

Introducing the Disassembler

Phil Hughes
PO Box 2847
Olympia WA 98507

An unassembler (or disassembler) is a program which takes machine language and converts it into what appears to be assembly language. The input to the disassembler is generally a program which is contained in computer memory and the output is generally printed output which looks much like an assembly language listing.

The one thing which a disassembler cannot do is generate the meaningful labels and comments which appeared in the original source program, but it can be very helpful as a debugging aid. For example, if you have purchased a new printer and want to interface it to the I/O routines of your computer system's BASIC interpreter you may have to modify those routines. Even though you do not have source code for those I/O routines, you can disassemble them and have something much easier to work with than an octal or hexadecimal memory dump. If you have written a program in machine code, you can also use the disassembler to document your program.

The Design

The disassembler presented here was designed to operate on a SWTPC 6800 system. The design should work with any computer and the program presented can be easily modified to work with any 6800 based system. Also, if Motorola comes out with the new 680X superchip to replace the 6800, the un-

assembler should be easy to update as the decoding is performed by the use of tables. The only portion of the decoding operation which is hard coded (instead of using tables) is the address mode (relative, indexed, etc.) of the instructions. If you desired to make the design even more flexible, this information could also be added to the tables.

In order to perform the disassembly function, two tables are used. Table 1, called the op code table, has an entry for every possible operation code (0 through 255). Each table entry is 2 bytes long and contains the following information:

1. An index into the mnemonic table.
2. Indicators for the accumulator used (if any).
3. The length of the instruction in bytes.
4. A flag to indicate if this instruction is an unconditional transfer of control (JMP, RTS, etc.).

An entry in this table is illustrated in Fig. 1.

The second table, called the mnemonic table, contains all the valid assembler mnemonics for the possible machine instructions. Note that pseudo operations such as EQU and FCC do not appear in the table. These do not generate instructions (although FCC does use memory space) and therefore could never be generated through the disassembly process. An entry in this table is illustrated in Fig. 2.

Table 1 shows the assumptions that were made con-

OP CODE				L	B	C
INDEX		A	B	A	Y	N
				B	T	T

- Op code index — Pointer into the Mnemonic table.
A — Indicates A register used if set.
B — Indicates B register used if set.
LAB — Indicates unconditional transfer instruction.
BYT CNT — Length of instruction in bytes.

Fig. 1. Layout of an op code Table Entry.

cerning the address mode of the instructions. The exception (we always have an exception) is that op code 8D is the BSR instruction and is therefore relative, not immediate as would be expected.

The Program

The program flow is straightforward. Once the user sets location A002 and A003 to the address of the program to be disassembled and initiates execution of the unassembler the following takes place.

The first byte of data (program to be disassembled) is picked up, multiplied by 2 and used to compute the address of the appropriate op code table entry. The data address (from A002 and A003) is converted to a character string and put into the output line buffer.

Next, using the op code table entry the mnemonic is picked up from the mnemonic table and transferred to the output line buffer. Then the address field is converted and placed in the print line, using the A and B fields and the length field from the op code table. If the address mode is relative, the offset is added to the address of the

next instruction and saved in the print line as the absolute address. The print routine is then called to print the data contained in the output buffer.

Finally, if the op code table entry had the lab bit set indicating that we have just processed an unconditional transfer instruction, a period is placed in position 7 of the print buffer so that the next line printed will have a flag indicating that it should have a label. This is because the only path to this instruction would be through a transfer of control instruction such as BRA.

If decoding the op code does not yield a valid mnemonic (mnemonic index of 0) three asterisks are printed in the operation field followed by the op code byte displayed in hexadecimal.

Using the Unassembler on a SWTPC 6800

If you have a SWTPC 6800 system configured with the AC-30 cassette interface on I/O port 1 and PR-40 printer on I/O port 7 and at least 8K bytes of memory then the following procedure would load the unassembler and cause it to unassemble itself:

First 4 bits of op code (hex)	Address Mode	1076	105	MIKBUG
1,3,4 or 5	Inherent	1077	JSP	1000
2	Relative	107C	LFX	1002
6,A or E	Indexed	107F	LDA H	0000
7,B or F	Extended	1041	STA H	1000
8,C	Immediate	1044	ASL A	
9,D	Direct	1045	LDA #1000	
		1048	STX	1000
		104B	BCC	1050
		104D	INC	1000
		1050	LDA B	1000
		1051	ASH	
		1054	STA A	1000
		1057	BCC	1050
		1059	INC	1000
		105C	LDA #1000	
		105F	LDA A	0002
		1062	JSP	1070
		1065	LDA A	0000
		1068	JSP	1070
		106C	LDA	1000
		106E	LDA A	0000
		1070	STA A	1000
		1072	LDA B	0100
		1075	STA B	1000
		1079	AND B	#00
		107A	LDA	0002
		107C	LDA A	0100
		107F	STA A	1000
		1082	LDA A	0000
		1084	STA A	1000
		1087	INC	
		1089	BCC B	
		108B	BNE	1087
		108E	STX	0002
		1090	LDA A	1000
		1092	TAB	
		1095	LDA #1000	
		1098	STX	1000
		109B	ASL H	
		109E	BCC	109F
		10A0	INC	1000
		10A3	CLC	
		10A5	ASH	
		10A8	BCC	10A5
		10AB	INC	1000
		10AD	CLC	
		10AF	ASH	
		10B0	BCC	10A5
		10B3	INC	1000
		10B5	CLC	
		10B8	ASH	
		10BB	BCC	10B5
		10BD	INC	1000
		10BF	CLC	
		10C0	ASH	
		10C3	BCC	10C0
		10C5	INC	1000
		10C8	CLC	
		10CB	ASH	
		10CD	BCC	10C0
		10CF	INC	1000
		10D1	LDA B	#41

Table 1. Address Mode Decoding.

1. Load the object tape using the MIKBUG L key-in.
2. Using the MIKBUG M key-in set location A002 to the address of the program to be disassembled. In this case it is 2B36, the first instruction of the disassembler.
3. Enter G to start program execution.
4. When the disassembler has decoded enough push reset to stop execution.

Typing G will cause the disassembler to resume operation at the next location. Also, by changing the contents of A002 and typing G the disassembler will resume operation at the new starting address.

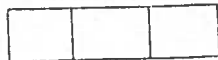
Modifications

With only minor modifications the disassembler can be changed to operate with almost any 6800 based system. To change it to print on the control interface it is only necessary to delete the PR-40 handler code and insert the following:

```
BINIT RTS
PDATA1 EQU $E07E
```

For anyone who does not use MIKBUG it is only necessary to replace the code at BINIT with the necessary printer initialization code and the code at PDATA1 with code which will print the ASCII string pointed to by the index register. Note that the end of string is signified by an ASCII EOT (hex 04) character.

If you have any occasional problems with strange looking output it is most likely an error in the tables. Check carefully, one missing byte can skew all the other data and generate very strange results. ■



3 ASCII character mnemonic

Fig. 2. Layout of Mnemonic Table Entry.

PAGE 001 UNASS

Address	Op Code	Op Name	UNASS
1076	MIKBUG		
1077	JSP	1000	
107C	LFX	1002	
107F	LDA H	0000	
1041	STA H	1000	
1044	ASL A		
1045	LDA #1000		
1048	STX	1000	
104B	BCC	1050	
104D	INC	1000	
1050	LDA B	1000	
1051	ASH		
1054	STA A	1000	
1057	BCC	1050	
1059	INC	1000	
105C	LDA #1000		
105F	LDA A	0002	
1062	JSP	1070	
1065	LDA A	0000	
1068	JSP	1070	
106C	LDA	1000	
106E	LDA A	0000	
1070	STA A	1000	
1072	LDA B	0100	
1075	STA B	1000	
1079	AND B	#00	
107A	LDA	0002	
107C	LDA A	0100	
107F	STA A	1000	
1082	LDA A	0000	
1084	STA A	1000	
1087	INC		
1089	BCC B		
108B	BNE	1087	
108E	STX	0002	
1090	LDA A	1000	
1092	TAB		
1095	LDA #1000		
1098	STX	1000	
109B	ASL H		
109E	BCC	109F	
10A0	INC	1000	
10A3	CLC		
10A5	ASH		
10A8	BCC	10A5	
10AB	INC	1000	
10AD	CLC		
10AF	ASH		
10B0	BCC	10A5	
10B3	INC	1000	
10B5	CLC		
10B8	ASH		
10BB	BCC	10B5	
10BD	INC	1000	
10BF	CLC		
10C0	ASH		
10C3	BCC	10C0	
10C5	INC	1000	
10C8	CLC		
10CB	ASH		
10CD	BCC	10C0	
10CF	INC	1000	
10D1	LDA B	#41	

Example A. Sample Program Execution (disassembly of portion of Disassembler Program).

1076	UNASS	MIKBUG	
2B36	FC	0002	LDS #0000
1039	ED	1007	JSP BINIT
103C	ASH		EQU #
			* START ADDRESS
103E	FC	0002	LDA #0002
103F	DC	00	LDA A 0000
1041	E7	100A	STA A SRVA
1044	46		ASL A

MIKBUG

Disassembler
in now located
at 2B00
set PGM CNTR
TO 2B36 TO
START

1845 CE 1002 LDX #DPS
 1846 FE 1000 STX TMP
 1848 24 03 BCC NOC
 1849 7C 1000 INC TMP
 1850 F6 1000 NOC LDR B TMP+1
 1853 18 ABR

 1854 E7 1000 STA A TMP+1
 1857 24 03 BCC NOC
 1859 7C 1000 INC TMP

 * ADDRESS
 185C EQU *
 185D CE 1000 LDX #LIN
 185F 00 1002 LDR A #D002
 1862 00 1070 JSR COH
 1865 00 1007 LDR A #D003
 1868 00 1070 JSR COH

 * OPS INDEX
 186B FE 1000 LDX TMP
 186E A6 03 LDR A D.X
 1870 E7 1000 STA A TMP

 * REG/LENGTH
 1873 00 01 LDR B 1.X
 1875 F7 1000 STA B BAR
 1878 04 03 AND B #03
 187A FE 1002 LDX #D002

 * SAVE POSSIBLE OPERANDS
 187D A0 01 LDR A 1.X
 187F E7 1000 STA A OPER
 1882 A6 03 LDR A 2.X
 1884 E7 1000 STA A OPER+1

 * UPDATE INST POINTER
 1887 00 NOPE INX

 1888 50 DEC B
 1889 26 FC BNE NOPE
 188B FF 1002 STX #D002
 188E 00 1000 LDR A TMP
 1891 16 TAB

 1892 00 1000 LDX #INCR
 1895 FF 1000 STX TMP
 1898 40 ASL A

 1899 24 04 BCC NOE
 189B 7C 1000 INC TMP
 189E 00 CLC

 189F 10 NOE ABR

 18A0 24 03 BCC NOF
 18A2 7C 1000 INC TMP
 18A5 00 CLC
 18A6 00 1000 NOF LDR A TMP+1
 18A9 24 03 BCC NOG
 18AB 7C 1000 INC TMP
 18AE E7 1000 NOG STA A TMP+1
 18B1 FE 1000 LDX TMP
 18B4 A6 03 LDR A D.X
 18B6 E7 101F STA A LIN+17
 18B9 A6 01 LDR A 1.X
 18BB E7 1000 STA A LIN+10
 18BE A6 03 LDR A 2.X
 18C0 E7 1021 STA A LIN+19
 18C3 06 20 LDR B #'

 18C5 F7 1002 STA B LIN+20
 18C8 00 1000 LDR A BAR
 18CB 05 00 BIT A #100
 18CD 27 00 BCR SREG
 18CF 2A 04 BPL DREG
 18D1 00 41 LDR B #1A
 18D3 20 00 BFA SREG
 18D5 0A 40 BFG LDR B #1B
 18D7 F7 1023 SREG STA B LIN+21

18A0 FE 1000 LDX #LIN+24
 18A2 00 101F LDR A LIN+17
 18A4 01 0A CMP A #'A
 18A6 20 00 BNE NOTBAD
 18A8 06 100A LDR A SAVA
 18AB 00 1070 JSR COH
 18AD 20 40 BFA NOOPR
 18AF 00 1000 NOPE LDR A BAR
 18B1 05 02 BCF BIT A #102
 18B3 27 30 BEQ NOOPR
 18B5 00 100A LDR A SAVA
 18B7 01 00 CMP A #100
 18B9 27 12 BCC NOTIMM
 18BB 24 0A AND A #100
 18BD 07 100A STA A SAVA
 18BF 01 00 CMP A #100
 18C1 27 04 BEQ IMM
 18C3 01 00 CMP A #100
 18C5 26 05 BNE NOTIMM

 1007 00 20 * FLAG IMM ADDRESS
 1009 07 00 IMM LDR A #'A
 100B 00 STA A D.X
 100E 00 INX

 100C NOTIMM COH *
 * CHECK FOR REL ADDRESS
 100E 00 100A LDR A SAVA GET
 100F 01 00 CMP A #100 BFA
 1011 27 05 BEQ REL YES
 1013 04 00 AND A #100
 1015 01 00 CMP A #100 DIF
 1017 27 4F BEQ REL YES
 1019 00 1000 LDR A OPER
 101B 00 1070 JSR COH
 101D 00 1000 LDR A BAR
 101F 05 01 BIT A #11
 1021 27 00 BEQ NOOPR
 1023 00 100A LDR A OPER+1
 1025 00 1070 JSR COH
 1027 1000 NOOPR EQU *
 1029 00 100A LDR A SAVA
 102B 01 0A CMP A #10A
 102D 27 00 BEQ INCRD
 102F 01 00 CMP A #100
 1031 27 04 BEQ INCRD
 1033 01 00 CMP A #100
 1035 27 04 BEQ INCRD
 1037 01 00 CMP A #100
 1039 26 0A BNE NOTI

 * PUT 'A' IN LINE
 103B 00 00 INCRD LDR A #'A
 103D 07 00 STA A D.X
 103F 00 INX

 1040 06 50 LDR A #1X
 1042 07 00 STA A D.X
 1044 00 INX

 1045 NOTI EQU *
 * PRINT LINE
 1045 00 00 LDR A #D00
 1047 07 00 STA A D.X
 1049 00 0A LDR A #D0A
 104B 07 01 STA A 1.X
 104D 00 04 LDR A #D04
 104F 07 00 STA A 2.X
 1051 00 1000 LDX #LIN
 1053 00 1000 JSR PDAT1

 * IF THE LAST INST WAS
 * UNCONDITIONAL XFER THEN
 * FLAG LABEL FIELD ON THE
 * NEXT LINE
 1057 06 20 LDR A #'A

 * CHECK LAB BIT
 1059 00 1000 LDR B BAR
 105B 04 10 AND B #10
 105D 27 00 BEQ INCLAB
 105F 00 20 LDR A #'A
 1062 07 1014 INCLAB STA A LIN+6

1065 7E 1030
 1066 4F
 1067 5E 1008
 1068 2A 01
 106E 4A
 106F 0C
 1070 8D 0003
 1071 89 0002
 1076 80 85
 1079 17
 1079 80 03
 107D 20 08
 107E

COMPUTE RELATIVE ADDR
 REL CLR A
 LDR B OFEF
 EPL MIREL
 DEC A
 MIREL CLC
 ADD B #0003 L 0
 ADD A #0002 H 0
 RFR CDH CONV
 TGR
 BSR CDH CONV
 BRH #0011
 CDH EQU *

1080 39
 108C 07 801C
 108F 0A 801C
 1092 70 8010
 1095 2A FE*
 10B7 39
 DUTCH EQU #
 STA A PAPIA
 LDR A PAPIA
 TST PAPIA+1
 EPL PLUP
 RTS

* END OF FF-40 HANDLER

1070 36
 107E 80 06
 1080 32
 1081 00
 1082 80 06
 1084 00
 1085 39
 1086 44
 1087 44
 1088 44
 1089 44
 108A 84 0F
 108C 88 39
 108E 81 79
 1090 87 02
 1092 80 07
 1094 87 00
 1096 39

CONVERT BINARY NUMBER
 IN A-REG TO 2 HEX
 CHRS AND SAV IN
 ADDR IN X-REG
 INC X-REG
 BSR A
 BSR CBHLH
 PUL A
 INK
 BSR CBHPH
 INK
 RTS
 CBHLH LSR A
 LSR A
 LSR A
 LSR A
 CBHPH AND A #1F
 AND A #10
 ORF A #139
 BLS CBHPH
 AND A #17
 CBHPH STA A 0,X
 RTS

1089 3A
 10C2 3A
 10CA 3A
 10CD 4E
 10D0 4F
 10D0 50
 10DE 54
 10E5 41
 10E0 50
 10E1 54
 10E2 50
 10E3 41
 10E4 49
 10E5 4E
 10E6 58
 10E7 44
 10E8 45
 10E9 58
 10EA 43
 10EB 40
 10EC 56
 10ED 53
 10EE 45
 10EF 56
 10F0 43
 10F1 40
 10F2 40
 10F3 53
 10F4 45
 10F5 43
 10F6 43
 10F7 40
 10F8 49
 10F9 52
 10FA 45
 10FB 43
 10FC 53
 10FD 42
 10FE 41
 10FF 43
 10E0 42
 10E1 41
 10E2 54
 10E3 41
 10E4 40
 10E5 54
 10E6 42
 10E7 41
 10E8 44
 10E9 41
 10EA 41
 10EB 41
 10EC 42
 10ED 41
 10FF 42
 10FF 52
 10F0 41
 10F1 40
 10F2 43
 10F3 49
 10F4 43
 10F5 40
 10F6 52
 10F7 42

MNET FCC /***/
 FCC /NOP/
 FCC /INX/
 FCC /DEK/
 FCC /CLV/
 FCC /SEV/
 FCC /CLC/
 FCC /SEC/
 FCC /CLI/
 FCC /SEI/
 FCC /SEB/
 FCC /CDB/
 FCC /TAB/
 FCC /TEA/
 * 410
 FCC /DAB/
 FCC /ABR/
 FCC /BRA/
 FCC /BHL/
 FCC /BSL/
 FCC /BCL/
 FCC /BCC/
 * 39 RTS

 * PR-40 HANDLER
 * REPLACE THIS CODE
 * WITH
 * BINIT RTS
 * PORTA1 EQU JDATE
 * TO OUTPUT TO CONTROL INTERFACE
 *
 * FF-40 INITIALIZATION
 * 39 RTS

1097 86 FF
 1099 87 801C
 109C 86 3E
 109E 87 8010
 10A1 39
 BINIT LDR A #1FF
 STA A PAPIA
 LDR A #13E
 STA A PAPIA+1
 RTS

10A2
 10A2 80 08
 10A4 08
 10A5
 10A5 80 00
 10A7 81 84
 10A9 20 F7
 PORTA2 EQU #
 BSR DUTCH
 INK
 PRINT LINE ON PR-40
 PORTA1 EQU #
 LDR A #1
 ORF A #1
 BNE PORTA2
 AS IS FOR TELETYPE
 BD EOTE JSR
 39 RTS

7
 12
 5
 IE

PROGRAM NAME _____

WRITTEN BY _____

DATE SUBMITTED _____

PAGE _____ OF _____

LOC. #	PSEUDO ADDRESS	#	MNEMONIC INSTRUCTIONS & DATA	OCTAL CODE	COMMENTS
2E55			2C		
2E56			A5.		
2E97		86	LDA.A.	03	
2E98			#03.		
2E99		B7	STA.A.	8000	
2E9A			80		
2E9B			00		
2E9C		86	LDA.A	45	
2E9D			#45		
2E9E		B7	STA.A.	8000	
2E9F			80		
2EA0			00		
2EA1		39	RTS		
2EAC		C6	LDA.B	02	
2EAD			02		
2EAE		F5	BIT B	8000	
2EAF			80		
2EB0			00		
2EB1		27	BEQ.	2EAE	
2EB2			FB.		

