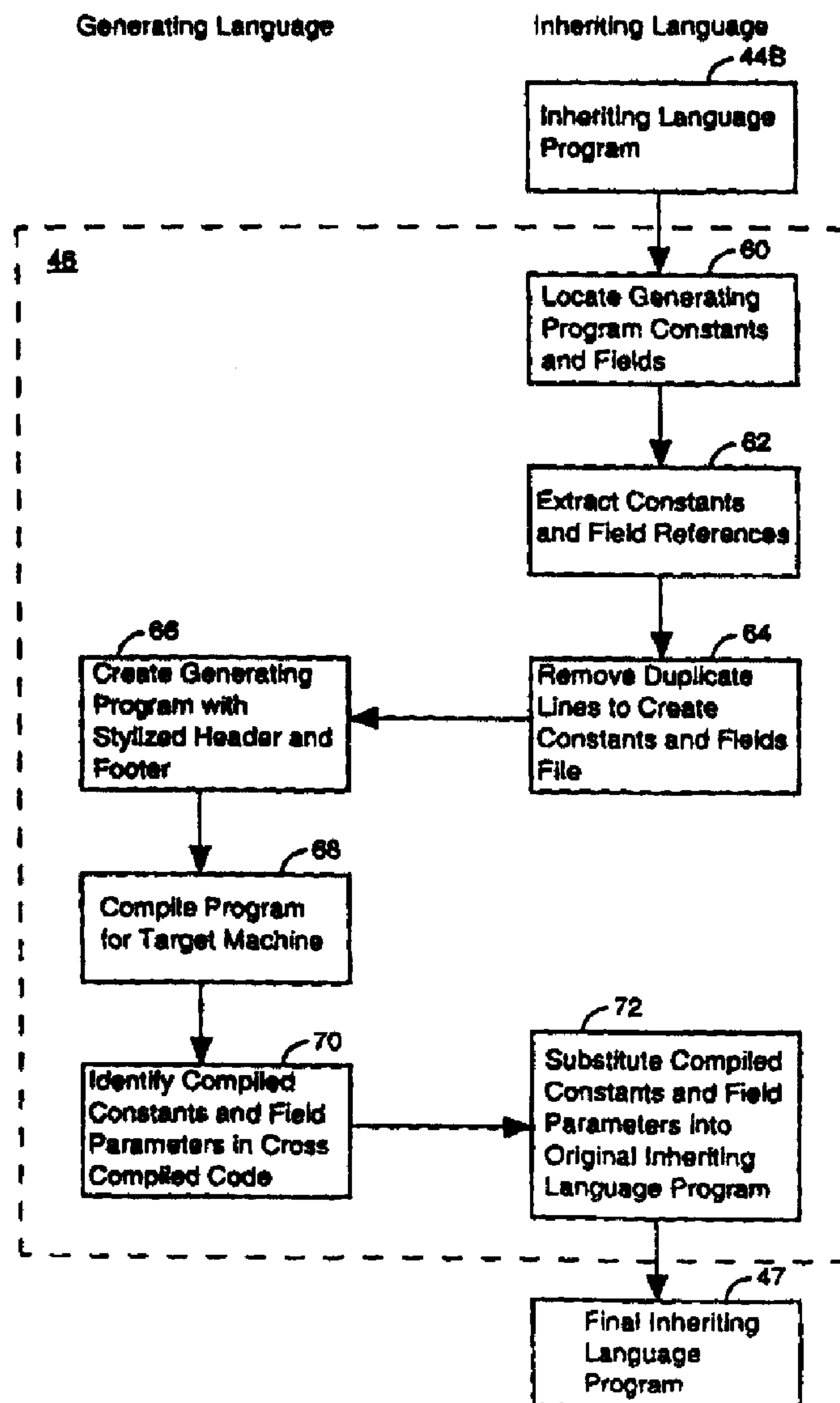




(72) YELLIN, FRANK, US
(72) LONG, DEAN R. E., US
(72) TUCK, RICHARD D., US
(71) SUN MICROSYSTEMS, INC., US
(51) Int.Cl.⁶ G06F 9/445
(30) 1997/12/12 (08/989,848) US
(54) **APPAREIL ET PROCEDE DE COMPILEATION CROISEE D'UN
CODE SOURCE**
(54) **APPARATUS AND METHOD FOR CROSS-COMPILING
SOURCE CODE**





(21) (A1) **2,280,698**
(86) 1998/12/10
(87) 1999/06/24

(57) L'invention concerne un procédé permettant la compilation croisée de programmes informatiques consistant à extraire des constantes et des références (62) de champ d'un programme informatique (44B) récepteur écrit dans un premier langage informatique. Les constantes et les références de champ extraites ont trait à un programme informatique (60) générateur écrit dans un second langage informatique. Un nouveau programme générateur (66) écrit dans le second langage informatique est ensuite créé à l'aide des constantes et des références de champ. Le nouveau programme est ensuite compilé pour un ordinateur cible (68) de manière à obtenir des valeurs (70) de constantes compilées. Ces valeurs de constantes compilées sont ensuite substituées à l'intérieur du programme informatique récepteur afin de produire un programme informatique (47) récepteur final.

(57) A method of cross-compiling computer programs includes the step of extracting constants and field references (62) from an inheriting computer program (44B) written in a first computer language. The extracted constant and field references refer to a generating computer program (60) written in a second computer language. A new generating program (66) in the second computer language is then created using the constants and field references. The new program is then compiled for a target computer (68) to ascertain compiled constant values (70). The compiled constant values are then substituted into the inheriting computer program to produce a final inheriting computer program (47).



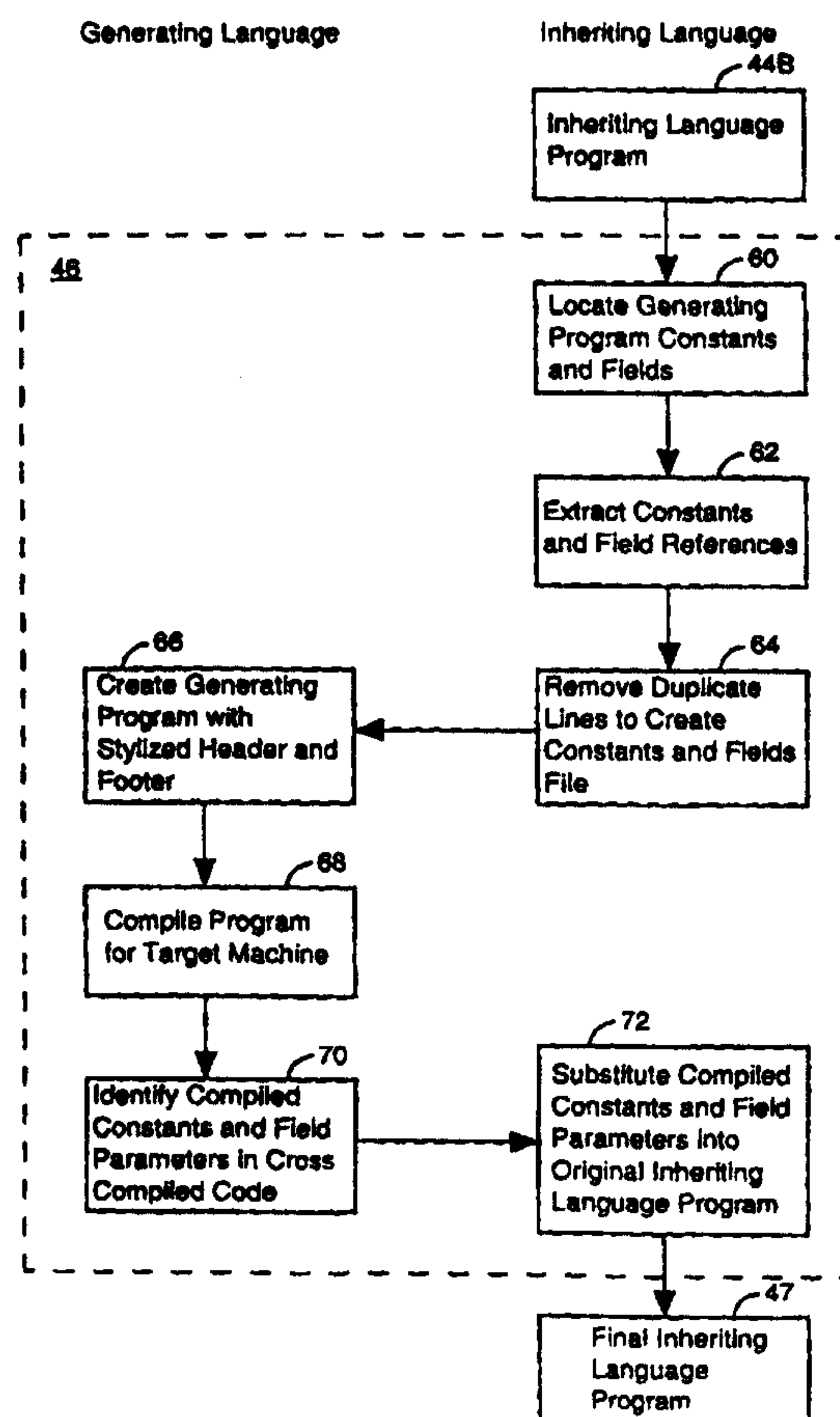
PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G06F 9/445	A1	(11) International Publication Number: WO 99/31583 (43) International Publication Date: 24 June 1999 (24.06.99)
(21) International Application Number: PCT/US98/26285 (22) International Filing Date: 10 December 1998 (10.12.98) (30) Priority Data: 08/989,848 12 December 1997 (12.12.97) US (71) Applicant: SUN MICROSYSTEMS, INC. [US/US]; 901 San Antonio Road, Palo Alto, CA 94303 (US). (72) Inventors: YELLIN, Frank; 510 Beresford Avenue, Redwood City, CA 94061 (US). LONG, Dean, R., E.; P.O. Box 268, Boulder Creek, CO 95006 (US). TUCK, Richard, D.; 343 Hill Street, San Francisco, CA 94114 (US). (74) Agents: MISROCK, S., Leslie et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036-2711 (US).		(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>

(54) Title: APPARATUS AND METHOD FOR CROSS-COMPILING SOURCE CODE**(57) Abstract**

A method of cross-compiling computer programs includes the step of extracting constants and field references (62) from an inheriting computer program (44B) written in a first computer language. The extracted constant and field references refer to a generating computer program (60) written in a second computer language. A new generating program (66) in the second computer language is then created using the constants and field references. The new program is then compiled for a target computer (68) to ascertain compiled constant values (70). The compiled constant values are then substituted into the inheriting computer program to produce a final inheriting computer program (47).



APPARATUS AND METHOD FOR CROSS-COMPILING SOURCE CODE

Brief Description of the Invention

This invention relates generally to software system management tools. More particularly, this invention relates to a software system management tool that is used to combine different source code programs written in multiple computer languages into a
5 single program.

Background of the Invention

Software system management tools define software modules, their interdependencies, and the rules that are required to combine the software modules
10 into a final software program. Software system management tools are necessary in view of the fact that a large software program typically includes a large number of individual software modules in the form of source code programs that are compiled and then linked into a single executable final software program. The large number of individual software modules and their individual processing requirements result in the
15 need for sophisticated software system management tools. For example, each source code program may require a distinct compiler. Thus, a mechanism must be developed to coordinate a source code program with the appropriate compiler. There are many commercially available software system management tools.

Software system management tools often combine computer code written in
20 one language with computer code written in a second language. It is often necessary for the code written in one language (referred to herein as the inheriting language) to make use of constants and data structures defined by code written in a second language (referred to herein as the generating language). In addition, the values of these

WO 99/31583

PCT/US98/26285

constants and the exact layout of these data structures often depend on the target machine for which the code is being compiled. The layout of the data structures and the value of the constants may be different between the machine doing the compiling (the host computer), and the machines for which the code is being compiled (the target computer). In some cases, the code may need to be compiled for multiple target machines that have different values and layouts.

The layout of these data structures and the value of the constants may change during the development process. It is important that it be easy to recompile the code written in the inheriting language. However, the specific constants and data structures referenced by the inheriting code are liable to change as the inheriting code is developed. In addition, new constants can be accessed and new fields in data structures might be accessed.

It would be highly desirable to provide a software system management tool that allows an inheriting language to make use of information from a generating language, regardless of the host computer or of the target computer. That is, the software system management tool should provide an automatic technique for determining the constants and fields needed by the inheriting code. Further, the software system management tool should make it easy to recompile the inheriting code automatically. Such a software system management tool would simplify the development of source code written in different languages.

Summary of the Invention

A method of cross-compiling computer programs includes the step of extracting constants from an inheriting computer program written in a first (inheriting) computer language. The extracted constants refer to a generating computer program written in a second (generating) computer language. A new program in the second computer language is then created using the constants. The new program is then compiled for a target computer to ascertain compiled constant values. The compiled constant values are then substituted into the inheriting computer program to produce a final inheriting computer program.

The invention provides a software system management tool

WO 99/31583

PCT/US98/26285

that allows an inheriting computer program to make use of information from a generating computer program, regardless of the computer on which the compilation will take place or of the target computer. That is, the cross-compiler of the invention provides an automatic technique for determining the constants and fields needed by the inheriting computer program. Further, the cross-compiler facilitates automatic recompilation of the inheriting computer program. The technique of the invention efficiently compiles source code written in different languages.

Brief Description of the Drawings

For a better understanding of the nature and objects of the invention, reference should be made to the following detailed description taken in conjunction with the accompanying drawings, in which:

FIG. 1 illustrates a software system management apparatus in accordance with an embodiment of the invention.

FIG. 2 illustrates processing operations associated with the software system management apparatus of Fig. 1.

FIG. 3 is a more detailed depiction of the operation of the cross compiler of Fig. 2.

FIG. 4 illustrates processing steps associated with an embodiment of a cross compiler constructed in accordance with the invention.

FIG. 5 illustrates a software system management apparatus in accordance with another embodiment of the invention.

Like reference numerals refer to corresponding parts throughout the several views of the drawings.

Detailed Description of the Invention

Fig. 1 illustrates a software system management apparatus 20. The apparatus 20 includes processor 22 which communicates with a memory module 24 and a set of input/output devices (e.g., keyboard, mouse, monitor, printer, etc.) 26 through a bus 28. An individual working at the input/output devices 26 inputs a software system description, which is down-loaded into a memory module 24. The software system description 30 is an integrated description of software modules, their interdependencies, and the rules that are used to construct an executable program.

The memory module 24 stores a number of executable programs including a software user interface 32 (such as a graphical user interface) and a rule execution

WO 99/31583

PCT/US98/26285

engine 34, which produces system construction commands 36. The executable programs stored in the memory module 24 also include a tool controller 38, compilers 40, linkers 42, and a cross compiler 46. The tool controller 38 receives the system construction commands 36 and existing program files 44. The tool controller 38
5 operates in conjunction with the compilers 40, linkers 42, and cross compiler 46 to execute the system construction commands 36 on the existing program files 44 to yield an executable program 48.

The processing associated with this software system management apparatus 20 is more fully appreciated with reference to Fig. 2. The input data to the apparatus is a
10 software system description specifying a set of software source code modules, their interdependencies, and the rules that are required to combine the software modules into a final software program. The software user interface 32 receives this information and passes it to a rule execution engine 34, which produces a set of system construction commands 36. For instance, if the rule-based software configuration
15 description 30 specifies that files A, B, and C are to be combined to form file D, then the system construction commands 36 would constitute a set of commands that could be executed by the tool controller 38 to compile and link existing files A, B, and C to form file D.

Thus, the tool controller 38 executes the system construction commands 36 by
20 interacting with the compilers 40, linkers 42, existing program files 44, and a cross compiler 46. After the tool controller 38 has executed the system construction commands 36, an executable program 48 is produced.

The present invention is directed toward the operation of the cross compiler 46. The other tools with which it operates (e.g., the other elements of Fig. 2) are generally
25 known in the art.

The cross compiler 46 interacts with the tool controller 38 to execute system construction commands 36. More particularly, the cross compiler 46 interprets the system construction commands 36 to identify when two existing program files written in different languages are to be compiled into a single program. That is, the cross
30 compiler 46 identifies when an existing program file written in a generating language is to be combined with an existing program file written in an inheriting language. The cross compiler 46 allows the inheriting language to make use of information from the generating language. That is, the cross compiler 46 provides an automatic technique

WO 99/31583

PCT/US98/26285

for determining the constants and fields needed by the inheriting code, regardless of the target machine on which the program will be compiled. The cross compiler 46 does this in such a manner that it is still easy to automatically recompile the inheriting code. In sum, the cross compiler 46 facilitates the compilation of source code written in different languages. As a result, the other software system management tools of Fig. 2 operate more efficiently to produce an executable program 48.

This operation is more fully appreciated with reference to Fig. 3. The cross compiler 46 processes commands from the tool controller 38 by accessing a generating language program 44A and an inheriting language program 44B of the existing program files 44. In particular, a constants and fields locator software module 50 identifies in the inheriting language program 44B references to the generating language program 44A. A constants and fields processor module 52 then extracts the constants and fields and preferably removes duplicate references, thereby producing a constants and fields file. A generating language program constructor module 54 then uses the constants and fields file to create a new and stylized generating language program. A constants and fields value identifier module 56 identifies compiled constant and field values in cross compiled code corresponding to the new generating language program. An inheriting language program modifier module 58 is then used to substitute the compiled constants and field parameters into the inheriting language program. This results in a final inheriting language program 47, which uses the constants and fields from the generating language program 44A. The tool controller 38 may then use the final inheriting language program 47 in connection with the compilers 40 and linkers 42 to produce an executable program 48.

The particular steps executed by an embodiment of the cross compiler 46 of the invention are illustrated in Fig. 4. Fig. 4 illustrates the cross compiler 46 of the invention receiving an inheriting language program 44B and generating a final inheriting language program 47, consistent with the description of Fig. 3.

The first processing step for the cross compiler 46 of Fig. 4 is to modify the inheriting language program to identify references to the generating language program (step 60). This operation may be performed by the constants and fields locator module 50. The references to the generating language program are in the form of constants and fields. Modifications to the constants and fields, discussed below, makes them implementation independent.

WO 99/31583

PCT/US98/26285

The next processing step executed by the cross compiler 46 is to extract the identified constants and fields (step 62). Thereafter, duplicate lines containing the identified constants and fields are preferably removed to create a constants and fields file (step 64). These operations may be performed with the constants and fields processor module 52.

As shown in Fig. 4, all of the processing up to this point has been directed toward the inheriting language. The next processing step is to create, from the constants and fields file written in the inheriting language, a corresponding generating language program with a stylized header and footer (step 66). This operation entails converting each line from the constants and fields file of step 64 into a macro, as discussed below. This operation may be performed by the generating language program constructor 54.

The program is then compiled for a target machine or machines (step 68). This aspect of the invention provides cross compiled code for multiple target machines. Observe that the cross compiled code only relates to the constants and fields of the original inheriting language program 44B. Thus, when changes are made to the original inheriting language program 44B, it is relatively efficient to recompile the code in accordance with the invention. That is, the entire inheriting language program 44B need not be recompiled, rather only the constants and fields file needs to be recompiled.

The next processing step is to identify compiled constants and field parameters in the code (step 70). This operation may be performed with the constants and fields value identifier 56. The identified compiled constants and field parameters are then combined with the original inheriting language program (step 72) to yield the final inheriting language program 47. This step may be performed with the inheriting language program modifier module 58.

While prior art techniques exist for automatically generating "constants files" for specific constants and field offsets, those skilled in the art will recognize that the present invention expands beyond such rudimentary implementations because it operates in an environment where compilation is for machines other than the machine on which the compilation is being performed. In accordance with the invention, the inheriting language code is automatically rewritten with correct constants and field offsets for the target machine.

WO 99/31583

PCT/US98/26285

The features and benefits of the invention are more fully appreciated with reference to a more specific example. The invention is disclosed through an example wherein the generating language is "C" and the inheriting language is SPARC Assembler. However, those skilled in the art will appreciate that the techniques of the invention can be used for any inheriting and generating languages.

In accordance with an embodiment of the invention, the software system description 30 includes a MAKE file with the following lines.

```

10 (1) $(OBJDIR)/make_struct_offsets.s: extract_offsets.nawk opcodes.wide opcodes.h
    (2) $(OBJDIR)/make_struct_offsets.s: executeJava_sparc.m4.s
    (3)  $(M4) -DJAVAS -DEXTRACT_OFFSETS $< | sort -u | nawk -f
        extract_offsets.nawk > $(OBJDIR)/make_struct_offsets.c
    (4)  $(CC) $(CFLAGS) $(INCLUDES) -S $(OBJDIR)/make_struct_offsets.c -o
15      $@

```

The foregoing and following computer code of this patent document contains material which is subject to copyright protection. The copyright owner has no objection to the facsimile reproduction of the computer code, as it appears in the Patent and Trademark Office files or records, but otherwise reserves all rights established under applicable copyright laws.

Lines (1) and (2) of the code indicate that if any of the four files (1) "extract_offsets.nawk", (2) "opcodes.wide", (3) "opcodes.h", or (4) "executeJava_sparc.m4.s" change, then line (3) should be executed. In other words, the software system description 30 will result in system construction commands 36 that are processed by the tool controller 38. If the tool controller 38 identifies that any of the four files have changed, it will invoke the cross compiler 46.

The file "extract_offsets.nawk" contains portions of the cross compiler 46 of the invention. In particular, it contains instructions to implement steps 62 and 66 of Fig. 4. If this file changes, then the result should be re-computed. If the files "opcodes.wide" or "opcodes.h" have changed, then it's possible that some of

WO 99/31583

PCT/US98/26285

the constants have changed value. Any change to the file "executeJava_sparc.m4.s" means that there may be new constants or field offsets of interest or that some previously interesting constants or field offsets are no longer important.

If one of the four enumerated files has changed, the following code is executed:

5

```

m4 -DJAVAOS -DEXTRACT_OFFSETS executeJava_sparc.m4.s |
    sort -u |
    nawk -f extract_offsets.nawk > $(OBJDIR)/make_struct_offsets.c
10 cc <various flags> -S $(OBJDIR)/make_struct_offsets.c -o $@

```

This code corresponds to the operations performed by the cross compiler 46. The code causes "executeJava_sparc.m4.s" to run through processor m4 with the two
15 flags JAVAOS and EXTRACT_OFFSETS set to 1. This causes the processor m4 to modify the inheriting language program to identify generating program references (step 60). Thus, the processor m4 can be thought of as the constants and fields locator module 50. For example, all references to constants are written as "DEFINED_CONSTANT (<constant>)". Further, all references and updates to fields
20 are written in such a way that it is unnecessary to know the size, signedness, or offset of the particular field in the record. For example, each "GET_FIELD (<reg>, <structure>, <field>, <result>)" term is converted into a line of the form
LOAD_STORE <structure> <field>.

These operations are implemented as follows. The file executeJava_sparc.m4.s
25 is written in a highly stylized format, suitable for m4. The first line of the file is "ifdef('EXTRACT_OFFSETS', 'divert(-1))", which says that if the flag EXTRACT_OFFSETS is defined, then throw out all input unless otherwise instructed. Later in the file, there are m4 directives such that if EXTRACT_OFFSETS is defined, the entire contents of the file are ignored except for occurrences of

```

30     DEFINED_CONSTANT(<baz>),
        EXTRACT_OFFSET (<mystruct>,<myfield>),
        GET_FIELD(<base>,<mystruct>,<myfield>,<reg>), and

```

WO 99/31583

PCT/US98/26285

SET_FIELD(<value>,<base>,<mystruct>,<myfield>), where each of the lower-case strings is, in fact, a text string.

Each occurrence of DEFINED_CONSTANT(<baz>) is output as "CONST <baz>". Each occurrence of EXTRACT_OFFSET(<mystruct>,<myfield>) is output
5 as "FIELD <mystruct>,<myfield>". Each occurrence of GET_FIELD(<base>,<mystruct>,<myfield>,<reg>) and SET_FIELD(<value>,<base>,<mystruct>,<myfield>) is output as "LDST <mystruct>,<myfield>".

The code in the file is written such that it makes use of these macros. For example, If reg1 contains a pointer to a methodblock structure, and its ClassClass field
10 is required, GET_FIELD(reg1, methodblock, ClassClass, reg2) is written. If reg3 is to be compared to the constant "opc_wide", "cmp reg3, DEFINED_CONSTANT (opc_wide)" is written. In particular, all references to constants are wrapped inside DEFINED_CONSTANT(..). To extract the field from a pointer, the GET_FIELD instruction is used. To set the field of a pointer, the SET_FIELD instruction is used.
15 To get a pointer to the field of a pointer, "STRUCTURE_OFFSET(..)" is added to the pointer.

The foregoing processing associated with steps 60 and 62 results in a line for each reference to a constant and a line for each reference to a field. If there are multiple references to a field, each of those references will generate a separate line. It
20 is easier, although not necessary, to process each defined constant and each field only once. Thus, preferably, duplicate lines are deleted to create a constants and fields file (step 64). This may be accomplished by passing the result through "sort -u", a standard utility, which sorts lines and deletes duplicate lines. The sorting of lines is an artifact of the "sort" utility and therefore is unimportant.

25 A new generating language program with a stylized header and footer is then created (step 66). Each line is associated with a macro call, which is defined in the header. For example, the line LOAD_STORE <structure> <field> becomes LOAD_STORE (<structure>,<field>). The specific header and trailer information and the exact code generated for each line is dependent on the generating language.

30 This operation is more fully appreciated with the following example. The constants and fields file created by step 64 is processed by a text processor called "nawk". The program that the text processor runs is contained in the file

WO 99/31583

PCT/US98/26285

“extract_offsets.nawk”. The “nawk” program generates a C program (a generating language program) that has three parts. The first part is a stylized header, for example:

```

#include "oobj.h"
5  #include "interpreter.h"
#include "opcodes.h"
#include "tree.h"
#include "typecodes.h"
#include "stddef.h"
10 #define SHOWME(structure, name) \
    { struct structure *temp; \
        asm("SET_STRUCTURE_INFO(`" #structure "`,`" #name \
            ",`%0`,`%1`,`%2`)" \
            :: "n" (sizeof(temp -> name)), \
15      "n" (offsetof(struct structure, name)), \
            "n" ((typeof(temp -> name))(~0 < 0)); \
    }
#define FIELDOFF(structure, name) \
    asm("SET_FIELD_OFFSET(`" #structure "`,`" #name "`,`%0`)" \
20      :: "n" (offsetof(struct structure, name)) )
#define CONSTANT(name) asm("SET_VALUE(`" #name "`,`%0`)" :: "n"
                                ((int)name) )

main(int argc, char ** argv){
    asm("! File automatically generated by m4 and nawk" );
25  asm("! Do not bother editing this! Find the source!" );

```

The processing of this header information is discussed below. After this stylized header is created, the following transformations are made for the constants and fields: “CONST <baz>” is assigned to “CONSTANT(<baz>”, “LDST <struct><myfield>” is assigned to “SHOWME(<mystruct>,<myfield>”, and “FIELD <mystruct><myfield>” is assigned to “FIELDOFF(<mystruct>,<myfield>”. A final closing “}” is then appended as a footer. The result is a program “make_struct_offsets.c”, which can be compiled by a C compiler. In other words, at

WO 99/31583

PCT/US98/26285

this point, step 66 of Fig. 4 is completed. The next processing step is to compile the program for the target machine (step 68).

In accordance with the invention, the compiler is given specific switches to tell it to generate assembly language for the target machine, rather than to generate a binary file. In particular, the stylized generating language file from step 66 is specifically designed so that it generates highly stylized assembly-language code. The code may be stylized to the point that it cannot be actually assembled into machine code. All that is necessary is that the resulting assembly language code be machine-parseable so that one can determine (1) the value that the compiler gave to each constant and (2) the size (number of bytes), offset, and signedness of each of the fields.

The result of this step will always be assembly language, regardless of the inheriting language. It is coincidental that in the present example, that the original inheriting language is also assembly language.

At this point, the .s file has the following form for each of the SHOWME, CONSTANT, and FIELDOFF items, respectively:

```
SET_STRUCTURE_INFO (<structure>, <field>, <size>, <offset>, <signedness>)
SET_VALUE (<name>, <value>)
SET_FIELD_OFFSET(<structure>, <field>, <offset>).
```

The process by which this transformation takes place is as follows. Because of the #define, each line of the form CONSTANT (<baz>) gets turned into:

```
"asm ( "SET_VALUE("`" "baz" "'", '%0')" : : "n" ( ( int ) baz ) )".
```


WO 99/31583

PCT/US98/26285

Similarly, each line of the form FIELDOFF(mystruct,myfield) gets turned into:

```
“asm ( "SET_FIELD_OFFSET(`" "mystruct" ",`" "myfield" ",`%0)" : : "n" (
  offsetof ( struct mystruct , myfield ) ) )”
```

5

Each line of the form SHOWME(mystruct,myfield) gets turned into:

```
“{ struct mystruct * temp ; asm ( "SET_STRUCTURE_INFO(`" "mystruct"
  ",`"
  "myfield" ",`%0`,`%1`,`%2)" : : "n" ( sizeof ( temp -> myfield ) ) ,
  "n" ( offsetof ( struct mystruct , myfield ) ) , "n" ( ( typeof ( temp -> myfield
    ) ) ( ~ 0 ) < 0 ) ) ; }”
```

10

These instructions make use of the Gnu C-compiler (or other similar C compiler) facility for generating specific assembly language. However, since the compilation to assembly code is solely for the purpose of defining constants and field offsets, the inline assembly code that is generated does not have to be executable. In particular, the inline assembly code:

15

```
20 “asm ( "SET_VALUE(`" "baz" ",`%0)" : : "n" ( ( int ) baz ) )”
```

generates, in assembly language:

```
“SET_VALUE(`baz`,`23`)”
```

25

where 23, in this case happens to be the value of baz.

The inline assembly code :

```
30 “asm ( "SET_FIELD_OFFSET(`" "mystruct" ",`" "myfield" ",`%0)" : : "n" (
  offsetof ( struct mystruct , myfield ) )”
```

WO 99/31583

PCT/US98/26285

generates:

```
"SET_FIELD_OFFSET(`mystruct`,`myfield`,`16`)"
```

5 where 16, for example, is the offset of "myfield" in the "mystruct" structure.

The inline assembly code:

```

    "{ struct mystruct * temp ; asm ( "SET_STRUCTURE_INFO(" "mystruct"
    "" ,`"
10    "myfield" "" ,`%0`,`%1`,`%2`)" : : "n" ( sizeof ( temp -> myfield ) ) ,
    "n" ( offsetof ( struct mystruct , myfield ) ) , "n" ( ( typeof ( temp -> myfield
                                   ) ) ( ~ 0 ) < 0 ) ) ; }"

```

turns into:

```
15    "SET_STRUCTURE_INFO(`mystruct`,`myfield`,`2`,`16`,`1`)"
```

In this case, the 16 is again the offset of the field. The "2" indicates that this is a 2-byte quantity. The "1" indicates that this is a signed value (an unsigned value would assign a "0" to this field).

20 Observe that the .s file generated will have additional information. It will specifically have the lines:

```

"! File automatically generated by m4 and nawk"
"! Do not bother editing this! Find the source!"

```

25

These lines are created by the asm declarations at the start of the main routine defined in step 66. The code will also include function prologues and other information.

If one intends to cross compile for a different platform, then the .c file is compiled into assembly for the target platform. The macros are written in such a way
 30 that all the constant values, field offsets, etc., are generated for the target machine.

The .s file is processed by extracting lines that contain the word SET. For example, one can create the file "executeJava_sparc.include" by extracting lines

WO 99/31583

PCT/US98/26285

containing the word SET. This can be done using the UNIX command "grep", e.g.,
 "grep SET extractoffsets.s executeJava_sparc.include".

The next processing step is to combine the compiled constants and field parameters with the original inheriting language program. That is, for each occurrence
 5 of `DEFINED_CONSTANT(<name>)`, replace it with the exact value of the constant. For each occurrence of `GET_FIELD(<reg>,<structure>, <field>, <result>)` convert it into specific inheriting language instructions necessary to access the value. Thus, the final inheriting language program 47 is a combination of the original program text with the result of the processing of the cross compiler 46.

10 These operations can be performed as follows. Run "executeJava_sparc.m4" through the processor m4 a second time. However, this time the `EXTRACT_OFFSET` flag is not turned-on. This causes m4 macros called `SET_FIELD_OFFSET`, `SET_STRUCTURE_INFO`, and `SET_VALUE` to be defined. In addition, the previously created file "ExecuteJavaStructOffsets.include" is read in. Each line of the
 15 file is interpreted using the macro definitions defined in step 60. The remainder of the file "executeJava_sparc.m4" is handled normally by the processor m4. In particular, the four macros

```

    DEFINED_CONSTANT(baz)
20    EXTRACT_OFFSET(mystruct,myfield)
    GET_FIELD(base,mystruct,myfield,reg)
    SET_FIELD(value,base,mystruct,myfield)
  
```

have completely different meanings than they did in step 60.

25 For constants, `SET_VALUE('name','value')` is turned into `define('name,value')` and `DEFINED_CONSTANT(name)` is turned into "name". Hence, if one has `SET_VALUE('baz','23')` from the include file, and the occurrence of `DEFINED_CONSTANT(baz)` somewhere in the text, the macro pre-processor will turn `DEFINED_CONSTANT(baz)` into `baz` and then that into 23.

30 For field offsets, `SET_FIELD_OFFSET(struct,field,offset)` is turned into the assembler directive: "`struct.field = offset`". Similarly, "`STRUCT_OFFSET(struct,field)`" is turned into "`struct.field`". The next effect is that every occurrence of "`STRUCT_OFFSET(struct,field)`" becomes "`struct.offset`", which the assembler can

WO 99/31583

PCT/US98/26285

then replace with the correct value. For field accessors and setters,

“SET_STRUCTURE_INFO(`mystruct`,`myfield`,`4`,`16`,`1`)” defines two macros:

“GET_FIELD.mystruct.myfield(base, result)” and “SET_FIELD.mystruct.myfield(base, value)”. The definition of:

5

“GET_FIELD.mystruct.myfield(base, result)” is “ld.2.1 [base+16], value”

“SET_FIELD.mystruct.myfield(base, value)” is “st.2.1 value,[base+16], ”

where the “4”, “16”, and “1” are extracted from the fields. Separately, the opcodes

10 st.1.0, st.1.1, st.2.0, st.2.1, st.4.0, st.4.1, ld.1.0, ld.1.1, ld.2.0, ld.2.1, ld.4.0, ld.4.1, are defined to be the appropriate opcodes for storing and loading the appropriately sized field of the appropriate sign. For example, on SPARC, “ld.2.1” is defined to be “ldsh” (load a signed half-word). Similarly,

15

GET_FIELD(base,mystruct,myfield, reg)

SET_FIELD(value,base,mystruct,myfield)

would be defined to be

20

GET_FIELD.mystruct.myfield(base,reg)

SET_FIELD.mystruct.myfield(base,value)

respectively. Hence

25

GET_FIELD(base,mystruct,myfield,reg)

turns into

GET_FIELD.myfield.mystruct(base,reg)

WO 99/31583

PCT/US98/26285

which turns into

ld.2.1 [base+16],reg

5 which turns into

ldsh [base+16], reg,

which is the desired result of an inheriting language program instruction that has
10 utilized information from a generating language program.

Fig. 5 illustrates an alternate apparatus for practicing the invention. Fig. 5 corresponds to Fig. 1, but includes a storage device 80, a communications interface 82, a network link 84, and a network 86. The programs stored in the memory 24 may be downloaded from a computer-readable medium associated with the storage device 80,
15 or alternately, may be executed from the computer-readable medium associated with the storage device 80. The term "computer-readable medium" refers to any medium that participates in providing instructions to the processor 22 for execution. Such a medium may take many forms, including but not limited to, non-volatile media, volatile media, and transmission media. Non-volatile media includes, for example,
20 optical or magnetic disks, associated with the storage device 110. Volatile media includes dynamic memory. Transmission media includes coaxial cables, copper wire and fiber optics, including the wires that comprise bus 28. Transmission media can also take the form of acoustic or light waves, such as those generated during radio-wave and infra-red data communications.

25 Common forms of computer-readable media include, for example, a floppy disk, a flexible disk, hard disk, magnetic tape, or any other magnetic medium, a CD-ROM, any other optical medium, punchcards, papertape, any other physical medium with patterns of holes, a RAM, a PROM, EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave as described below, or any other medium
30 from which a computer can read.

Various forms of computer readable media may be involved in carrying one or more sequences of one or more instructions to processor 22 for execution. For example, the instructions may initially be carried on a magnetic disk of a remote

WO 99/31583

PCT/US98/26285

computer. The remote computer can load the instructions into its dynamic memory and send the instructions over a telephone line using a modem. A modem local to the computer system 20 can receive the data on the telephone line and use an infra-red transmitter to convert the data to an infra-red signal. An infra-red detector coupled to the bus 28 can receive the data carried in the infra-red signal and place the data on bus 28. The bus 28 then carries the data to the memory 24, from which the processor 22 retrieves and executes the instructions. The instructions received by the memory 24 may optionally be stored on the storage device 80 either before or after execution by the processor 22.

10 The computer system 20 also includes a communication interface 82 coupled to the bus 28. The communication interface 82 provides a two-way data communication coupled to a network link 84 that is connected to a network 86. For example, the communication interface 82 may be an integrated services digital network (ISDN) card or a modem to provide a data communication connection to a corresponding type of telephone line. As another example, the communication interface 82 may be a local area network (LAN) card to provide a data communication connection to a compatible LAN. Wireless links may also be implemented. In any such implementation, the communication interface 82 sends and receives electrical, electromagnetic or optical signals that carry digital data streams representing various types of information.

20 The network link 84 typically provides data communication through one or more networks, represented by the network 86. For example, the network link 84 may provide a connection to a network 88 that includes a host computer operated as an Internet Service Provider (ISP). The ISP in turn provides data communication services through the world wide packet data communication network now commonly referred to as the "Internet". The network 86 uses electrical, electromagnetic or optical signals that carry digital data streams. The signals through the various networks and the signals on the network link 84 and through the communication interface 82, which carry the digital data to and from the computer system 20, are exemplary forms of carrier waves transporting the information.

30 The computer system 20 can send messages and receive data, including program code, through the network 86, the network link 84, and the communication interface 82. In the Internet example, a server on the network 86 may transmit a

WO 99/31583

PCT/US98/26285

requested code for an application program through the network 86, the network link 84, and the communication interface 82. The received code may be executed by the processor 22 as it is received and/or stored in the storage device 80, or other non-volatile storage for subsequent execution. In this manner, the computer system 20
5 may obtain application code in the form of a carrier wave.

The foregoing description, for purposes of explanation, used specific nomenclature to provide a thorough understanding of the invention. However, it will be apparent to one skilled in the art that the specific details are not required in order to practice the invention. In other instances, well known circuits and devices are shown
10 in block diagram form in order to avoid unnecessary distraction from the underlying invention. Thus, the foregoing descriptions of specific embodiments of the present invention are presented for purposes of illustration and description. They are not intended to be exhaustive or to limit the invention to the precise forms disclosed, obviously many modifications and variations are possible in view of the above
15 teachings. The embodiments were chosen and described in order to best explain the principles of the invention and its practical applications, to thereby enable others skilled in the art to best utilize the invention and various embodiments with various modifications as are suited to the particular use contemplated. It is intended that the scope of the invention be defined by the following Claims and their equivalents.

20

WO 99/31583

PCT/US98/26285

IN THE CLAIMS:

1. A method of cross-compiling computer programs, said method comprising the steps of:
 - 5 extracting constants from an inheriting computer program written in a first computer language that refer to a generating computer program written in a second computer language;
creating a new program in said second computer language using said constants;
compiling said new program for a target computer to ascertain compiled
10 constant values; and
substituting said compiled constant values into said inheriting computer program to produce a final inheriting computer program.
2. The method of claim 1 wherein said extracting step includes the step of
15 extracting fields from said inheriting computer program.
3. The method of claim 2 further comprising the step of removing duplicate fields and constants.
- 20 4. The method of claim 2 wherein said creating step includes the step of creating a new program that is machine-parseable to determine the value of each of said constants and the attributes of each of said fields without assembling said new program into machine code.
- 25 5. The method of claim 1 wherein said compiling step includes the step of compiling said new program for a plurality of target computers.
6. A computer readable medium to direct a computer to function in a specified manner, comprising:
30 a first set of instructions to extract constants from an inheriting computer program written in a first computer language that refer to a generating computer program written in a second computer language;

WO 99/31583

PCT/US98/26285

a second set of instructions to create a new program in said second computer language using said constants;

a third set of instructions to compile said new program for a target computer to ascertain compiled constant values; and

5 a fourth set of instructions to substitute said compiled constant values into said inheriting computer program to produce a final inheriting computer program.

7. The apparatus of claim 6 wherein said first set of instructions include instructions to extract fields from said inheriting computer program.

10

8. The apparatus of claim 7 wherein said first set of instructions include instructions to remove duplicate fields and constants.

9. The apparatus of claim 7 wherein said second set of instructions include
15 instructions to create a new program that is machine-parseable to determine the value of each of said constants and the attributes of each of said fields without assembling said new program into machine code.

10. The apparatus of claim 6 wherein said third set of instructions include
20 instructions to compile said new program for a plurality of target computers.

11. A computer readable medium to direct a computer to function in a specified manner, comprising:

a constants and fields locator module to identify in an inheriting computer
25 program references to constants and fields associated with a generating computer program;

a generating language program constructor module to create a new generating language program incorporating said constants and fields;

a constants and fields value identifier module to identify compiled constants
30 and fields in compiled code corresponding to said new generating language program;
and

WO 99/31583

PCT/US98/26285

an inheriting language program modifier module to substitute said compiled constants and fields into said inheriting language program to produce a final inheriting language program.

5 12. The apparatus of claim 11 wherein said constants and fields locator module removes duplicate fields and constants.

13. The apparatus of claim 11 wherein said generating language program
10 constructor module creates a new program that is machine-parseable to determine the value of each of said constants and the attributes of each of said fields without assembling said new program into machine code.

14. The apparatus of claim 11 further comprising a tool controller, a compiler, and a linker.

15

15. The apparatus of claim 11 wherein said computer readable medium is a memory connected to a system bus of a general purpose computer.

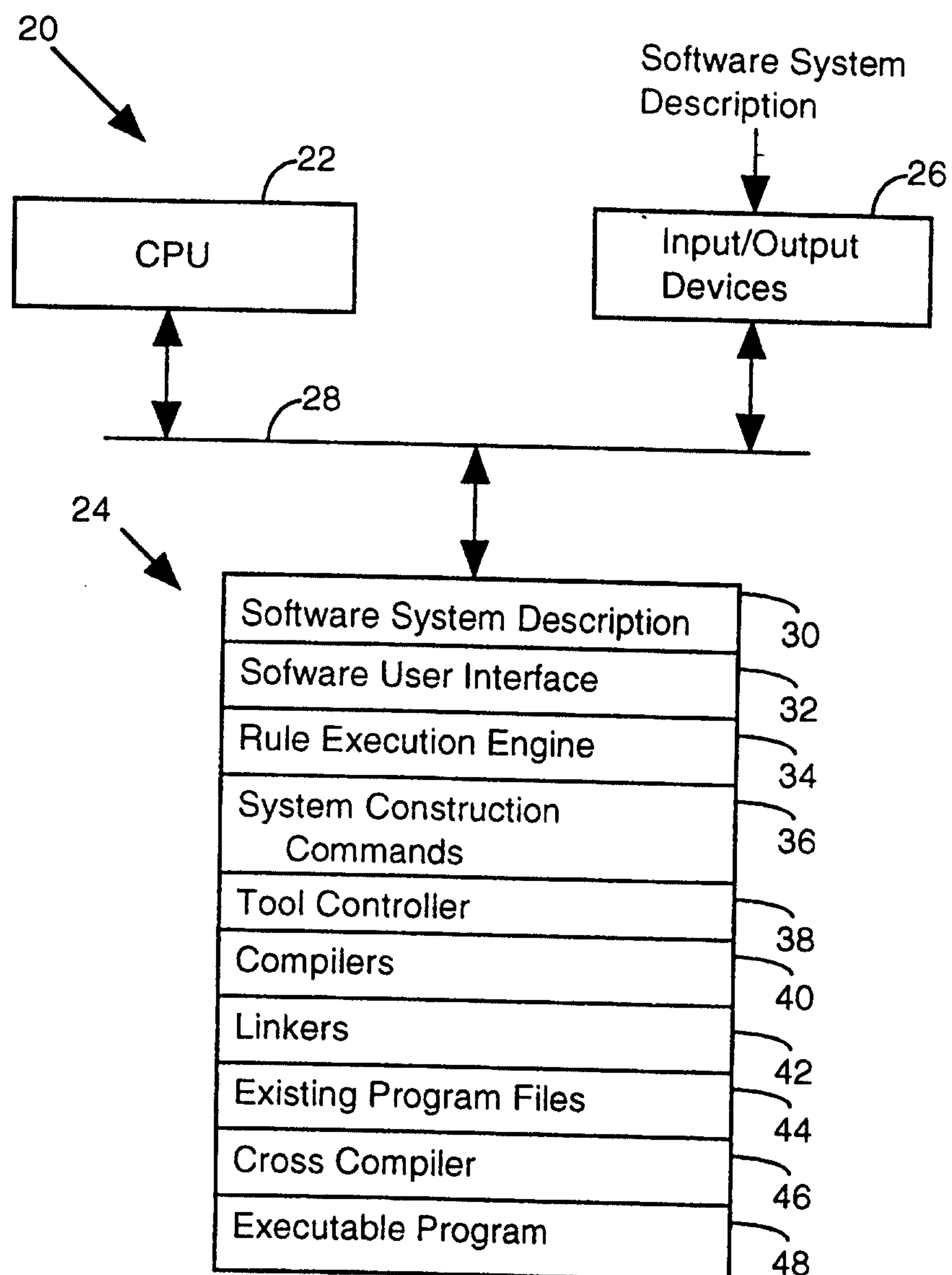


Fig. 1

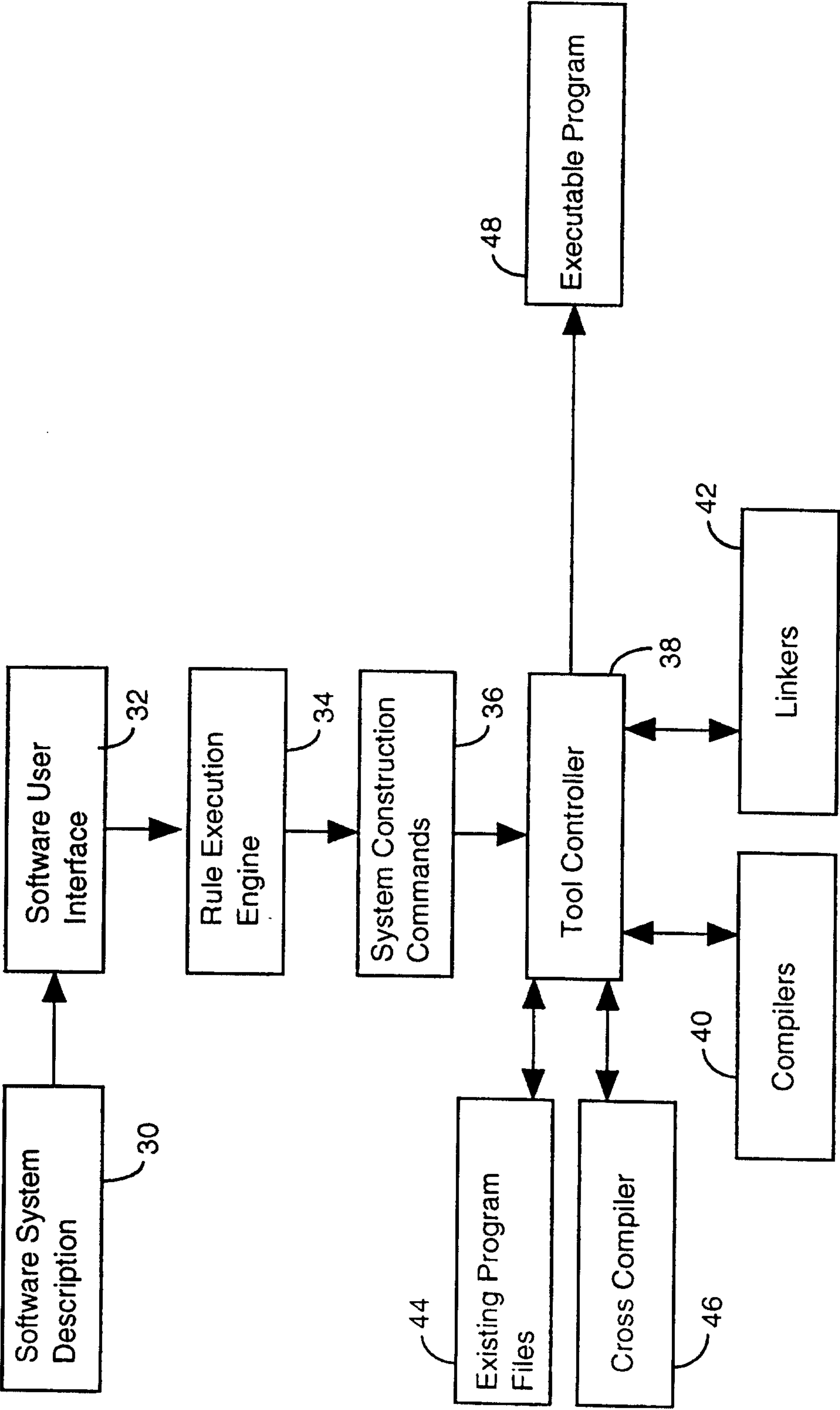


Fig. 2

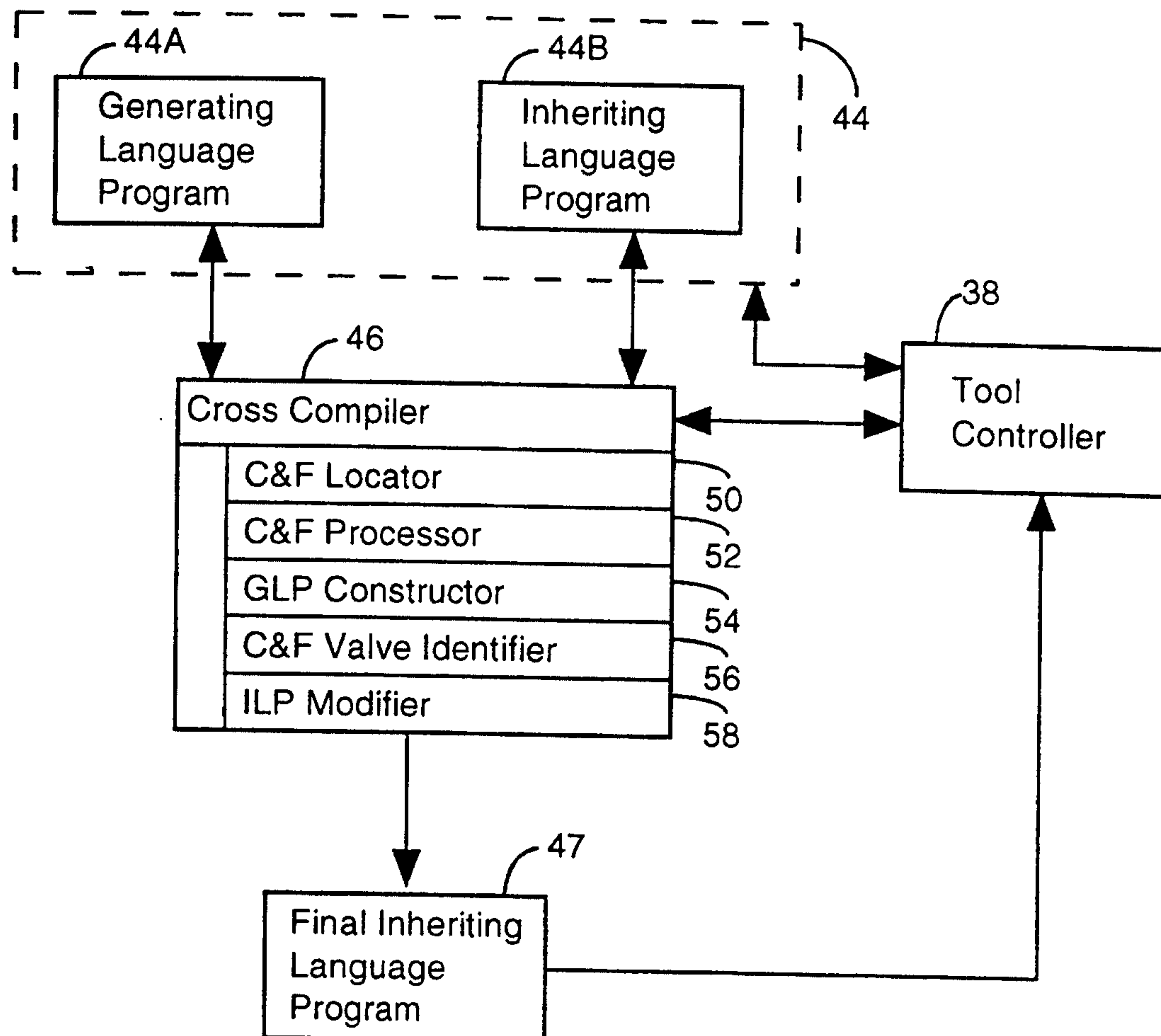


Fig. 3

Generating Language

Inheriting Language

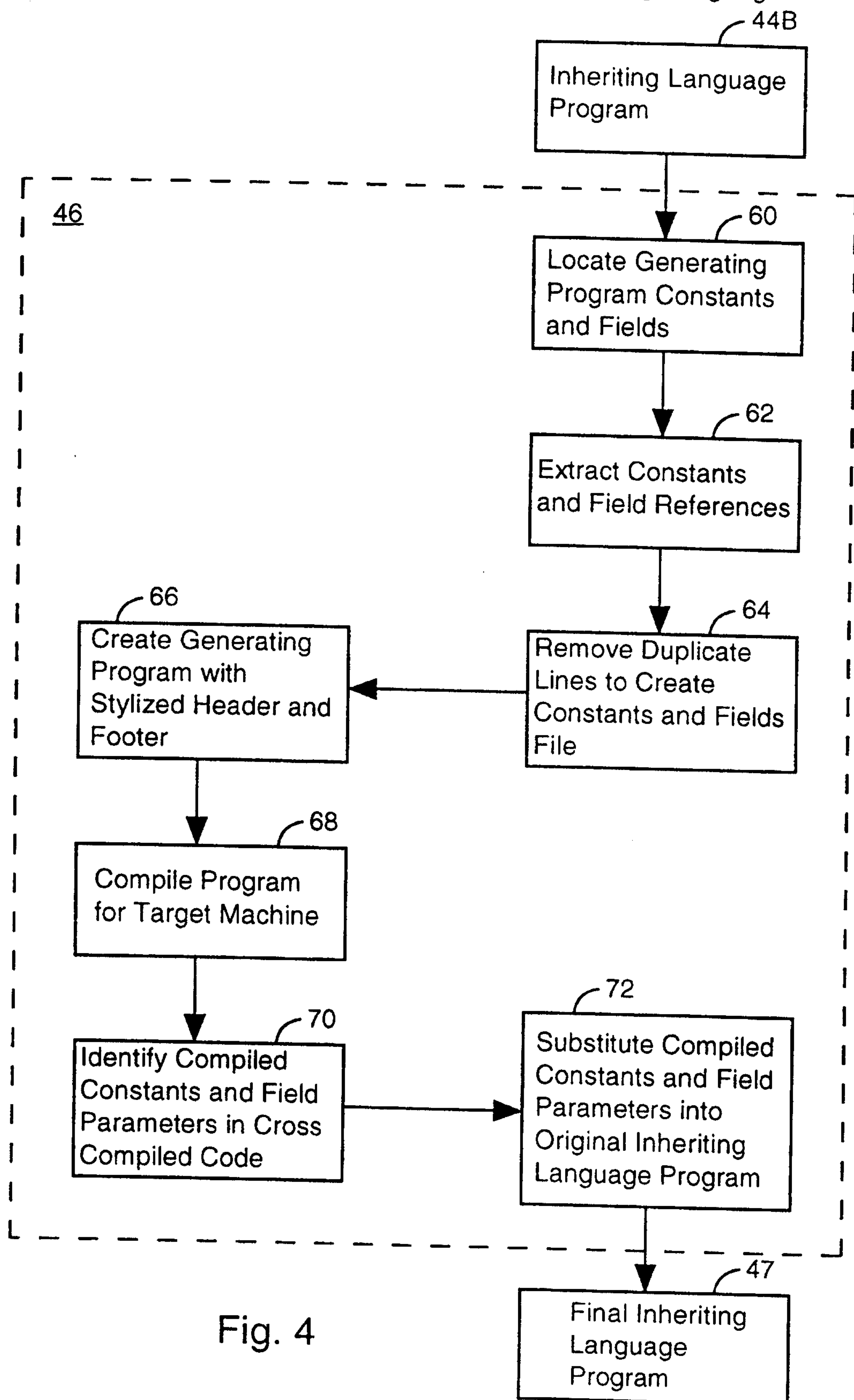


Fig. 4

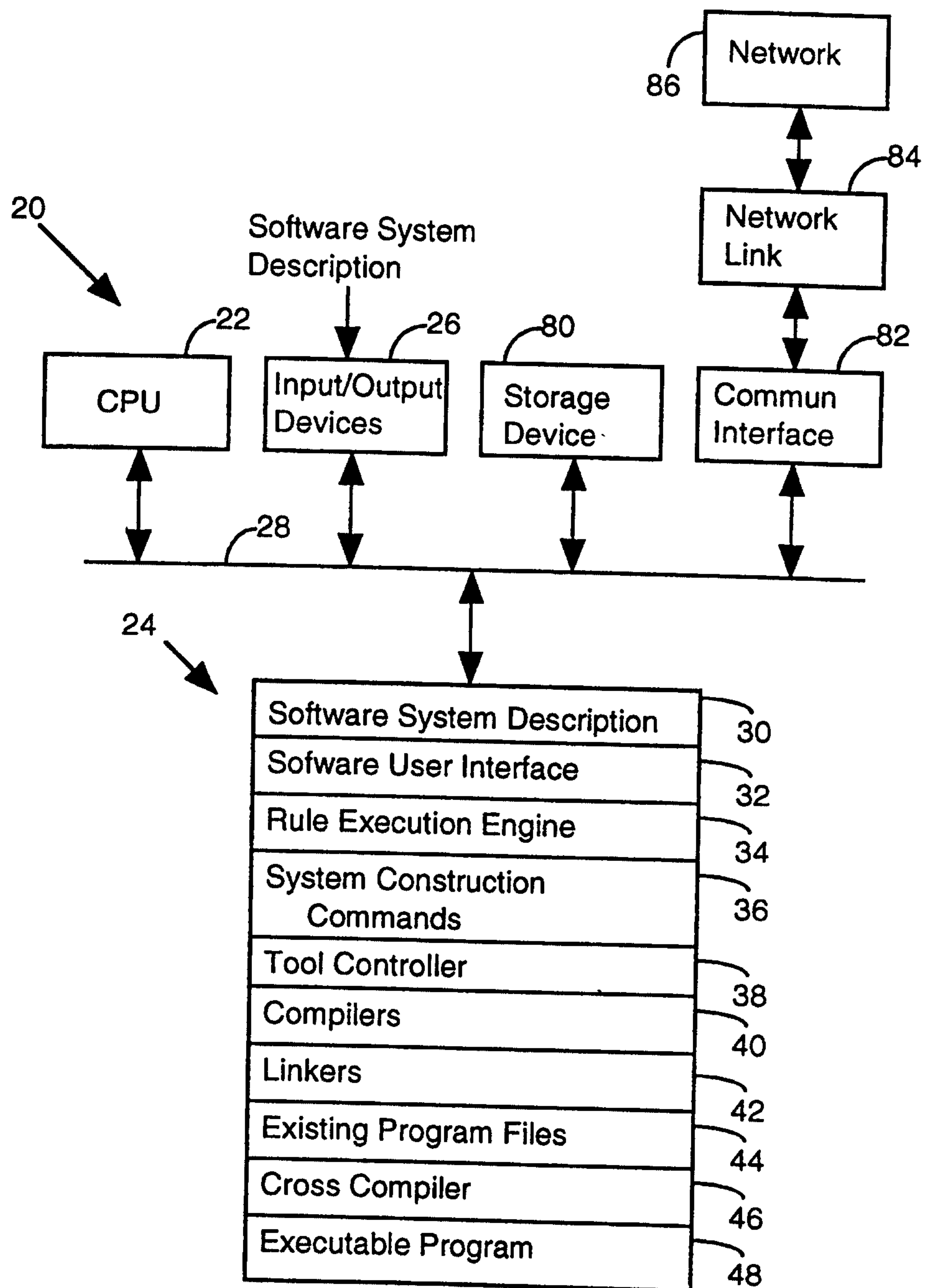


Fig. 5