Chapter 3 Loaders and Linkers
Outline

- 3.1 Basic Loader Functions
- 3.2 Machine-Dependent Loader Features
- 3.3 Machine-Independent Loader Features
- 3.4 Loader Design Options
- 3.5 Implementation Examples
Introduction

- **Loading**
  - Brings the object program into memory for execution

- **Relocation**
  - Modify the object program so that it can be loaded at an address different from the location originally specified

- **Linking**
  - Combine two or more separate object programs and supplies the information needed to allow references between them
Overview of Chapter 3

- Type of loaders
  - Assemble-and-go loader
  - Absolute loader (bootstrap loader)
  - Relocating loader (relative loader)
  - Direct linking loader

- Design options
  - Linkage editors
  - Dynamic linking
  - Bootstrap loaders
3.1 Basic Loader Functions

- The most fundamental functions of a loader:
  - Bringing an object program into memory and starting its execution

- Design of an Assemble-and-Go Loader

- Design of an Absolute Loader

- A Simple Bootstrap Loader
3.1.0 Assemble-and-Go Loader

- **Characteristic**
  - The object code is produced directly in memory for immediate execution after assembly

- **Advantage**
  - Useful for program development and testing

- **Disadvantage**
  - Whenever the assembly program is to be executed, it has to be assembled again
  - Programs consist of many control sections have to be coded in the same language
3.1.1 Design of an Absolute Loader

- **Absolute Program (e.g. SIC programs)**
  - **Advantage**
    - Simple and efficient
  - **Disadvantages**
    - The need for programmer to specify the actual address at which it will be loaded into memory
    - Difficult to use subroutine libraries efficiently

- **Absolute loader only performs** `loading` **function**
  - Does not need to perform `linking` and `program relocation`.
  - All functions are accomplished *in a single pass*. 
Design of an Absolute Loader (Cont.)

- In a single pass
  - Check the Header record for program name, starting address, and length
  - Bring the object program contained in the Text record to the indicated address
  - No need to perform program linking and relocation
  - Start the execution by jumping to the address specified in the End record
Loading of an Absolute Program (Fig 3.1 a)

- Object program contains
  - H record
  - T record
  - E record

(a) Object program

```plaintext
COPY 0010000107A
T0010001E1410334820390010362810303010154820613C100300102A0C1039001020
T00101E150C10364820610810334C0000454F4600000300000
T0020391E041030001030E0205D30203FD8205D2810303020575490392C205E38203F
T0020571C1010364C0000F1001000041030E02079302064509039DC20792C1036
T002073073820644C000005
E001000
```
Loading of an Absolute Program (Fig 3.1 b)

<table>
<thead>
<tr>
<th>Memory address</th>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx</td>
</tr>
<tr>
<td>0010</td>
<td>xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx</td>
</tr>
<tr>
<td>1000</td>
<td>14103348 20390010 36281030 30101548</td>
</tr>
<tr>
<td>1010</td>
<td>20613C10 0300102A 0C103900 102D0C10</td>
</tr>
<tr>
<td>1020</td>
<td>36482061 0810334C 0000454F 46000003</td>
</tr>
<tr>
<td>1030</td>
<td>00000000 xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx</td>
</tr>
<tr>
<td>2030</td>
<td>xxxxxxxxxx xxxxxxxxxx xx041030 001030E0</td>
</tr>
<tr>
<td>2040</td>
<td>205D0320 3FD8205D 28103030 20575490</td>
</tr>
<tr>
<td>2050</td>
<td>392C205E 38203F10 10364C00 00F10010</td>
</tr>
<tr>
<td>2060</td>
<td>00041030 E0207930 20645090 39DC2079</td>
</tr>
<tr>
<td>2070</td>
<td>2C103638 20644C00 0005xxxx xxxxxxxxxx</td>
</tr>
<tr>
<td>2080</td>
<td>xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx xxxxxxxxxx</td>
</tr>
</tbody>
</table>

(b) Program loaded in memory
Algorithm for an Absolute Loader (Fig. 3.2)

```
begin
  read Header record
  verify program name and length
  read first Text record
  while record type ≠ ‘E’ do
    begin
      move object code to specified location in memory
      read next object program record
    end
  end
  jump to address specified in End record
end
```

E.g., convert the pair of characters “14” (two bytes) in the object program to a single byte with hexadecimal value 14

**Figure 3.2** Algorithm for an absolute loader.
Object Code Representation

- Figure 3.1 (a)
  - Each byte of assembled code is given using its hexadecimal representation in *character* form
    - For example, 14 (opcode of STL) occupies two bytes of memory
    - Easy to read by human beings
  - Each pair of bytes from the object program record must be *packed together into one byte during loading.*
    - Inefficient in terms of both space and execution time

- Thus, most machine store object programs in a *binary form*
3.1.2 A Simple Bootstrap Loader

- Bootstrap Loader
  - When a computer is first turned on or restarted, a special type of absolute loader, called a bootstrap loader is executed
  - In PC, BIOS acts as a bootstrap loader

- This bootstrap loads the first program to be run by the computer -- usually an operating system
Example: a simple SIC/XE bootstrap loader (Fig. 3.3)

- The bootstrap itself begins at address 0 in the memory of the machine.
- It loads the OS (or some other program) starting address 0x80.
  - No Header record, End record, or control information.
  - The object code from device F1 is always loaded into consecutive bytes of memory, starting at address 80.
- After all the object code from device F1 has been loaded, the bootstraps jumps to address 80, which begins the execution of the program that was loaded.
Bootstrap loader for SIC/XE (Fig. 3.3)

BOOT  START  0  BOOTSTRAP LOADER FOR SIC/XE

. This bootstrap reads object code from device F1 and enters it into memory starting at address 80 (hexadecimal). After all of the code from DEVF1 has been seen entered into memory, the bootstrap executes a jump to address 80 to begin execution of the program just loaded. Register X contains the next address to be loaded.

CLEAR  A  CLEAR REGISTER A TO ZERO
LDX  #128  INITIALIZE REGISTER X TO HEX 80
LOOP  JSUB  GETC  READ HEX DIGIT FROM PROGRAM BEING LOADED
      RMO  A,S  SAVE IN REGISTER S
      SHFTL  S,4  MOVE TO HIGH-ORDER 4 BITS OF BYTE
      JSUB  GETC  GET NEXT HEX DIGIT
      ADDR  S,A  COMBINE DIGITS TO FORM ONE BYTE
      STCH  0,X  STORE AT ADDRESS IN REGISTER X
      TIXR  X,X  ADD 1 TO MEMORY ADDRESS BEING LOADED
      J  LOOP  LOOP UNTIL END OF INPUT IS REACHED
Bootstrap loader for SIC/XE (Fig. 3.3)

- SUBROUTINE TO READ ONE CHARACTER FROM INPUT DEVICE AND
- CONVERT IT FROM ASCII CODE TO HEXADECIMAL DIGIT VALUE. THE
- CONVERTED DIGIT VALUE IS RETURNED IN REGISTER A. WHEN AN
- END-OF-FILE IS READ, CONTROL IS TRANSFERRED TO THE STARTING
- ADDRESS (HEX 80).

GETC   TD   INPUT    TEST INPUT DEVICE
JEQ    GETC  LOOP UNTIL READY
RD     INPUT  READ CHARACTER
COMP   #4   IF CHARACTER IS HEX 04 (END OF FILE),
JEQ    80   JUMP TO START OF PROGRAM JUST LOADED
COMP   #48  COMPARE TO HEX 30 (CHARACTER '0')
JLT    GETC  SKIP CHARACTERS LESS THAN '0'
SUB    #48  SUBTRACT HEX 30 FROM ASCII CODE
COMP   #10  IF RESULT IS LESS THAN 10, CONVERSION IS
JLT    RETURN  COMPLETE. OTHERWISE, SUBTRACT 7 MORE
SUB    #7   (FOR HEX DIGITS 'A' THROUGH 'F')
RETURN  RSUB  RETURN TO CALLER
INPUT   BYTE 'X'F1'  CODE FOR INPUT DEVICE
END     LOOP

Figure 3.3 Bootstrap loader for SIC/XE.
Bootstrap loader for SIC/XE (Fig. 3.3)

begin
  X=0x80 ; the address of the next memory location to be loaded
Loop
  A←GETC ; read one char. From device F1 and convert it from the
           ; ASCII character code to the value of the hex digit
  save the value in the high-order 4 bits of S
  A←GETC
  A←(A+S) ; combine the value to form one byte
store the value (in A) to the address represented in register X
  X←X+1
end
3.2 Machine-Dependent Loader

Features

- **Drawback of absolute loaders**
  - Programmer needs to specify the actual address at which it will be loaded into memory.
  - Difficult to run several programs concurrently, sharing memory between them.
  - Difficult to use subroutine libraries.

- **Solution: a more complex loader that provides**
  - *Program relocation*
  - *Program linking*
Machine-Dependent Loader Features (Cont.)

- 3.2.1 Relocation
- 3.2.2 Program Linking
- 3.2.3 Algorithm and Data Structures for a Linking Loader
Review

Section 2.2.2
Program Relocation
Program Relocation

- **Relocatable program**
  
  An object program that contains the information necessary to perform address modification for relocation.

  The assembler can identify for the loader those parts of object program that need modification.

  No instruction modification is needed for
  - immediate addressing (not a memory address)
  - PC-relative, Base-relative addressing

  The only parts of the program that require modification at load time are those that specify **direct addresses**.
Instruction Format vs. Relocatable Loader

- In SIC/XE
  - *Relative* and *immediate* addressing
    - Do not need to modify their object code after relocation
  - *Extended format*
    - Whose values are affected by relocation
    - Need to modify when relocation

- In SIC
  - Format 3 with address field
    - Should be modified
    - SIC does not support PC-relative and base-relative addressing
3.2.1 Relocation

- Loaders that allow for program relocation are called **relocating loaders** or **relative loaders**.

- Two methods for specifying relocation as part of the object program
  
  - **Modification records**
    - Suitable for a **small** number of relocations required when relative or immediate addressing modes are extensively used
  
  - **Relocation bits**
    - Suitable for a **large** number of relocations required when only direct addressing mode can be used in a machine with fixed instruction format (e.g., the standard SIC machine)
Relocation by Modification Record

- A *Modification record* is used to describe each part of the object code that must be changed when the program is relocated.

- Fig 3.4 & 3.5
  - The only portions of the assembled program that contain addresses are the *extended format* instructions on lines 15,35,65
  - The only items whose values are affected by relocation.
Example of a SIC/XE Program (Fig 3.4, 2.6)

<table>
<thead>
<tr>
<th>Line</th>
<th>Loc</th>
<th>Source statement</th>
<th>Object code</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0000</td>
<td>COPY START 0</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0000</td>
<td>FIRST STL RETADR</td>
<td>17202D</td>
</tr>
<tr>
<td>12</td>
<td>0003</td>
<td>LDB #LENGTH</td>
<td>69202D</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>BASE LENGTH</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0006</td>
<td>CLOOP +JSUB RDREC</td>
<td>4E101036</td>
</tr>
<tr>
<td>20</td>
<td>000A</td>
<td>LDA LENGTH</td>
<td>032026</td>
</tr>
<tr>
<td>25</td>
<td>000D</td>
<td>COMP #0</td>
<td>290000</td>
</tr>
<tr>
<td>30</td>
<td>0010</td>
<td>JEQ ENDFIL</td>
<td>332007</td>
</tr>
<tr>
<td>35</td>
<td>0013</td>
<td>+JSUB WRREC</td>
<td>4B10105D</td>
</tr>
<tr>
<td>40</td>
<td>0017</td>
<td>J CLOOP</td>
<td>3F2FEC</td>
</tr>
<tr>
<td>45</td>
<td>001A</td>
<td>ENDFIL LDA EOF</td>
<td>032010</td>
</tr>
<tr>
<td>50</td>
<td>001D</td>
<td>STA BUFFER</td>
<td>0F2016</td>
</tr>
<tr>
<td>55</td>
<td>0020</td>
<td>LDA #3</td>
<td>010003</td>
</tr>
<tr>
<td>60</td>
<td>0023</td>
<td>STA LENGTH</td>
<td>0F200D</td>
</tr>
<tr>
<td>65</td>
<td>0026</td>
<td>+JSUB WRREC</td>
<td>4B10105D</td>
</tr>
<tr>
<td>70</td>
<td>002A</td>
<td>J @RETADR</td>
<td>3F2003</td>
</tr>
<tr>
<td>80</td>
<td>002D</td>
<td>BOF BYTE C’EOF’</td>
<td>454F46</td>
</tr>
<tr>
<td>95</td>
<td>0030</td>
<td>RETADR RESW 1</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0033</td>
<td>LENGTH RESW 1</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>0036</td>
<td>BUFFER RESB 4096</td>
<td></td>
</tr>
</tbody>
</table>

Only three addresses need to be relocated.
Example of a SIC/XE Program (Fig 3.4, 2.6) (Cont.)

```
110          .                      SUBROUTINE TO READ RECORD INTO BUFFER
115          .
120          .
125   1036   RDREC   CLEAR   X   B410
130   1038   CLEAR   A   B400
132   103A   CLEAR   S   B440
133   103C   +LDT   #4096   75101000
135   1040   RLOOP   TD   INPUT   E32019
140   1043   JEQ   RLOOP   332FFA
145   1046   RD   INPUT   DB2013
150   1049   COMP   A,S   A004
155   104B   JEQ   EXIT   332008
160   104E   STCH   BUFFER,X   57C003
165   1051   TIXR   T   B850
170   1053   JLT   RLOOP   3B2FEA
175   1056   EXIT   STX   LENGTH   134000
180   1059   RSUB   X'1000'   4F0000
185   105C   INPUT   BYTE   X'F1'   F1
```
Example of a SIC/XE Program (Fig 3.4,2.6) (Cont.)

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

```

**Figure 2.6** Program from Fig. 2.5 with object code.```
Relocatable Program

Pass the address – modification information to the relocatable loader

- **Modification record**
  - Col 1 M
  - Col 2-7 Starting location of the address field to be modified, relative to the beginning of the program (hex)
  - Col 8-9 length of the address field to be modified, in half-bytes
  - E.g M\^000007\^05

Beginning address of the program is to be added to a field that begins at addr ox000007 and is 5 bytes in length.
Object Program with Relocation by Modification Records for Fig 3.5 (Fig 2.8)

Chapter 3  Loaders and Linkers

Add the starting address of the program

There is one modification record for each address need to be relocated.

Figure 3.5 Object program with relocation by Modification records.
The Modification record scheme is a convenient means for specifying program relocation.

However, it is not well suited for use with all machine architectures

- See Fig. 3.6.
  - Relocatable program for a SIC machine
- Most instructions use direct addressing
  - *Too many modification records*
Relocatable program for a standard SIC machine (Fig. 3.6)

<table>
<thead>
<tr>
<th>Line</th>
<th>Loc</th>
<th>Source statement</th>
<th>Object code</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0000</td>
<td>COPY START 0</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0000</td>
<td>FIRST STL RETADR</td>
<td>140033</td>
</tr>
<tr>
<td>15</td>
<td>0003</td>
<td>CLOOP JSUB RDREC</td>
<td>481039</td>
</tr>
<tr>
<td>20</td>
<td>0006</td>
<td>LDA LENGTH</td>
<td>000036</td>
</tr>
<tr>
<td>25</td>
<td>0009</td>
<td>COMP ZERO</td>
<td>280030</td>
</tr>
<tr>
<td>30</td>
<td>000C</td>
<td>JEQ ENDFIL</td>
<td>300015</td>
</tr>
<tr>
<td>35</td>
<td>000F</td>
<td>JSUB WRREC</td>
<td>481061</td>
</tr>
<tr>
<td>40</td>
<td>0012</td>
<td>J CLOOP</td>
<td>3C0003</td>
</tr>
<tr>
<td>45</td>
<td>0015</td>
<td>ENDFIL LDA EOF</td>
<td>00002A</td>
</tr>
<tr>
<td>50</td>
<td>0018</td>
<td>STA BUFFER</td>
<td>0C0039</td>
</tr>
<tr>
<td>55</td>
<td>001B</td>
<td>LDA THREE</td>
<td>00002D</td>
</tr>
<tr>
<td>60</td>
<td>001E</td>
<td>STA LENGTH</td>
<td>0C0036</td>
</tr>
<tr>
<td>65</td>
<td>0021</td>
<td>JSUB WRREC</td>
<td>481061</td>
</tr>
<tr>
<td>70</td>
<td>0024</td>
<td>LDL RETADR</td>
<td>080033</td>
</tr>
<tr>
<td>75</td>
<td>0027</td>
<td>RSUB</td>
<td>4C0000</td>
</tr>
<tr>
<td>80</td>
<td>002A</td>
<td>EOF BYTE C’EOF’</td>
<td>454F46</td>
</tr>
<tr>
<td>85</td>
<td>002D</td>
<td>THREE WORD 3</td>
<td>000003</td>
</tr>
<tr>
<td>90</td>
<td>0030</td>
<td>ZERO WORD 0</td>
<td>000000</td>
</tr>
<tr>
<td>95</td>
<td>0033</td>
<td>RETADR RESW 1</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>0036</td>
<td>LENGTH RESW 1</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>0039</td>
<td>BUFFER RESB 4096</td>
<td></td>
</tr>
<tr>
<td>110</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td></td>
<td></td>
<td>SUBROUTINE TO READ RECORD INTO BUFFER</td>
</tr>
</tbody>
</table>
Relocatable program for a standard SIC machine (Fig. 3.6) (Cont.)

<table>
<thead>
<tr>
<th>Line</th>
<th>Address</th>
<th>Instruction</th>
<th>Label</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>1039</td>
<td>RDREC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>103C</td>
<td>LDX ZERO</td>
<td></td>
<td>040030</td>
</tr>
<tr>
<td>135</td>
<td>103F</td>
<td>RLOOP TD</td>
<td></td>
<td>000030</td>
</tr>
<tr>
<td>140</td>
<td>1042</td>
<td>JEQ RLOOP</td>
<td></td>
<td>E0105D</td>
</tr>
<tr>
<td>145</td>
<td>1045</td>
<td>RD INPUT</td>
<td></td>
<td>30103F</td>
</tr>
<tr>
<td>150</td>
<td>1048</td>
<td>COMP ZERO</td>
<td></td>
<td>D8105D</td>
</tr>
<tr>
<td>155</td>
<td>104B</td>
<td>JEQ EXIT</td>
<td></td>
<td>280030</td>
</tr>
<tr>
<td>160</td>
<td>104E</td>
<td>STCH BUFFER,X</td>
<td></td>
<td>301057</td>
</tr>
<tr>
<td>165</td>
<td>1051</td>
<td>TIX MAXLEN</td>
<td></td>
<td>548039</td>
</tr>
<tr>
<td>170</td>
<td>1054</td>
<td>JLT RLOOP</td>
<td></td>
<td>2C105E</td>
</tr>
<tr>
<td>175</td>
<td>1057</td>
<td>EXIT STX</td>
<td>LENGTH</td>
<td>38103F</td>
</tr>
<tr>
<td>180</td>
<td>105A</td>
<td>RSUB</td>
<td></td>
<td>100036</td>
</tr>
<tr>
<td>185</td>
<td>105D</td>
<td>INPUT BYTE</td>
<td>X’F1’</td>
<td>4C0000</td>
</tr>
<tr>
<td>190</td>
<td>105E</td>
<td>MAXLEN WORD</td>
<td>4096</td>
<td>F1</td>
</tr>
<tr>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td>001000</td>
</tr>
<tr>
<td>200</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td></td>
<td></td>
<td>SUBROUTINE TO WRITE RECORD FROM BUFFER</td>
</tr>
</tbody>
</table>
Relocatable program for a Standard SIC Machine (fig. 3.6) (Cont.)

<table>
<thead>
<tr>
<th>Line</th>
<th>Address</th>
<th>Instruction</th>
<th>Operand</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>205</td>
<td>1061</td>
<td>WRREC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>1064</td>
<td>WLOOP</td>
<td>LDX</td>
<td>ZERO</td>
</tr>
<tr>
<td>215</td>
<td>1067</td>
<td>JEQ</td>
<td>TD</td>
<td>OUTPUT</td>
</tr>
<tr>
<td>220</td>
<td>106A</td>
<td>LDCH</td>
<td>JEQ</td>
<td>WLOOP</td>
</tr>
<tr>
<td>225</td>
<td>106D</td>
<td></td>
<td>LDCH</td>
<td>BUFFER, X</td>
</tr>
<tr>
<td>230</td>
<td>1070</td>
<td></td>
<td>WD</td>
<td>OUTPUT</td>
</tr>
<tr>
<td>235</td>
<td>1073</td>
<td></td>
<td>TIX</td>
<td>LENGTH</td>
</tr>
<tr>
<td>240</td>
<td>1076</td>
<td></td>
<td>JLTL</td>
<td>LOOP</td>
</tr>
<tr>
<td>245</td>
<td>1079</td>
<td>OUTPUT</td>
<td>RSUB</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td></td>
<td>OUTPUT</td>
<td>BYTE</td>
<td>X’05’</td>
</tr>
<tr>
<td>255</td>
<td></td>
<td>END</td>
<td>END</td>
<td>FIRST</td>
</tr>
</tbody>
</table>

**Figure 3.6**  Relocatable program for a standard SIC machine.

This SIC program does not use relative addressing. The addresses in all the instructions except RSUB must be modified. This would require 31 Modification records.
Relocation by Modification Bit

- If a machine primarily uses *direct addressing* and has a *fixed instruction format*
  - There are many addresses needed to be modified
  - It is often more efficient to specify relocation using *relocation bit*

- *Relocation bit* (Fig. 3.6, 3.7)
  - Each instruction is associated with *one relocation bit*
    - Indicate the corresponding word should be modified or not.
  - These relocation bits in a Text record is gathered into *bit masks*
Relocation by Modification Bit (Fig. 3.7)

- Relocation bit
  - 0: no modification is needed
  - 1: modification is needed

Text record
- col 1: T
- col 2-7: starting address
- col 8-9: length (byte)
- col 10-12: relocation bits
- col 13-72: object code

Figure 3.7  Object program with relocation by bit mask.

F1 is one-byte
Relocation Bits (Cont.)

- Each bit mask consists of 12 relocation bit in each Text record
  - Since each text record contains less than 12 words
  - Unused words are set to 0
    - E.g. FFC=111111111100 for line 10-55
    - However, only 10 words in the first text record
Relocation Bits (Cont.)

- Note that, any value that is to be modified during relocation must coincide with one of these 3-byte segments
  - E.g. Begin a new Text record for line 210
    - Because line 185 has only 1-byte object code (F1)
    - Make the following object code does not align to 3-byte boundary
Some computers provide a hardware relocation capability that eliminates some of the need for the loader to perform program relocation.

- E.g. some such machine consider all memory references to be relative to the beginning of the user’s assigned area of memory.
- The conversion of these relative addresses to actual addresses is performed as the program is executed.
Mapped Memory (Skip)

Boundary value

< → Enable/Error

Logical address

+ → Physical address

Base address

Hardware memory mapping
Base & Bound (Skip)
3.2.2 Program Linking

- Control sections
  - Refer to segments of codes that are translated into *independent* object program units
  - These control sections could be assembled together or independently of one another
  - It is necessary to provide some means for *linking control sections together*
    - External definitions
    - External references
Illustration of Control Sections and Program Linking (Fig 2.15 & 2.16)

<table>
<thead>
<tr>
<th>Line</th>
<th>Loc</th>
<th>Source statement</th>
<th>Object code</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0000</td>
<td>COPY START</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0000</td>
<td>EXTDDEF BUFFER, BUEND, LENGTH</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0000</td>
<td>EXTRREF RDREC, WRREC</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0000</td>
<td>FIRST STG RETADR</td>
<td>1F2027</td>
</tr>
<tr>
<td>15</td>
<td>0003</td>
<td>CLOOP +JSUB RDREC</td>
<td>4B100000</td>
</tr>
<tr>
<td>20</td>
<td>0007</td>
<td>LDA LENGTH</td>
<td>032023</td>
</tr>
<tr>
<td>25</td>
<td>000A</td>
<td>COMP #0</td>
<td>290000</td>
</tr>
<tr>
<td>30</td>
<td>000D</td>
<td>IEQ ENDFIL</td>
<td>332007</td>
</tr>
<tr>
<td>35</td>
<td>0010</td>
<td>+JSUB WRREC</td>
<td>4B100000</td>
</tr>
<tr>
<td>40</td>
<td>0014</td>
<td>J CLOOP</td>
<td>3F2FEC</td>
</tr>
<tr>
<td>45</td>
<td>0017</td>
<td>ENDFIL LDA =C'EDF'</td>
<td>032016</td>
</tr>
<tr>
<td>50</td>
<td>001A</td>
<td>STA BUFFER</td>
<td>0F2016</td>
</tr>
<tr>
<td>55</td>
<td>001D</td>
<td>LDA #3</td>
<td>010003</td>
</tr>
<tr>
<td>60</td>
<td>0020</td>
<td>STA LENGTH</td>
<td>0F2007</td>
</tr>
<tr>
<td>65</td>
<td>0023</td>
<td>+JSUB WRREC</td>
<td>4B100000</td>
</tr>
<tr>
<td>70</td>
<td>0027</td>
<td>J RETADR</td>
<td>3E2000</td>
</tr>
<tr>
<td>95</td>
<td>002A</td>
<td>RETADR RESW</td>
<td>1</td>
</tr>
<tr>
<td>100</td>
<td>002D</td>
<td>LENGTH RESW</td>
<td>1</td>
</tr>
<tr>
<td>103</td>
<td>0030</td>
<td>LTORG =C'EDF'</td>
<td>454F46</td>
</tr>
<tr>
<td>105</td>
<td>0033</td>
<td>BUFFER RESB</td>
<td>4096</td>
</tr>
<tr>
<td>106</td>
<td>1033</td>
<td>BUFOEND RQU</td>
<td>*</td>
</tr>
<tr>
<td>107</td>
<td>1000</td>
<td>MAXLEN RQU</td>
<td>BUFEND-BUFFER</td>
</tr>
</tbody>
</table>
Illustration of Control Sections and Program Linking (Fig 2.15 & 2.16) (Cont.)

109  0000  RDREC  CSECT
110
115
116  SUBROUTINE TO READ RECORD INTO BUFFER
120
121
122  EXTREF  BUFFER, LENGTH, BUFEND
125  0000  CLEAR  X    B410
130  0002  CLEAR  A    B400
132  0004  CLEAR  S    3440
133  0006  LDT  MAXLEN  77201F
135  0009  RLOOP  TD  INPUT  E3201B
140  000C  JEQ  RLOOP  332FFA
145  000F  RD  INPUT  DB2015
150  0012  COMPR  A, S  A004
155  0014  JEQ  EXIT  332009
160  0017  +STCH  BUFFER, X  5790000C
165  001B  TIXR  T  B850
170  001D  JLT  RLOOP  3E2FES
175  0020  EXIT  +STX  LENGTH  1310000C
180  0024  RSUB  4F0000
185  0027  INPUT  BYTE  X,'$1'  F1
190  0028  MAXLEN  WORD  BUFEND-BUFFER  000000
Illustration of Control Sections and Program Linking (Fig 2.15 & 2.16) (Cont.)

**Figure 2.16** Program from Fig. 2.15 with object code.
Control Sections and Program Linking (Cont.)

- **Assembler directive: secname CSECT**
  - Signals the start of a new control section
  - E.g. 109 RDREC CSECT
  - e.g. 193 WRREC CSECT

- **External references**
  - References between control sections
  - The assembler generates information for each external reference that will allow the loader to perform the required linking.
How the Assembler Handles Control Sections?

- The assembler must include information in the object program that will cause the loader to insert proper values where they are required

Define record

- Col. 1  D
- Col. 2-7 Name of external symbol defined in this control section
- Col. 8-13 Relative address within this control section (hex)
- Col. 14-73 Repeat information in Col. 2-13 for other external symbols

Refer record

- Col. 1  R
- Col. 2-7 Name of external symbol referred to in this control section
- Col. 8-73 Name of other external reference symbols
How the Assembler Handles Control Sections? (Cont.)

- **Modification record** *(revised)*
  - Col. 1 M
  - Col. 2-7 Starting address of the field to be modified (hex)
  - Col. 8-9 Length of the field to be modified, in half-bytes (hex)
  - Col. 10 Modification flag (+ or - )
  - Col. 11-16 External symbol whose value is to be added to or subtracted from the indicated field.

- **Example (Figure 2.17)**
  - M000004_05_+RDREC
  - M000011_05_+WRREC
  - M000024_05_+WRREC
  - M000028_06_+BUFEND
  - M000028_06_-BUFFER
Program Linking (Cont.)

- **Goal of program linking**
  - Resolve the problems with EXTREF and EXTDEF from different control sections

- **Example:**
  - Fig. 3.8 and Fig. 3.9
Sample Programs Illustrating Linking and Relocation (Fig. 3.8) – PROGA

<table>
<thead>
<tr>
<th>Loc</th>
<th>Source statement</th>
<th>Object code</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>PROGA</td>
<td></td>
</tr>
<tr>
<td>0020</td>
<td>REF1 LDA LISTA</td>
<td>03201D</td>
</tr>
<tr>
<td>0023</td>
<td>REF2 +LDT LISTB+4</td>
<td>77100004</td>
</tr>
<tr>
<td>0027</td>
<td>REF3 LDX #ENDA-LISTA</td>
<td>050014</td>
</tr>
<tr>
<td>0040</td>
<td>LISTA EQU *</td>
<td></td>
</tr>
<tr>
<td>0054</td>
<td>ENDA EQU *</td>
<td>000014</td>
</tr>
<tr>
<td>0054</td>
<td>REF4 WORD ENDA-LISTAC+LISTC</td>
<td>FFFFF6</td>
</tr>
<tr>
<td>0055</td>
<td>REF5 WORD ENDC-LISTC-10</td>
<td></td>
</tr>
<tr>
<td>005A</td>
<td>REF6 WORD ENDC-LISTC+LISTA-1</td>
<td>00003F</td>
</tr>
<tr>
<td>005D</td>
<td>REF7 WORD ENDA-LISTA-(ENDB-LISTB)</td>
<td>000014</td>
</tr>
<tr>
<td>0060</td>
<td>REF8 WORD LISTB-LISTA</td>
<td>FFFFFFFC0</td>
</tr>
<tr>
<td></td>
<td>END REF1</td>
<td></td>
</tr>
</tbody>
</table>
Sample Programs Illustrating Linking and Relocation (Fig. 3.8) – PROGB

<table>
<thead>
<tr>
<th>Loc</th>
<th>Source statement</th>
<th>Object code</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>PROGB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>START 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EXTDFF LISTB, ENDF</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EXTRREF LISTA, ENDA, LISTC, ENDC</td>
<td></td>
</tr>
<tr>
<td>0036</td>
<td>REF1 +LDA LISTA</td>
<td>03100000</td>
</tr>
<tr>
<td>003A</td>
<td>REF2 LIST LISTB+4</td>
<td>772027</td>
</tr>
<tr>
<td>003D</td>
<td>REF3 +LDX #ENDA-LISTA</td>
<td>05100000</td>
</tr>
<tr>
<td>0060</td>
<td>LISTB EQU *</td>
<td></td>
</tr>
<tr>
<td>0070</td>
<td>ENDB EQU *</td>
<td></td>
</tr>
<tr>
<td>0070</td>
<td>REF4 WORD ENDA-LISTA+LISTC</td>
<td>000000</td>
</tr>
<tr>
<td>0073</td>
<td>REF5 WORD ENDC-LISTC-10</td>
<td>FFFFF6</td>
</tr>
<tr>
<td>0076</td>
<td>REF6 WORD ENDC-LISTC+LISTA-1</td>
<td>FFFFFF</td>
</tr>
<tr>
<td>0079</td>
<td>REF7 WORD ENDA-LISTA- (ENDB-LISTB)</td>
<td>FFFFFF0</td>
</tr>
<tr>
<td>007C</td>
<td>REF8 WORD LISTB-LISTA</td>
<td>000060</td>
</tr>
</tbody>
</table>

Figure 3.8 Sample programs illustrating linking and relocation.
### Sample Programs Illustrating Linking and Relocation (Fig. 3.8) – PROGC

<table>
<thead>
<tr>
<th>Loc</th>
<th>Source statement</th>
<th>Object code</th>
</tr>
</thead>
<tbody>
<tr>
<td>0000</td>
<td>PROGC START 0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EXTDDEF LISTC ENDC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EXTRREF LISTA ENDA LISTB ENDS</td>
<td></td>
</tr>
<tr>
<td>0018</td>
<td>REF1 +LDA LISTA</td>
<td>03100000</td>
</tr>
<tr>
<td>001C</td>
<td>REF2 +LDT LISTB +4</td>
<td>77100004</td>
</tr>
<tr>
<td>0020</td>
<td>REF3 +LDX #ENDA-LISTA</td>
<td>05100000</td>
</tr>
<tr>
<td>0030</td>
<td>LISTC EQU *</td>
<td></td>
</tr>
<tr>
<td>0042</td>
<td>ENDC EQU *</td>
<td></td>
</tr>
<tr>
<td>0042</td>
<td>REF4 WORD ENDA-LISTA+LISTC</td>
<td>000030</td>
</tr>
<tr>
<td>0045</td>
<td>REF5 WORD ENDC-LISTC-10</td>
<td>000008</td>
</tr>
<tr>
<td>0048</td>
<td>REF6 WORD ENDC-LISTC+LISTA-1</td>
<td>000011</td>
</tr>
<tr>
<td>004B</td>
<td>REF7 WORD ENDA-LISTA-(ENDB-LISTB)</td>
<td>000000</td>
</tr>
<tr>
<td>004E</td>
<td>REF8 WORD LISTB-LISTA</td>
<td>000000</td>
</tr>
</tbody>
</table>

*Figure 3.8 (cont’d)*
Sample Programs Illustrating Linking and Relocation

- Each control section defines a list:
  - Control section A: LISTA --- ENDA
  - Control section B: LISTB --- ENDB
  - Control section C: LISTC --- ENDC

- Each control section contains exactly the same set of references to these lists
  - REF1 through REF3: instruction operands
  - REF4 through REF8: values of data words
Sample Programs Illustrating Linking and Relocation (Fig. 3.9) – PROGA

Figure 3.9  Object programs corresponding to Fig. 3.8.
Sample Programs Illustrating Linking and Relocation (Fig. 3.9) – PROGB

```
HPROGB 00000000007F
DLISTB 000060ENDB 000070
KLISTA ENDA LISTC ENDC

T0000360B0310000077202705100000

T0000700F00000000FFFFF6FFFFFFFFFF000060

MO0003705+LISTA
MO0003E05+ENDA
MO0003E05-LISTA
MO0007006+ENDA
MO0007006-LISTA
MO0007006+LISTC
MO0007306+ENDC
MO0007306-LISTC
MO0007606+ENDC
MO0007606-LISTC
MO0007606+LISTA
MO0007906+ENDA
MO0007906-LISTA
MO0007C06+PROGB
MO0007C06-LISTA
```
Sample Programs Illustrating Linking and Relocation (Fig. 3.9) – PROGC

```
HROGIC 00000000000051
DLISTC 000030ENDC 000042
RLISTA ENDA LISTB ENDB

*
T0000180C03100007710000405100000
*
T0000420E000030000080000110000000000
M00001905+LISTA
M00001D05+LISTB
M00002105+ENDA
M00002105-LISTA
M00004206+ENDA
M00004206-LISTA
M00004206+PROGC
M00004806+LISTA
M00004B06+ENDA
M00004B06-LISTA
M00004B06-ENDB
M00004B06+LISTB
M00004E06+LISTB
M00004E06-LISTA
E
```

Figure 3.9  (cont’d)
REF1 (LISTA)

- Control section A
  - LISTA is defined within the control section.
  - Its address is available using **PC-relative addressing**.
  - No modification for relocation or linking is necessary.

- Control sections B and C
  - LISTA is an *external reference*.
  - Its address is not available
    - An *extended-format instruction* with *address field set to 00000* is used.
  - A modification record is inserted into the object code
    - Instruct the loader to *add the value of LISTA to this address field*. 
REF2 (LISTB+4)

- Control sections A and C
  - REF2 is an *external reference* (LISTB) plus a constant (4).
  - The address of LISTB is not available
    - An *extended-format instruction* with *address field set to 00004* is used.
  - A modification record is inserted into the object code
    - Instruct the loader to add the value of LISTB to this address field.

- Control section B
  - LISTB is defined within the control section.
  - Its address is available using *PC-relative addressing*.
  - No modification for relocation or linking is necessary.
Control section A
- ENDA and LISTA are defined within the control section.
- The difference between ENDA and LISTA is immediately available.
- No modification for relocation or linking is necessary.

Control sections B and C
- ENDA and LISTA are external references.
- The difference between them is not available
  - An extended-format instruction with address field set to 00000 is used.
- Two modification records are inserted into the object code
  - +ENDA
  - -LISTA
REF4 (ENDA-Lista+ListaC)

- **Control section A**
  - The values of ENDA and LISTA are internal. Only the value of LISTC is unknown.
  - The address field is initialized as 000014 (ENDA-Lista).
  - **One** Modification record is needed for LISTC:
    - +LISTC

- **Control section B**
  - ENDA, LISTA, and LISTC are all unknown.
  - The address field is initialized as 000000.
  - **Three** Modification records are needed:
    - +ENDA
    - -LISTA
    - +LISTC

- **Control section C**
  - LISTC is defined in this control section but ENDA and LISTA are unknown.
  - The address field is initialized as the *relative address of LISTC (000030)*
  - **Three** Modification records are needed:
    - +ENDA
    - -LISTA
    - +PROGC (**for relocation***) // Thus, relocation also use modification record
## Program Linking Example

<table>
<thead>
<tr>
<th>Label</th>
<th>Expression</th>
<th>Type</th>
<th>Program A</th>
<th>Program B</th>
<th>Program C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>LISTA, ENDA</td>
<td>LISTB, ENDB</td>
<td>LISTC, ENDC</td>
</tr>
<tr>
<td>REF1</td>
<td>LISTA</td>
<td>R</td>
<td>local, R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REF2</td>
<td>LISTB+4</td>
<td>R</td>
<td></td>
<td>local, R</td>
<td></td>
</tr>
<tr>
<td>REF3</td>
<td>ENDA-LISTA</td>
<td>A</td>
<td>local, A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REF4</td>
<td>ENDA-LISTA+LISTC</td>
<td>R</td>
<td>local, A</td>
<td></td>
<td>local, R</td>
</tr>
<tr>
<td>REF5</td>
<td>ENDC-LISTC-10</td>
<td>A</td>
<td></td>
<td>local, A</td>
<td></td>
</tr>
<tr>
<td>REF6</td>
<td>ENDC-LISTC+LISTA-1</td>
<td>R</td>
<td>local, R</td>
<td></td>
<td>local, A</td>
</tr>
<tr>
<td>REF7</td>
<td>ENDA-LISTA-(ENDB-LISTB)</td>
<td>A</td>
<td>local, A</td>
<td>local, A</td>
<td></td>
</tr>
<tr>
<td>REF8</td>
<td>LISTB-LISTA</td>
<td>R</td>
<td>local, R</td>
<td>local, R</td>
<td></td>
</tr>
</tbody>
</table>
Suppose the loader sequentially allocate the address for object programs

- See Fig. 3.10
- Load address for control sections
  - PROGA 004000 63
  - PROGB 004063 7F
  - PROGC 0040E2 51

- Fig. 3.10
  - Actual address of LISTC: 0030+PROGC=4112
Values of REF4, REF5, ..., REF8 in three places are all the same.
Relocation and Linking Operations Performed on REF4 from PROGA (Fig. 3.10b)

Figure 3.10(b) Relocation and linking operations performed on REF4 from PROGA.

3.2.3 Algorithm and Data Structures for a Linking Loader
Calculation of REF4 (ENDA-LISTA+LISTC)

- **Control section A**
  - The address of REF4 is 4054 (4000 + 54)
  - The address of LISTC is:
    \[
    \text{0040E2} + \text{000030} = \text{004112}
    \]
    (starting address of PROGC) (relative address of LISTC in PROGC)
  - The value of REF4 is:
    \[
    \text{000014} + \text{004112} = \text{004126}
    \]
    (initial value) (address of LISTC)

- **Control section B**
  - The address of REF4 is 40D3 (4063 + 70)
  - The value of REF4 is:
    \[
    \text{000000} + \text{004054} - \text{004040} + \text{004112} = \text{004126}
    \]
    (initial value) (address of ENDA) (address of LISTA) (address of LISTC)

Target Address are the same
After these control sections are linked, relocated, and loaded

- Each of REF4 through REF8 should have the same value in each of the three control sections.
- They are data labels and have the same expressions.

But not for REF1 through REF3 (instruction operation)

- Depends on PC-relative, Base-relative, or direct addressing used in each control section
  - In PROGA, REF1 is a PC-relative
  - In PROGB, REF1 is a direct (actual) address
- However, the target address of REF1~REF3 in each control section are the same
  - Target address of REF1 in PROGA, PROGB, PROGC are all 4040
3.2.3 Algorithm and Data Structure for a Linking Loader

- Algorithm for a *linking (and relocating) loader*
  - *Modification records* are used for relocation
    - Not use the *modification bits*
    - So that *linking* and *relocation* functions are performed using the same mechanism.

- This type of loader is often found on machines (e.g. SIC/XE)
  - Whose *relative addressing* makes relocation unnecessary for most instructions.
Implementation of An Assembler

- Data Structure
  - Operation Code Table (OPTAB)
  - Symbol Table (SYMTAB)
  - Location Counter (LOCCTR)

Source Program

Pass 1 assembler

Intermediate file

Pass 2 assembler

Object Program

LOCCTR

OPTAB

SYMTAB
Two-pass process (similar to the Assembler):
- Pass 1: assigns addresses to all external symbols
- Pass 2: performs the actual loading, relocation, and linking
Algorithm for a Linking Loader

- Input is a set of object programs, i.e., control sections
  - Since a control section may make an *external reference* to a symbol whose definition will appear later in the input stream

- Thus, a linking loader usually makes two passes over its input, just as an assembler does
  - Pass 1: assign addresses to all *external symbols*
  - Pass 2: perform the actual loading, relocation, and linking
Data Structures

- **External Symbol Table (ESTAB)**
  - For each external symbol, ESTAB stores
    - its name
    - its address
    - in which control section the symbol is defined
  - *Hashed organization*

- **Program Load Address (PROGADDR)**
  - PROGADDR is the beginning address in memory where the linked program is to be loaded (supplied by OS).

- **Control Section Address (CSADDR)**
  - CSADDR is the starting address assigned to the control section currently being scanned by the loader.

- **Control section length (CSLTH)**
Pass 1 Program Logic (Fig. 3.11a)

- Assign addresses to all external symbols
- The loader is concerned only with Header and Define record types in the control sections
- To build up ESTAB
  - Add control section name into ESTAB
  - Add all external symbols in the Define record into ESTAB
Pass 1 Program Logic (Fig. 3.11a)
Pass 1: (only Header and Define records are concerned)

begin
get PROGADDR from operating system
set CSADDR to PROGADDR (for first control section)
while not end of input do
  begin
    read next input record (Header record for control section)
    set CSLTH to control section length
    search ESTAB for control section name
    if found then
      set error flag (duplicate external symbol)
    else
      enter control section name into ESTAB with value CSADDR
    end
    while record type ≠ ‘E’ do
      begin
        read next input record
        if record type = ‘D’ then
          for each symbol in the record do
            begin
              search ESTAB for symbol name
              if found then
                set error flag (duplicate external symbol)
              else
                enter symbol into ESTAB with value (CSADDR + indicated address)
              end
            end
          end
        add CSLTH to CSADDR (starting address for next control section)
      end
    end
  end
end (Pass 1)

Figure 3.11(a) Algorithm for Pass 1 of a linking loader.
Load Map

- ESTAB (External Symbol Table) may also look like Load MAP

<table>
<thead>
<tr>
<th>Control section</th>
<th>Symbol name</th>
<th>Address</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROGA</td>
<td>LISTA</td>
<td>4040</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ENDA</td>
<td>4054</td>
<td></td>
</tr>
<tr>
<td>PROGB</td>
<td>LISTB</td>
<td>40C3</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>ENDB</td>
<td>40D3</td>
<td></td>
</tr>
<tr>
<td>PROGC</td>
<td>LISTC</td>
<td>4112</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ENDC</td>
<td>4124</td>
<td>0051</td>
</tr>
</tbody>
</table>
Pass 2 Program Logic (Fig. 3.11b)

- Perform the actual loading, relocation, and linking
- When *Text record* is encountered
  - Read into the specified address (+CSADDR)
- When *Modification record* is encountered
  - Lookup the symbol in ESTAB
  - This value is then added to or subtracted from the indicated location in memory
- When the *End record* is encountered
  - Transfer control to the loaded program to begin execution
- Fig. 3.11(b)
Pass 2 Program Logic (Fig. 3.11b)
Pass 2:

\[
\begin{align*}
\text{begin} & \quad \text{set CSADDR to PROGADDR} \\
& \quad \text{set EXECADDR to PROGADDR} \\
\text{while not end of input do} & \quad \text{begin} \\
& \quad \text{read next input record \{Header record\}} \\
& \quad \text{set CSLTH to control section length} \\
& \quad \text{while record type \neq 'E' do} & \quad \text{begin} \\
& \quad \text{read next input record} \\
& \quad \text{if record type = 'T' then} & \quad \text{begin} \\
& \quad \{\text{if object code is in character form, convert into internal representation}\} \\
& \quad \text{move object code from record to location (CSADDR + specified address)} \\
& \quad \text{end} \{\text{if 'T'}\} \\
& \quad \text{else if record type = 'M' then} & \quad \text{begin} \\
& \quad \text{search ESTAB for modifying symbol name} \\
& \quad \text{if found then} & \quad \text{add or subtract symbol value at location (CSADDR + specified address)} \\
& \quad \text{else} & \quad \text{set error flag (undefined external symbol)} \\
& \quad \text{end} \{\text{if 'M'}\} \\
& \quad \text{end} \{\text{while \neq 'E'}\} \\
& \quad \text{if an address is specified \{in End record\} then} & \quad \text{begin} \\
& \quad \text{set EXECADDR to (CSADDR + specified address)} \\
& \quad \text{add CSLTH to CSADDR} & \quad \text{end} \{\text{while not EOF}\} \\
& \quad \text{jump to location given by EXECADDR \{to start execution of loaded program\}} \\
& \quad \text{end} \{\text{Pass 2}\}
\end{align*}
\]

\textbf{Figure 3.11(b)} \quad \text{Algorithm for Pass 2 of a linking loader.}
Improve Efficiency

- Use *local searching* instead of multiple searches of ESTAB for the same symbol
  - Assign a *reference number* to each external symbol referred to in a control section
  - The reference number (instead of symbol name) is used in Modification records

- Avoiding *multiple searches* of ESTAB for the same symbol during the loading of a control section.
  - Search of ESTAB for each external symbol can be performed *once* and the result is *stored in a new table* indexed by the *reference number*.
  - The values for code modification can then be obtained by simply *indexing* into the table.
Improve Efficiency (Cont.)

- Implementation
  - 01: control section name
  - other: external reference symbols

- Example
  - Fig. 3.12
Object Programs Corresponding to Fig. 3.8 Using Reference Numbers for Code Modification (Fig. 3.12)

Figure 3.12 Object programs corresponding to Fig. 3.8 using reference numbers for code modification. (Reference numbers are underlined for easier reading.)
Object Programs Corresponding to Fig. 3.8 Using Reference Numbers for Code Modification (Fig. 3.12) (Cont.)

```
HPROC B 00000000007F
DLISTB 000060ENDB 000070
R02LISTA 03ENDA 04LISTC 05ENDC
*  
T0000360B0310000077202705100000
*  
T00007000F0000000FFFFF6FFFFFFFFFFFFFFF0000060
M00003705+02
M00003E05+03
M00003E05+02
M00007006+03
M00007006+02
M00007006+04
M00007306+05
M00007306+04
M00007606+05
M00007606+04
M00007606+02
M00007906+03
M00007906+02
M00007C06+01
M00007C06+02
E
```
Object Programs Corresponding to Fig. 3.8 Using Reference Numbers for Code Modification (Fig. 3.12) (Cont.)

```
HPROC 00000000051
DLISTC 000030ENDC 000042
R02LISTA Q3ENDA Q4LISTB Q5ENDB
0
T0000180C03100000771000040510000
0
T0000420F0000300000080000110000000000
M00001905+02
M00001D05+04
M00002105+03
M00002105-02
M00004206+03
M00004206-02
M00004206+01
M00004806+02
M00004B06+03
M00004B06-02
M00004B06-05
M00004B06+04
M00004EO6+04
M00004EO6-02
E
```

**Figure 3.12 (cont’d)**
# New Table for Figure 3.12

<table>
<thead>
<tr>
<th>Ref No.</th>
<th>Symbol</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PROGA</td>
<td>4000</td>
</tr>
<tr>
<td>2</td>
<td>LISTB</td>
<td>40C3</td>
</tr>
<tr>
<td>3</td>
<td>ENDB</td>
<td>40D3</td>
</tr>
<tr>
<td>4</td>
<td>LISTC</td>
<td>4112</td>
</tr>
<tr>
<td>5</td>
<td>ENDC</td>
<td>4124</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ref No.</th>
<th>Symbol</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PROGB</td>
<td>4063</td>
</tr>
<tr>
<td>2</td>
<td>LISTA</td>
<td>4040</td>
</tr>
<tr>
<td>3</td>
<td>ENDA</td>
<td>4054</td>
</tr>
<tr>
<td>4</td>
<td>LISTC</td>
<td>4112</td>
</tr>
<tr>
<td>5</td>
<td>ENDC</td>
<td>4124</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ref No.</th>
<th>Symbol</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PROGC</td>
<td>4063</td>
</tr>
<tr>
<td>2</td>
<td>LISTA</td>
<td>4040</td>
</tr>
<tr>
<td>3</td>
<td>ENDA</td>
<td>4054</td>
</tr>
<tr>
<td>4</td>
<td>LISTB</td>
<td>40C3</td>
</tr>
<tr>
<td>5</td>
<td>ENDB</td>
<td>40D3</td>
</tr>
</tbody>
</table>