Upward Migration

Part 1: Translators

Using translation programs to move CP/M-86 programs to CP/M and MS-DOS

Roger Taylor and Phil Lemmons
c/o BYTE Publications Inc.
POB 372
Hancock, NH 03449

In software migrations as in any other, the first thing you must know is where you are and where you are going. If a Californian decides to move to Australia, he can call an airline, ask for a ticket from San Francisco to Sydney, and the airline clerk will be happy to reserve a seat for him. If the Californian asks for a ticket from Los Angeles-or-San Francisco to Sydney, however, the airline clerk will find the request confusing. If the Californian asks for a ticket from Los Angeles-or-San Francisco to Sydney under Catholic rule, or a ticket from Los Angeles-or-San Francisco to Sydney under Protestant rule, the airline clerk will probably say, "You've already spent too much time in the hot tub, buddy. Stay in California. You're right where you belong."

In this little parable of modern times, Los Angeles is the 8080, San Francisco is the Z80, Sydney is the 8086, Catholicism is CP/M-86, and Protestantism is MS-DOS. The operating systems are represented by religions because they generate similar passions, controversies, true believers, and skeptics.

XLT86 takes 8080 source code and converts it into 8086 source code in an intelligent manner using data-flow-analysis techniques.

Is there really any need to explain why Los Angeles stands for the 8080 and San Francisco for the Z80? Or that the airlines stand for the software houses that have written translation programs?

You've probably guessed who the guy is who's trying to buy the ticket. He's the experienced 8080 or Z80 programmer, and the hot tub symbolizes his strong preference for staying right where he is—on one familiar processor with one familiar operating system. What could be cozier?

The programmer may not have the urge for going, but he has to go to one unfamiliar processor and to both operating systems. And so do the rest of us.

We're lucky to have software from at least three companies to help us along. In this article, the first of two parts, we will review three CP/M-80-to-8086 translator programs. We'll

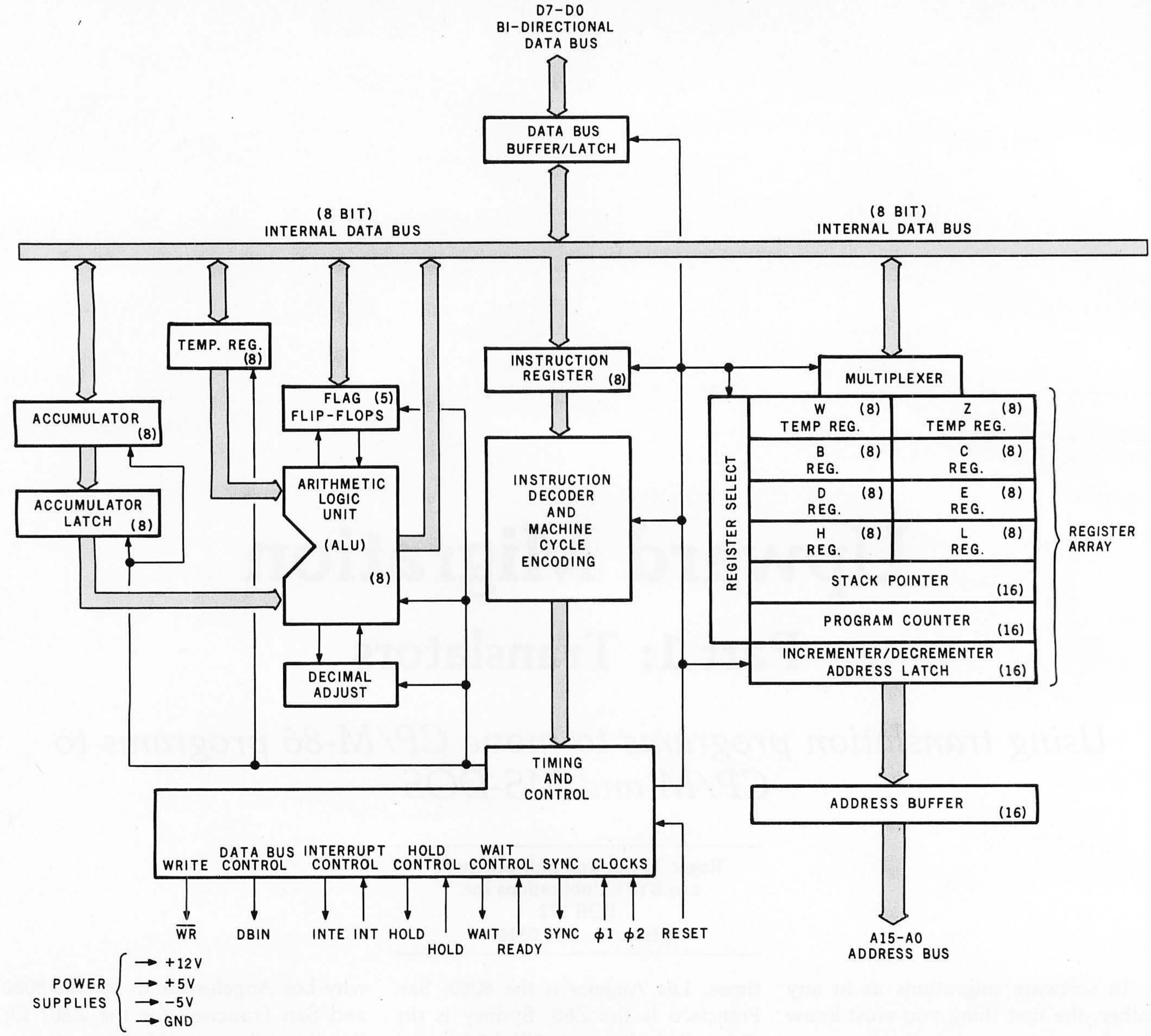

Figure 1: *Architecture of the 8080 microprocessor showing the internal registers.*

show the output that each of the three translators produced from the same original code. In addition, we'll make some observations about the differences in adapting the translated code to CP/M-86 and MS-DOS. Next month, we'll take a closer (although still not comprehensive) look at CP/M-86 and MS-DOS.

Orientation to the 8086

The first thing we have to do is examine the differences between the familiar 8080 and Z80 microprocessors and the 8086. For reference, figure 1 shows the registers and architecture of the 8080; figure 2 shows the registers and architecture of the Z80. We'll make few comments about these registers because they are familiar to you if you have 8080 or Z80 source code that you want to translate.

Figure 3 shows the registers and architecture of the 8086. Since the 8086 is less familiar, we'll take a brief look at it for orientation. (For further enlightenment, see *The 8086 Book*, Russell Rector and George Alexy, Osborne/McGraw-Hill, 1980, and *The 8086 Primer*, Stephen P. Morse, Hayden, 1980.) The 8086 is, of course, a 16-bit microprocessor. The 8088 is the same as the 8086 internally. Externally, however, they appear different due to the 8-bit bus of the 8088 and the 16-bit bus of the 8086. This means that programs that run on the 8088 will also run on the 8086 assuming that the memory resources and peripheral resources are the same. In general, statements in this article that apply to the 8086 apply to the 8088 as well.

The 8086 can access up to 1 megabyte of memory and as many as 65,000 input/output ports. The megabyte of memory is 2^{20} 8-bit bytes; any two consecutive bytes are a 16-bit word. Some 8086 instructions access bytes; others access words.

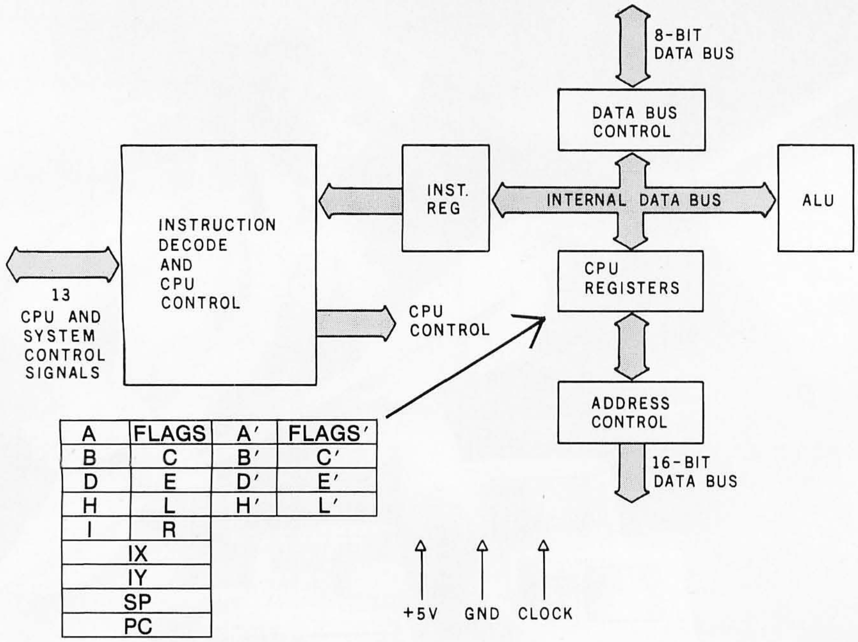

Figure 2: *Architecture of the Z80 microprocessor showing the internal registers.*

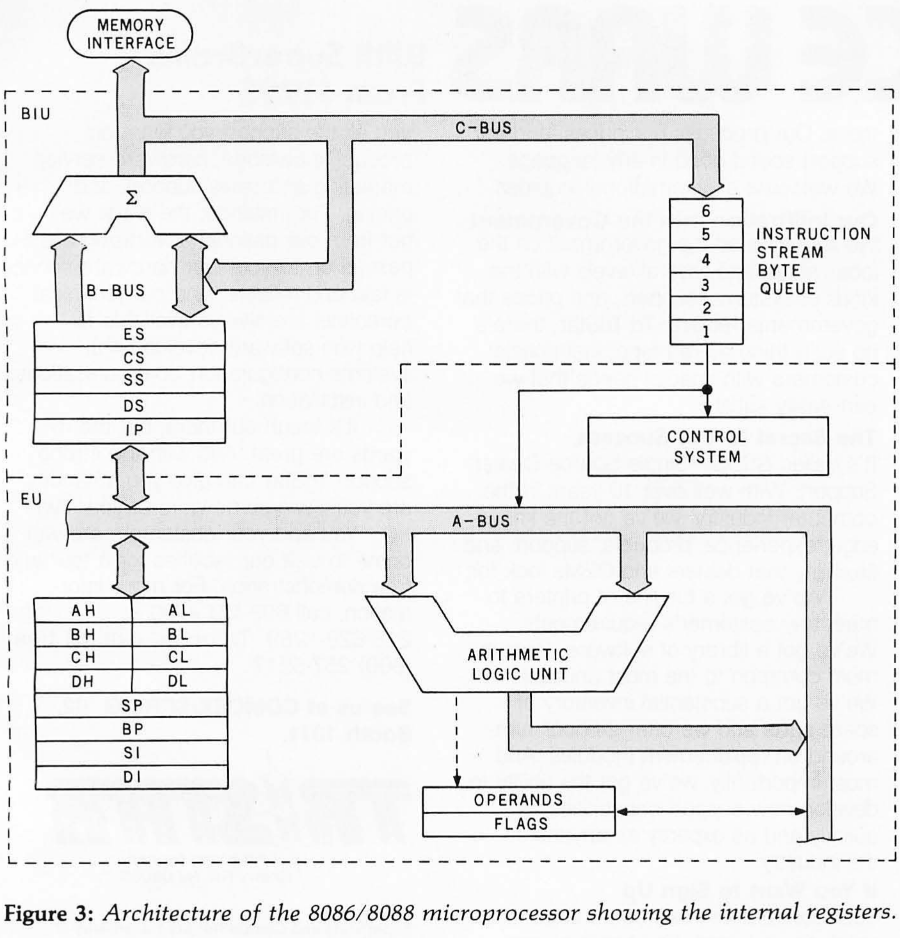

Figure 3: *Architecture of the 8086/8088 microprocessor showing the internal registers.*

Sixteen bits are not enough to address a megabyte of memory. The 8086 manages to do so, however, by dividing the megabyte of memory into a number of segments of 64K bytes each. Each segment begins at an address that yields an even result when divided by 16.

All calculations of memory addresses in the 8086 involve four special registers called segment registers. The 8080 family has a 16-bit address bus that allows addressing of 65,536 bytes of memory. While the internal registers of the 8088/8086 family also have 16 bits, the external address bus has 20 bits. To get the 20-bit address, the 8086 extends a segment register with 4 low-order bits of 0, and adds the segment register to a 16-bit address from another register, as shown in figure 4.

Each segment register defines what is known as its own "current" segment. Each instruction specifies an offset into a segment. The segment registers, which *cannot* be used interchangeably, are as follows:

CS—The Code Segment register defines the 64K-byte current code segment. When an instruction is fetched, the contents of the program counter are added to the CS register contents to calculate the address of the instruction to be fetched.

DS—The Data Segment register defines the current data segment. With three exceptions, all data memory references are understood in relation to the DS register. (The exceptions are that the stack pointer is used to calculate stack addresses, any data memory addresses calculated using the BP register are taken in relation to the stack segment, and any string operations involving the destination are taken in relation to the extra segment. See *SS* and *ES* immediately below.)

SS—The Stack Segment register identifies the current stack segment. References to data memory that use the BP or SP register in calculating the address are understood in relation to the SS register. For example, the PUSH, POP, CALL, INT, and RET instructions use the SS register.

ES—The Extra Segment register plays

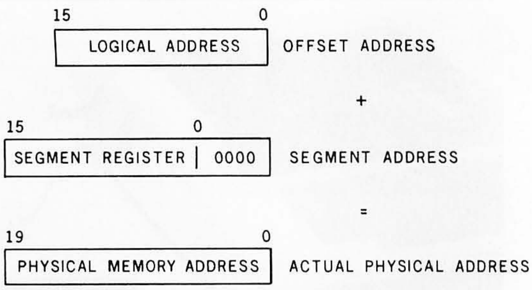

Figure 4: *The 8086 extends a segment register with 4 low-order bits of 0. It then adds the segment to a 16-bit address from another register to achieve a 20-bit address.*

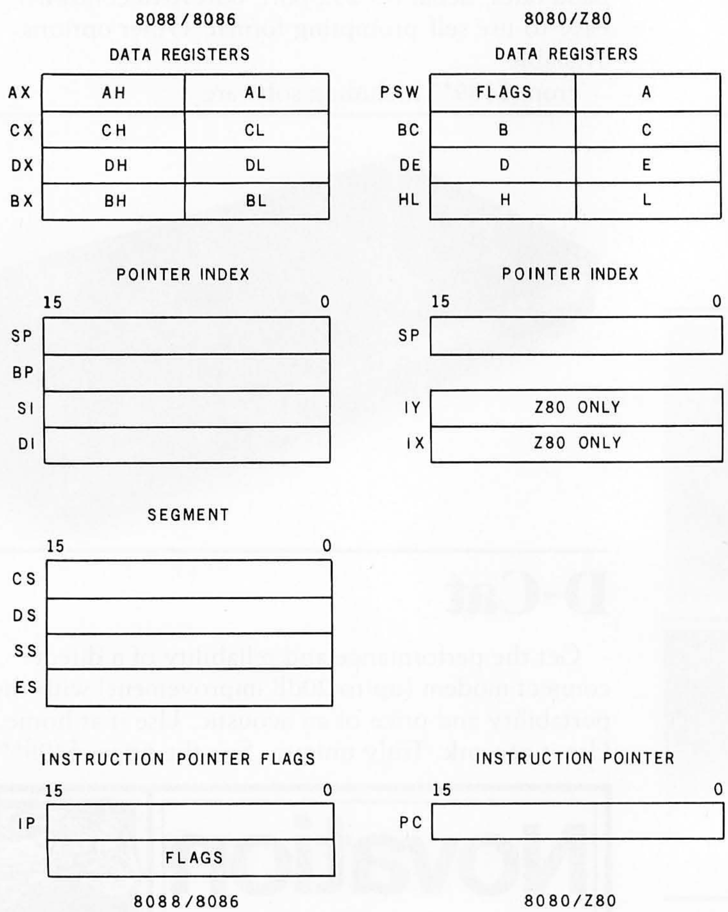

Figure 5: *Register-usage mapping between the two processor families is shown. Note that the IX and IY registers of the Z80 do not exist for the 8080/8085. More registers are in the 8088/8086 than in either the 8080 or Z80. Observe that the 8088/8086 segment registers have no parallel in the 8080/Z80 family. The 8080 family has a 16-bit address bus that allows addressing of 65,536 bytes of memory. While the internal registers of the 8088/8086 family also have 16 bits, the external address bus has 20 bits, formed by extending a segment register with 4 low-order bits of 0, and adding it to a 16-bit address from another register.*

a role in string operations. All destinations of string operations use the DI register in calculating addresses, and are taken relative to the ES register.

Besides the four segment registers, the 8086 has the following:

- Four general-purpose registers: AX, BX, CX, and DX. Each is addressable as a 16-bit register or as two 8-bit registers. When addressed as 8-bit registers, the pairs are called: AH, AL; BX, BL; CH, CL; and DH, DL. The general-purpose registers hold the intermediate results of operations.
- Four pointer and index *registers;* these locate data within a specified segment of memory. SP is the stack pointer, BP the base pointer, SI the source index, and DI the destination index.
- One program counter.
- One 16-bit flag register (program status word, or status register) containing nine flags.

Now that we've looked inside the 8086, we can take a look at how register usage in the 8080/Z80 processor family corresponds generally with that in the 8086. Figure 5 shows the sets of registers alongside each other so that you can see the general correspondence clearly. Note that the IX and IY registers of the Z80 do not exist for the 8080/8085. On the other hand, the alternate register set of the Z80 (not shown) as well as the I and R registers have no parallel in the 8086 register set; operations involving these will require special attention from the programmer after the conversion is done.

Clearly, the 8086 has more registers than either the 8080 or Z80. Since the 8086 also has a more powerful instruction set, translation should be possible with minor restrictions.

Complications

Since all CP/M-80 programs had to exist in a 64K-byte region, there should be little trouble fitting a translated program into a 1-megabyte (1,048,576-byte) region. If you're translating from the Z80, however, things are complicated slightly because its extensive instruction set has to be mapped to the 8086 instructions, as well as the 8080 registers.

The 8086 does, however, have a rich instruction set that covers all the bit instructions and most of the block instructions. The problems encountered in mapping instructions and registers can be formalized and solved by using a variety of software tools.

To convert a source file running under CP/M-80 so as to assemble it with an 8086 assembler requires either the use of a code translator or a considerable effort with a program editor. At least three commercial products come under the translator category:

- XLT86 from Digital Research Inc.
- TRANS86/ACT86 from Sorcim
- The Z80-to-8086 translator from Seattle Computer Products

Intel also has a translator, CONVERT 86, but it is sold only as part of a large software package intended for its OEM customers and doesn't run under CP/M. In addition, some reports indicate that Microsoft may soon bring out a translator.

Two issues must be addressed when converting from a CP/M-80 source:

1. Where is the source coming from?

   Z80 instruction mnemonics?
   (a) Zilog
   (b) TDL (Technical Design Labs)
   (c) Intel (Digital Research with macros)
   (d) Sorcim/ACT-80

   or

   Straight 8080/8085 code with no macros?

2. Where is the resultant code going?

   CP/M-86 or MS-DOS?

The three translators under review do not always approach these issues in the same way. Furthermore, they all handle register mapping somewhat differently. We'll look at the translators now one at a time. Then we'll see how they translated the same CP/M-80 listing. With the listing in hand, we'll see how each of the translators handled register mapping. Finally, we'll turn to the subject of how the two different 16-bit operating systems affect program translation.

XLT86 from Digital Research

Digital Research's XLT86 takes standard 8080 source code in a format compatible with ASM, MAC, or RMAC assemblers and converts the 8080 source code to 8086 source code in a format compatible with *ASM86* operating under either CP/M-80 or CP/M-86. Since XLT86 is written in PL/I-80, the translator can run either stand-alone under CP/M-80 or for cross development under VAX/VMS. It produces optimized 8086 code in a five-phase, multipass process, doing global data-flow analysis to determine optimal register usage.

Although macro definitions are not supported, conditional-assembly directives are. The *XLT86 User's Guide* suggests that if you want macro expansion, you can use a pass through MAC or RMAC to produce a PRN file that can be edited (removing the first few columns of generated hexadecimal code) to produce an expanded source file for input acceptable to XLT86. XLT86 does not recognize Z80 instructions. XLT86 passes repeat loops through to the 8086 source code.

XLT86 analyzes the source program in its entirety, determining the block structure and the register/flag usage. Working from this information, it translates the code to 8086 assembler code in an optimized way. The decision algorithm for each instruction type is given in a section of the manual to allow the user to see what happens in each situation.

Register mapping generally follows the correspondence shown in figure 4, with a loose relationship between the 8086 AX and the 8080 PSW; the exact relationship is determined from register usage at translate time.

Many run-time options are available to control the translation process, both on the command line and embedded in the 8080 source text. The options control the disks that the work and output files are on, whether the block-analysis information is output to disk, whether code and data segments are to be intermixed or kept separate, and whether the condition flags are active on exiting from subroutines.

XLT86 is a sophisticated program that does a reasonable job of optimizing the translation of 8080 source code to 8086 source code. BDOS calls from CP/M-80 are mapped into BDOS calls that are compatible with CP/M-86.

XLT86 has special features for handling translation of conditional JMP and CALL instructions in 8080 source code. In the 8080 instructions, JMP and CALL instructions are capable of reaching any address within the 64K-byte region. The 8086 conditional JMP instructions can reach only 128 bytes on either side of the IP (Instruction Pointer) register. XLT86 examines the target of the conditional JMP. If the target cannot be reached, XLT86 changes the sense of the conditional JMP and skips over a long JMP to the target address. Since there are no conditional CALL or RET instructions in the 8086, the sense of the condition is changed and a short conditional JMP is performed

The 8086 can access up to 1 megabyte of memory and as many as 65,000 input/output ports.

to skip over an unconditional CALL or RET.

As noted earlier, the segment registers allow for separation of code and data regions. To reference data, you have to tell the 8086 whether data is in the code segment (CS) or the data segment (DS). For the Digital Research ASM86 assembler, the Offset directive handles this chore. XLT86 examines an expression and determines the proper segment for the particular instruction.

XLT86 does have limits on the size of the 8080 source files that it can translate because the flow-analysis information must be in memory. In a 64K-byte CP/M system, the maximum source file that can be translated is approximately 6K bytes, depending on the structure of the program. Nothing is said in the manual about being able to deal with modular code using RMAC and external references. This implies that the entire source program must be converted at once, limiting the size of the program that can be translated by using XLT86 to 6K bytes.

In summary, if you're starting from 8080/8085 assembly code written for ASM or MAC and you want to go to CP/M-86, and if the source program does not exceed 6K bytes, XLT86 is the most useful translator. Code written for Z80s using MAC requires careful examination after the translation process to make sure that no flags were inadvertently changed. You have to expand macros before using XLT86 unless the number of invocations was small. In that event, expanding the macros by hand with an editor might be just as easy.

Sorcim's TRANS86

TRANS86 is an 8080/ACT80-to-8086 translator. It takes 8080 or ACT80 source code as input and creates a file compatible with the input to ACT86, an 8086 assembler.

The output of TRANS86 is incompatible with any assembler other than ACT86. The ACT86 mnemonics are different enough so that, unless the programmer has a sophisticated text processor and the talent (or patience) to do a great deal of text manipulation, TRANS86 should be used only with ACT86.

Both TRANS86 and ACT86 run on either 8080 or Z80 processors under CP/M-80, MP/M, or CDOS with a minimum of 24K bytes of RAM (random-access read/write memory). TRANS86 consists of an executable file, an overlay file, and a translation table. The input assembly source code must be in a form acceptable to the standard CP/M-80 assemblers (ASM, MAC, or RMAC), or to ACT80, Sorcim's Assembly Code Translator for 8080/Z80 processors.

Translation occurs on an instruction-by-instruction basis with some optimization rules applied to conditional jumps. There appears to be no limit as to the size of the source file that can be translated. A file is produced on the same disk as the source file with the same name and an .ASN extension. If a file by that name already exists, the user is asked if the file should be deleted or if the program should be aborted.

TRANS86 flags the following Z80 instructions as errors:

ACT80 code	Zilog/Mostek equivalent code
Mov A,R	Ld A,R
Mov R,A	Ld R,A
Mov A,I	Ld A,I
Mov I,A	Ld I,A
In,C reg	In reg,(C)
Inir	Inir
Otir	Otir
Rld	Rld
Rrd	Rrd

TRANS86 supports macros in the ACT80 format. Although TRANS86 acts on macros in the MAC format, there is no guarantee that the macros will expand correctly. The user is cautioned to examine the result of a macro expansion to determine if the sense of the macro has been maintained. Examples are given of some macros that work and some that do not. The *TRANS86 User's Reference Manual* includes a section that gives hints on how to hand-optimize the output of TRANS86; specifically, accumulator indirect loads through the DE and BC registers, 8080 conditional jumps, and Z80 block instructions.

Another section describes the differences between ACT80 and Z80 mnemonics. This information allows the programmer to manually convert assembly source code to a form acceptable to TRANS86. The ACT80 instruction set has some ASM-style instructions, some Z80-style instructions, and some instructions that are unique to the ACT80 assembler. If the source code is written in 8080 ASM mnemonics, TRANS86 will process it and output ACT86 assembler code. The 8080 instruction SPHL, however, was translated incorrectly in the current version of TRANS86.

Another section in the manual contains suggested constructions that can be manually entered to deal with Z80 op codes that are flagged as errors. Block input/output instructions and input/output through register C are described in detail.

Seattle Computer System

Seattle Computer Products' (SCP) translation system consists of two programs: a Z80 translator (called TRANS86 on the disk, though *not* the same as Sorcim's) and a compatible 8086 cross assembler (ASM86). Both programs run on Z80 processors under CP/M-80 with a minimum of 24K bytes of RAM. The programs do not run on 8080/8085 processors.

The translator accepts source files in Zilog/Mostek mnemonics and produces an 8086 assembler source file in a form acceptable to the SCP 8086 cross assembler. Since the 8086 source format required by the SCP 8086 cross assembler is different from any other 8086 assembler that we know of, you must use these two programs together. The translator places its output on the same disk as the Z80 source code and gives it an .A86 file extension.

The translation is on an instruction-by-instruction basis with no optimization. There appears to be no limit on the file size that may be translated. Not all Z80 instructions are translated, however. Those in the following list will produce an op code error:

Cpd	Ldi
Cpi	Otdr
Ind	Otir
Indr	Outd
Ini	Outi
Inir	Rld
Ldd	Rrd

These op codes are mostly in the block-manipulation set of instructions. Although programmers do use these instructions, they must be manually coded when converting to the 8086. The SCP translator does not support macros and permits use of the following pseudo-ops only:

Db
Dm
Ds
Dw
Equ
If/Endif
Org

If the Z80 index registers IX and IY are used, they are mapped into memory locations with the labels IX: and IY:. The programmer has to define these locations; otherwise, they will show up on the assembly listing as undefined labels. The Z80 alternate register set (BC',DE',HL') is treated the same way, as memory locations that the programmer must define.

Either the DI or SI register can be used as a temporary IX register by loading one of them from the location IX when required to do indexed instructions. The programmer has to take care of this substitution; the translator does not.

When using the DI register, you must always keep in mind that the only 8086 segment base that can be used with the DI register is the ES segment; the SI register, on the other hand, can reference all the segment bases, defaulting to the DS segment. If this 8086 source code is going to be run under CP/M-86, you have to be careful about using the ES segment register. The CP/M-86 documentation specifically states that ES is not saved through a BDOS call.

For the SCP Z80-to-8086 translator's register usage, see figure 6b.

Translating the Test File

To determine how the three programs actually translate source text, we prepared a file acceptable as input to an assembler and containing all the op codes of the 8080 and Z80. Since the SCP translator could accept only Zilog/Mostek mnemonics, the test text was run through an 8080-to-Z80 filter program before the translation.

Listing 1 presents, side by side and line by line, the original 8080 code, the Sorcim TRANS86 translation, the Seattle translation, and the XLT86 translation. Here are reminders of some things to consider when you examine the translations:

- Because of the differences in the architectures of the 8086 and the 8080/Z80, some choices must be made when translating from one architecture to the other. Therefore, some difference in translation is to be expected.

(a)

	8080/8085	TRANS86	8086/8088 Seattle	XLT86
8-bit registers	A	AL	AL	AL
	B	CH	CH	CH
	C	CL	CL	CL
	D	DH	DH	DH
	E	DL	DL	DL
	H	BH	BH	BH
	L	BL	BL	BL
16-bit register pairs	PSW	AX(1)	AX	AX
	BC	CX	CX(2)	CX(2)
	DE	DX	DX(2)	DX(2)
	HL	BX	BX	BX
16-bit register pairs	SP	SP	SP	SP
	PC	IP	IP	IP

(1) TRANS86 does not preserve 8080 byte order on the stack.
(2) The Seattle translator uses SI on loads from memory and DI for stores to memory. TRANS86 and XLT86 do a register exchange between BX and the appropriate register to allow indirect addressing through BX, then a register exchange to fix up BX and the appropriate register.

(b)

	Z80	TRANS86	8086/8088 Seattle	XLT86
8-bit	R	(3)	(3)	(3)
	I	(3)	(3)	(3)
16-bit	IX	DI	(4)	(5)
	IY	SI	(4)	(5)
alternate registers set	BC'	(5)	(4)	(5)
	DE'	(5)	(4)	(5)
	HL'	(5)	(4)	(5)

(3) Since the 8086 does not have equivalent registers, none of the translators support these registers. However, they can be mapped to a memory location by the programmer.
(4) Seattle's TRANS86 handles these registers by generating memory references to storage locations defined by the programmer.
(5) Although these registers are not mapped by the translators, the programmer can define storage locations and deal with them through macro definitions.

Figure 6: *8080/8085/Z80-to-8086/8088 register mapping. Figure 6a shows 8080/8085-to-8086 register mapping by the three translator programs. Figure 6b shows Z80-to-8086 register mapping by the three translator programs.*

- In general, when the 8080 does 16-bit arithmetic, only the carry bit is affected; this is definitely not so in the 8086.
- The Z80 and 8086 do string and block operations differently; the 8080 has no primitive block operations at all.
- As noted earlier, the segment registers in the 8086 allow addressing of up to 1 megabyte; no corresponding registers exist in either the 8080 or Z80.
- Registers used for indirect memory references in the 8080/Z80 are different from the corresponding mapped registers in the 8086.
- Conditional jumps in the 8080/Z80 can reach anywhere in its address space; conditional jumps in the 8086 can reach only 128 bytes on either side relative to the IP register.
- No conditional calls in the 8086 correspond to the conditional calls of the 8080/Z80.

Listing 1 makes it apparent that the three translators treat most instructions the same way, allowing for the differences in the target instruction set. The following comments highlight the differences found.

The only incorrect translation is TRANS86's rendering of the SPHL instruction. The transfer is in the wrong direction. The comment field of the instruction was wrong in the

Listing 1: *The original 8080 code, the Sorcim TRANS86 translation, the Seattle translation, and the XLT86 translation.*

Original 8080 code	TRANS86	Seattle	XLT86
Stax B	Xchg BX,CX Sto AL,[BX] Xchg BX,CX	Mov DI,CX Stob	Mov SI,CX Mov [SI],AL
Inx B	Inc CX	Lahf Inc CX Sahf	Lahf Inc CX Sahf
Dad B	Add BX,CX	Lahf Add BX,CX Rcr SI Sahf Rcl SI	Lahf Add BX,CX Sahf
Ldax B	Xchg BX,CX Ld AL,[BX] Xchg BX,CX	Mov SI,CX Lodb	Mov SI,CX Mov AL,[SI]
Stax D	Xchg BX,DX Sto AL,[BX] Xchg BX,DX	Mov DI,DX Stob	Mov SI,DX Mov [SI],AL
Ldax D	Xchg BX,DX Ld AL,[BX] Xchg BX,DX	Mov SI,DX Lodb	Mov SI,DX Mov AL,[SI]
Dad H	Add BX,BX	Lahf Add BX,BX Rcr SI Sahf Rcl SI	SHL BX,1
Dec H	Dec BX	Lahf Dec BX Sahf	Dec BX
Inr M	Incb [BX]	Inc B,[BX]	Inc Byte Ptr 0[BX]
Mvi M,3	Stob #3,[BX]	Mov B,[BX],3	Mov Byte Ptr 0[BX],3
Dad SP	Add BX,SP	Lahf Add BX,SP Rcr SI Sahf Rcl SI	Add BX,SP
Dcx SP	Dec SP	Lahf Dec SP Sahf	Dec SP
Mov M,A	Sto AL,[BX]	Mov [BX],AL	Mov Byte Ptr 0[BX],AL
Rnz	Jz :G1 Ret :G1:	Jnz Ret	Jz L 1 Ret L 1:
Jnz Help	Jz :G2 Jmp Help :G2:	Jnz Help	Jz L 2 Jmp Help L 2:

Listing 1 continued on page 340

Listing 1 continued:

```
Cnz Help        Jz :G3          Jz X000         Jz L 3
                  Call Help      Call Help       Call Help
                :G3:            X000:           L 3:

Xthl            Mov BP,SP       Pop SI          Mov BP,SP
                Xchg BX,[BP]    Xchg SI,BX      Xchg BX,[BP]
                                Push SI

Pop PSW         Pop AX          Pop AX          Pop AX
                Sahf            Xchg AH,AL      Xchg AL,AH
                                Sahf            Sahf

Push PSW        Lahf            Lahf            Lahf
                Push AX         Xchg AH,AL      Xchg AL,AH
                                Push AX         Push AX
                                Xchg AH,AL      Xchg AL,AH

Sphl            Mov BX,SP       Mov SP,BX       Mov SP,BX

(Z80 opcodes)

Djnz Label      Dec CX          Dec CH          -------
                Jnz Label       Jnz Label

Cpir            Xchg BX,DI      UP              -------
                Cld             Mov DI,BX
                Repnz           Repnz
                Scasb           Scab
                Xchg BX,DI      Mov BX,DI

Ldir            Xchg BX,SI      Up              -------
                Xchg DX,DI      Mov SI,BX
                Cld             Mov DI,DX
                Repnz           Rep
                Movsb           Movb
                Xchg BX,SI      Mov DX,DI
                Xchg DX,DI      Mov BX,SI

Push Ix         Push DI         Push [Ix]       -------

Pop  Iy         Pop  SI         Pop [Iy]        -------
```

translated text, also. We therefore suspect that the program has a minor bug with regard to this particular instruction.

Since this test program was nonsense as far as logical program flow is concerned, XLT86 was at a loss to determine the active registers, and sometimes chose simpler instructions than it would have in a real program. The results for XLT86 could be different in different situations.

The conditional jumps for TRANS86 and XLT86 can vary, depending on the distance of the target label from where the jump is.

The conditional return in Seattle's translator references a label called RET. This refers to any RET within 128 bytes on either side of the statement. This is one reason why Seattle's translator should be used with ASM86; no other assembler will take advantage of this feature.

Note that TRANS86 and the Seattle translator treat the DJNZ instruction differently. TRANS86 uses a 16-bit register, CX, and the Seattle translator uses CH, an 8-bit register. A warning message comes out of the Seattle translator reminding the programmer that DJNZ does not affect the flags in the Z80 but that this sequence of instructions will affect the 8086 flags.

Register Mapping

Figure 6 shows a detailed, side-by-side comparison of the differences in register mapping performed by the three translators. Figure 6a deals with the 8080/8085-to-8086 mapping; figure 6b, with Z80-to-8086 mapping.

As the notes there state, TRANS86 does not preserve 8080 byte order on the stack.

The Seattle translator uses SI on loads from memory and DI for stores to memory.

TRANS86 and XLT86 do a register exchange between BX and the appropriate register to allow indirect addressing through BX, then a register exchange to fix up BX and the appropriate register.

Since the 8086 does not have some of the registers of the Z80, the translators can't support them. The programmer can, however, map those registers to a memory location.

TRANS86 generates memory references to storage locations supplied by the programmer to take care of the Z80's IX, IY, BC', DE', and HL' registers.

Summing Up the Translators

A general view is that Sorcim's TRANS86 is a useful product if the original source is in 8080 or ACT80 form and the user has ACT86 as a target 8086 assembler. The register and flag usage appear to be a little looser than for the other two programs. This requires more knowledge and more involvement from the programmer to make sure that the sense of the translated code is maintained. No limitations exist as to the size of the source file and macros are supported if the input is in ACT80 format. Sorcim's TRANS86 is sold separately from ACT86, but they should be used together.

The Seattle Computer Products' Z80-to-8086 translator is a straightforward code translator that uses Zilog mnemonics and runs only on Z80-based processors. There appear to be no limitations as to the size of the source program that may be translated since the program translates one instruction at a time. Register and flag usage are very conservative, protecting the source architecture as much as possible and providing warnings when potential problems could arise. The converted program has more of a chance of working the first time than a less conservative translation would have.

Once the program is working correctly, the programmer can reedit the 8086 source to trim out the unnecessary register saves and flag manipulations to speed up the code and make it smaller. If the source code is in Zilog mnemonics, or if a filter program is available to convert to Zilog mnemonics, we recommend the pair of programs from Seattle Computer Products. The Seattle translator and its ASM86 assembler should be used together for optimum benefit.

XLT86 from Digital Research is a program that takes 8080 source code and converts it into 8086 source code in an intelligent manner using data-flow-analysis techniques. It will produce better code than either of the other two translators if two conditions are met: (1) no Z80 instructions are used, and (2) the source code is not bigger than 6K bytes (assuming the CP/M system is a 64K-byte system). XLT86 should be used with ASM86 under either CP/M-80 or CP/M-86.

MS-DOS versus CP/M-86

In the trade press and various advertisements, we see claims that conversion of CP/M-80 programs to MS-DOS (IBM PC-DOS, Lifeboat Associates' SB-86, Seattle Computer Products' 86-DOS) is as easy, or even easier, than from CP/M-80 to CP/M-86. In fact, some differences are seen in interfacing to the two 16-bit operating systems. With the assumption that the program had to operate in a 64K-byte region under CP/M-80, we will bypass memory-allocation questions. The remaining issues that have to be addressed are:

- How does a program gain access to operating system resources?
- How are file control blocks used to manipulate files?

Gaining access to the operating system under CP/M-80 requires placing the function code in the C register, placing the information address in the DE registers, and calling location 05 hexadecimal, the CP/M-80 entry point. If the system call returns a value, it returns the value to the A register. A so-called warm boot under CP/M-80 is accessed by jumping to location 00 hexadecimal, which reads in the operating system and resets the disk system.

Now let's look at similar functions under CP/M-86 and MS-DOS.

Gaining Access to CP/M-86

Gaining access to CP/M-86 requires placing the function code in the CL register, placing the byte parameter in the DL register or placing the word parameter in the DX register, placing the data segment in the DS register (the data segment is usually not changed for a converted program), and executing a software interrupt, INT #224. The result is returned in the AL register if it is a byte value; if the result is a word value, it is returned in both the AX and BX registers. Double-word values are returned with the offset in the BX registers and the segment in the ES register. Conversion of programs from CP/M-80 to CP/M-86, then, requires replacing the call to location 5 with the software interrupt INT #224.

Another necessary change involves the warm boot. Under CP/M-80, the warm boot may be accessed by a system call with a function code of 0 for a jump to location 0. CP/M-86, however, does not support the jump to location 0. As a result, you must change this program exit in the translated program if the program is to run correctly.

Provided that the call to location 5 is replaced with INT #224, that the warm boot change is made, and that the registers are mapped correctly, there should be little problem in getting the translated program to access the CP/M-86 system functions.

Gaining Access to MS-DOS

Although MS-DOS has a "preferred" mechanism through a software interrupt, INT #33, for accessing the system, an additional mechanism is provided for "preexisting" programs that is compatible with CP/M-80 calling conventions, at least for functions in the range of

0-36. As far as system calls within the allowed function range are concerned, the programmer doesn't have to do anything to translated programs to get them to run under MS-DOS other than to correctly map the registers.

MS-DOS also supports the warm boot function of CP/M-80. A jump to location 0 under MS-DOS executes a software interrupt, INT #32, which is functionally a program end and the normal way to exit from a program.

Manipulating File Control Blocks

The file control block used in CP/M-80 consists of a 36-byte block, which describes the disk drive on which to find or create the file, the file name, and information relating to which record of the file is desired.

At least so far as normal file-access requests are concerned, both MS-DOS and CP/M-86 treat this block of information the same.

System-level information is quite different in the two cases, and programs that look at system bytes within the file control block need to be changed for MS-DOS to function correctly. The MS-DOS file control block has many more features, including the date the file was created or last updated, the logical record size, and the file size. These system-information bytes are in areas within the file control block that application programs normally do not access. Nevertheless, converting programs to make use of MS-DOS file control blocks should take little effort.

The problems encountered in mapping instructions and registers can be formalized and solved by using a variety of software tools.

Conclusion

There is, in fact, little if any difference in the difficulty of translating sound CP/M-80 programs to CP/M-86 or MS-DOS. With CP/M-86, the programmer will have to make minor changes to gain access to the operating system. With MS-DOS, the programmer will have to make minor changes to handle the extra features of the MS-DOS file control blocks.

Next month, we will make further comparisons between MS-DOS and CP/M-86. We will include some benchmarks made with the Compupro 8085/8088 dual-processor S-100 system. We will report not only the results of running programs under both CP/M-86 and MS-DOS on the same 8088 in the same machine, but also the results of running the same programs under MS-DOS running Emulator-86 on the same 8088 in the same machine. Although that may sound more like a cat chasing its own tail than a test of operating systems, we will try to keep it all straight. ■