

修正可能な逆アセンブラについて

金 野 靖 英*

A Modifiable Disassembler

Yasuhide KONNO

要旨

修正可能な逆アセンブラを作成したもので報告する。この逆アセンブラは、出力リストに名前や注釈の追加、数値の出力指定ができるものである。応用例として、正弦・余弦関数ルーチンの逆アセンブラを示した。

Synosis

This paper describes a modifiable disassembler. It has a function of the comment insertion, the name substitution and the value designation. The disassembling of sin(con)-routine is indicated as its application.

1. ま え が き

逆アセンブラは、原始プログラム・リストのないシステムの解説や、コンパイラ作成時のオブジェクトコード発生の研修等に利用される。逆アセンブルとは、機械語によって構成された目的プログラムを元の原始プログラムに復元することである。この場合、問題となるのは、目的プログラム中では、命令部分と数値部分が明確に区分されていず、原始プログラムにはあった定数や作業場所の番地を表わす名前が喪失していることである。したがって、この名前の復元が逆アセンブラの大きな課題である。命令のオペランドの参照番地によって、自動的に数値部に名前をつけることは可能である。しかしながら、この名前は、単なる記号であり、意味情報としての価値は少ない。さらにプログラムの途中で、ベースレジスの値を替えたり、命令を書き替えるような、高級なプログラムにあっては、名前づけは非常な困難をとまなう。このような点から、逆アセンブラの自動化には、多くの問題が残る。名前は後から与えるも

のとして、命令についてのみ、逆アセンブルすることは容易である。数値部や作業場所では、意味不明の命令に変換したり、単なる数値を出力することとなり、原始プログラムとしての体裁は悪いが、解説は可能である。解説後、適当な名前をつけ、名前情報をつけ加えて再度アセンブルすると、元の原始プログラムが得られる。この方法は、人間が名前づけに関係しているのが特徴である。より完全な復元や能率の面から後者が望ましいと思われる。これを修正可能な逆アセンブラと名づけた。ハードウェアは、HITAC-8250である。

2. 逆アセンブラの概要

処理手順として、目的プログラムを、前もって所定のファイルに格納しておくものとする。逆アセンブルの手順は2パスとなるが、処理は1パスであり、ユーザの利用法はパラメータのセットによって行なわれる。1パス目はパラメータのセットを行なわず、目的プログラムを解説し、命令の変換のみを行なう。この手順では、数値や作業領域も命令とみなして変換するので、この部分は意味不明の命令や数値を出力する。又レジスタの初

*助教授 電気工学科

期値は、すべて0として取り扱う。出力リストには、1パス目に限り、オペランドの参照番地を16進で計算して出力する。2パス目は、出力リストにより、レジスタの設定、名前の設定、数値の出力、コメントの挿入等のパラメータをセットして、逆アセンブルする。

3. パラメータ

◎ R パラメータ

R0=05AFのように、Rの次に汎用レジスタの番号を、=の次に16進4桁で、初期値を与える。カンマで区切って、複数のレジスタの設定が可能である。R 5=0005 (0F20)のように、カッコでくくって、16進4桁で、設定される番地を指定することができる。この番地は先頭よりの相対番地である。

◎ NAME パラメータ

名前の設定は、N-NAMAE=0F21 (4 X)のように、Nの次に名前を表わし、=の次に、16進4桁で設定される相対番地を表わす。さらにカッコでくくって、その長さをバイトで表わし、Xは16進で数値を出力することを表わす。この部分は、Cなら文字、Pなら固定小数点数、Fなら浮動小数点数、SならDS命令を表わす。複数の指定の場合は、カンマで区切って続ける。

◎ C パラメータ

コメントの挿入は、C0F23'COMENT'のように、Cの次に挿入する相対番地を設定し、' 'の中に挿入するコメントを記す。命令の出力の後に、このコメントを出力する。挿入コメントは73文字までである。又、挿入指定番地の次に+の場合は、その次の行に、-の場合は、その前の行にコメントが挿入される。したがって、ステートメントの間は最大2行が挿入される。例外として C0000-とした場合は、最初の3行がとられ、2行目にコメントが挿入される。

4. 逆アセンブルの処理

パラメータがセットされない場合は、所定のファイル内の目的プログラムを、逆アセンブルして出力する。パラメータがセットされると、これをそれぞれのテーブルにセットし目的プログラムのアドレスカウンタにより、コメント処理、レジスタ処理、名前処理、数値処理を行なう。この処理では、出力領域への文字設定が主である。命令

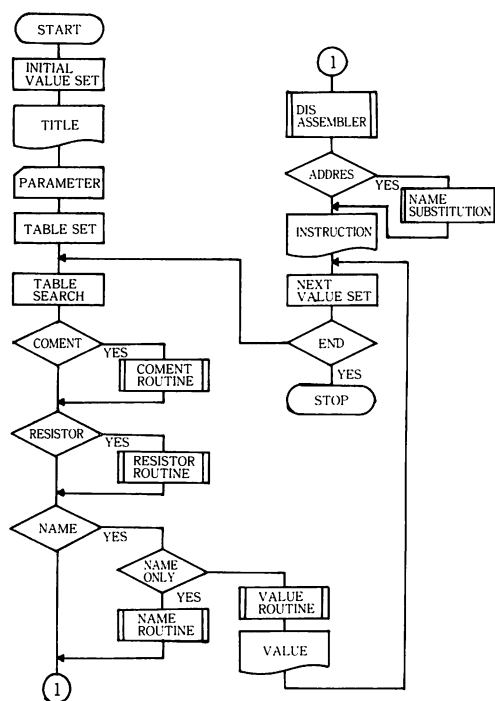


図1 逆アセンブラのフローチャート

の場合は、逆アセンブルーチンで解析後、アドレス情報にもとづいて、アドレス部と名前の置換を行なっている。目的プログラムは、所定のファイルより、主記憶に移されてから、処理される。処理のフローチャートを図1に示す。

5. 応用例

FÖRTRANの基本外部関数であるSIN(COS)ルーチンの逆アセンブル例を図2に示す。

名前づけ、コメントの挿入で、解読しやすい出力リストとなっている。この逆アセンブルリストによると、ルーチンは、次のような数値計算を行なっている。

◎ SIN の場合

$$\sin\left(-\frac{\pi}{4}x\right) = a_0x + a_1x^3 + a_2x^5 + a_3x^7$$

$$(-1 \leq x < 1)$$

$$a_0 = 0.78539819$$

$$a_1 = -0.08074546$$

$$a_2 = 0.00249001$$

$$a_3 = -0.00003595$$

LGC	OBJECT CODE	STMT	SOURCE STATEMENT
*** SIN COS ROUTINE ***			
000	47F0 F008	1	BC 15,BGN1
004	03C3D6E2	2	COS DC X'03C3D6E2'
008	90D1 D008	3	BGN1 STM 13,1,8(13)
00C	9202 F0F3	4	MVI 243(15),X'02'
010	58E1 0000	5	L 14,0(1,0)
014	45F0 F038	6	BAL 15,ENTRANCE
018	47F0 F008	7	BC 15 BGN2
01C	03E2C9D5	8	SIN DC X'03E2C9D5'
020	90D1 D008	9	BGN2 STM 13,1,8(13)
024	9200 F0DB	10	MVI 219(15),X'00'
028	58E1 0000	11	L 14,0(1,3)
02C	9180 E000	12	TM 0(14),X'80'
030	4780 F020	13	BC 8,ENTRANCE
034	9204 F0DB	14	MVI 219(15),X'04'
038	6840 F0D0	15	ENTRANCE LD 4,01
03C	6820 F0D8	16	LD 2,FLGAREA
040	2802	17	LDR 0,2
042	780E 0000	18	LE 0,0(14,0)
046	3000	19	LPER 0,0
048	7900 F0EC	20	CE 0,DMAx
04C	47A0 F0A0	21	BC 10,FC0N
050	6C00 F0E0	22	MD 0, PA14
054	3904	23	CER 0,4
056	4740 F048	24	BC 4,SC1
05A	2E02	25	AWR 0,2
05C	3820	26	LER 2,0
05E	2802	27	SDR 0,2
060	7020 F0AC	28	SC1 STE 2,DATA1
064	9101 E0AF	29	TM DATA2,X'01'
068	4780 F059	30	BC 8,SC2
06C	2804	31	SDR 0,4
06E	3000	32	LPER 0,0
070	4110 0004	33	SC2 LA 1,4
074	9103 F0AF	34	TM DATA1,X'03'
078	4740 F063	35	BC 4,SC3
07C	1811	36	SR 1,1
07E	3840	37	LER 4,0
080	7900 F0E8	38	SC3 CE 0,DMIN
084	4720 F072	39	BC 2,SC4
088	3800	40	SER 0,0
08A	3C00	41	SC4 MER 0,0
08C	3820	42	LER 2,0
08E	7C01 F0B4	43	ME 0,180(1,15)
09C	7A01 F0BC	44	AE 0,18B(1,15)
096	3C02	45	MER 0,2
098	7A01 F0C4	46	AE 0,196(1,15)
09C	3C02	47	MER 0,2
09E	7A01 F0CC	48	AE 0,204(1,15)
0A2	3C04	49	MER 0,4
0A4	9104 F0AF	50	TM DATA2,X'04'
0A8	4780 F096	51	BC 8,SC5
0AC	3100	52	LNER 0,0
0AE	58ED 000C	53	SC5 L 14,12(13,0)
0B2	92FF D00C	54	MVI 12(13),X'FF'
0B6	07FE	55	BCR 15,14
0B8	58F0 F0B0	56	FC0N L 15,FC
0BC	45EF 003C	57	BAL 14,60(15,0)
0C0	234C0000	58	AD DC X'234C0000'
0C4	000000	59	DATA1 DS CL3
0C7	00	60	DATA2 DS CL1
0CA	000082D8	61	FC DC X'000082D8'
0CC	BD25B368	62	A3 DC X'BD25B368'
0D0	BE14F17D	63	B3 DC X'BE14F17D'
0D4	3EA32F62	64	A2 DC X'3EA32F62'
0D8	3F40ED0F	65	B2 DC X'3F40ED0F'
0DC	C014A8BC	66	A1 DC X'C014A8BC'
0E0	C04EF4EE	67	B1 DC X'C04EF4EE'
0E4	40C90FDD	68	A0 DC X'40C90FDD'
0EA	4110000000000000	69	D1 DC X'4110000000000000'
0F6	4600000000000000	70	FLGAREA DC X'4600000000000000'
0F8	41145F306DC9C883	71	PA14 DC X'41145F306DC9C883'
100	3E100000	72	DMIN DC X'3E100000'
104	45C90FDA	73	DMAx DC X'45C90FDA'

図2 SIN-COSルーチンの逆アセンブルリスト

◎ COS の場合

$$\cos\left(\frac{\pi}{4}x\right) = b_0 + b_1x^2 + b_2x^4 + b_3x^6$$

$$(-1 \leq x \leq 1)$$

$$b_0 = 1.0$$

$$b_1 = -0.030842483$$

$$b_2 = 0.01585107$$

$$b_3 = -0.00031957$$

6. む す び

修正可能なアセンブラは、実際に使用してみると、非常に便利なものである。自動的に名前づけを行なう逆アセンブラも使用してみたが、命令の不自然さや、数値部や作業領域が、インデックス・レジスタに関係してアドレスづけられている場合は、意味不明となってしまう、1パスで完全な復元リストが得られない。特にインデックス修飾は、原始プログラム作成には欠くことのできないものである、問題は大きい。この点で、人間との

協力で完全なリストを得る2パス方式が秀れていると思われる。又、判明している部分のパラメータをセットすることにより、その事点でのリストを出力させることも可能で、有効である。今後、システムの解読やオブジェクトコードの研修等で活躍が期待できる。最後に、種々のルーチンの作成に協力いただいた卒研生、本間宏君(10期)に深謝いたします。

参 考 文 献

- (1) 山内二郎 電子計算機のための数値計算法III
培風館(1972)
- (2) HITAC 8250 NDOS アセンブラ・プログラミング文法 日立製作所(1974)
- (3) HITAC 8250 NDOS アセンブラ・プログラマズガイド 日立製作所(1974)
- (4) HITAC 8250 形 処理装置 日立製作所(1974)
- (5) bit VOL.12 NO 4 ソフトウェア考古学——逆アセンブラについて

(昭和55年11月29日受理)