

Byte Code Interpreter for 8051 Microcontroller

Suchart Khuntaweethep¹ and Narakorn Jeenjun^{2*}

1. Introduction

To use microcontroller, user has to understand its architecture and language. Microcontroller of each family has its own architecture and instruction sets. For example PIC family is RISC and 8051 family is CISC. When user wants to change microcontroller to the other, user has to develop the new program to support new microcontroller. It is waste of time because user has to study architecture and instruction sets of new microcontroller as well.

Interpreter for microcontroller may be the alternative way to solve the problem. When user wants to change the new microcontroller, user only installs interpreter to new microcontroller but do not write the new program.

Example for byte code is Java language. The Java as introduced by Sun Microsystems [1] in 1994 has spread throughout the computer industry and has reached all domains. As good as Java is for providing “write once, run everywhere” software. In figure 1, programmer writes Java program. Java compiler will compile Java program to Java byte code. To use Java program, user has to install Java virtual Machine [2] to operating system target. In case user wants to change to new operating system,

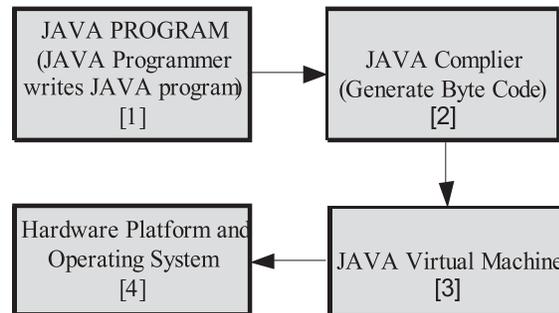


Figure 1 Java operation.

users just installs Java virtual Machine and Java program can be ran as normal. For this case Java virtual Machine is interpreter of Java system.

The favorite toy in the world, Lego mind storm [3]. It has H8 microcontroller with Lego interpreter inside. User interface is based on graphic block. It is easy for beginner even children can develop their own applications.

For the other examples of interpreter are MetaCricket [4] developed by MIT, GOGO BOARD [5], BasicStamp by Parallax Inc. [6], picoJava-I [7] picoJava-II [8] by Sun Microsystems, An Embedded Java Virtual Machine [9], [10], The Byte code Firmware Design for Microcontroller Device [11].

¹ Associate Professr, Department of Electronics, Faculty of Engineering King Mongkut’s Institute of Technology Ladkrabang.

² Student, Department of Electronics, Faculty of Engineering King Mongkut’s Institute of Technology Ladkrabang.

* Corresponding Author, Tel.08-7700-7502, E-mail: narakorn.jeenjun@spansion.com

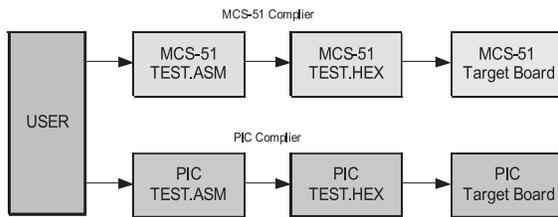


Figure 2 Standard MCU.

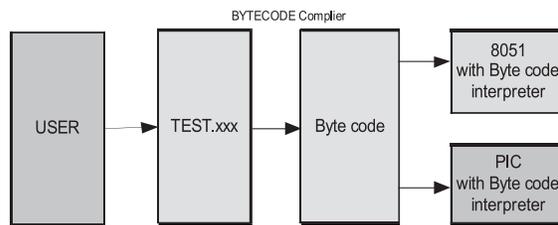


Figure 3 MCU with byte code interpreter.

2. Comparison

Standard Microcontroller (MCU) language design is depicted in figure 2, user writes assembly program to target MCU and compiles to machine code of each system. In this case, user has to develop 2 programs for both 8051 and PIC.

In figure 3, user installs byte code interpreter into both 8051 and PIC microcontroller then writes the program with byte code instruction sets. In this case user writes the program only 1 time.

3. Interpreter Diagram

The interpreter for MCU has 3 parts, they are depicted in figure 4.

- VM Controller is the core of interpreter, it controls fetch/execute and condition commands, IF, ELSE, REPEAT and GOTO.
- VM EXECUTE is byte code decoder.
- Peripheral is hardware modules. It consists of, Parallel Port, Serial Port, I²C Port, Timer, LCD Port and Arithmetic functions.

The system flow chart is depicted in figure 5.

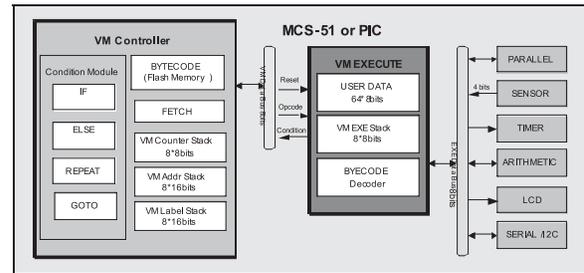


Figure 4 Interpreter diagram.

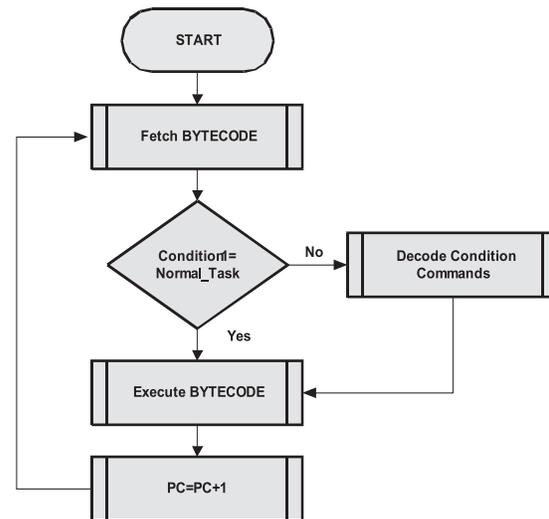


Figure 5 Interpreter flow chart.

After MCU is reset, interpreter fetches byte code from built in Flash memory. Condition1 will be checked, whatever it is normal task or condition task. If it is normal task, next byte code will be decoded as normal. If not, interpreter will check condition before go to next command. After the operation is done, program counter (PC) will be increased to next line address.

3.1 Data Stack

Interpreter is designed to have 16 levels of data stack. The advantage of data stack is, store the result of condition command IF/ELSE and WAITUNITL and the result from arithmetic commands.

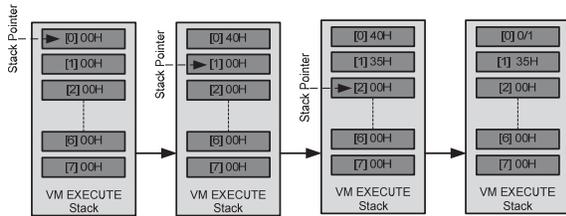


Figure 6 Data stack.

Data stack example is depicted in figure 6
 00,02,09,01,00,06,00,01,40,01,35,1C,0A

- 00,02 : Line address of program
- 09 : SUM
- 01 : Condition1 (IF statement)
- 00 : No Condition2
- 06 : Number slave of IF statement
- 00 : This is master line
- 01,40 : Push data 40H to current stack
- 01,35 : Push data 35 to next stack
- 1C : Equal comparison byte code
- 0A : Comparison data in stack and stack+1,

the result will be in stack

3.2 The Byte Code Format

The byte code format consists of 7 parts, it is depicted in figure 7

- AddrH and AddrL are 16 bits line address. (Maximum is 65,535 lines)
- SUM is used byte in each line (AddrL, AddrH and SUM are not included)
- Condition1 is the primary condition.
- Condition2 is the secondary condition.
- CondListH, in case if Condition1 is not 00H
 - Equal 00H : It is slave line address.
 - Not equal 00H : It is master line address.
- CondListL, in case if Condition1 is not 00H
 - Equal 00H : It is master line address.
 - Not equal 00H : It is slave line address.



Figure 7 Byte code format.



Figure 8 No condition command format.



Figure 9 Condition command format.

- Byte code is 8 bits command.

3.3 Byte Code Instruction Sets

The instruction sets are categorized to 2 sections.

3.3.1 No Condition Commands

For no condition commands, Condition1 will be 00H, Condition2 CondListH and CondListL will not be appeared in the line. Format is depicted in figure 8. Example for this commands are TX232, RX232, I²CSTART, I²CWRITE, LCD, WAITTIME, OUTPORT and INPORT.

3.3.2 Condition Commands

For condition commands consist of, IF, ELSE, REPEAT and GOTO. The format is depicted in figure 9. Condition1 will not be 00H and Condition2 is 00H reserved for further use.

Interpreter is designed to have 8 levels of IF and ELSE commands. In figure 10 shows 3 levels of IF command and 2 levels of ELSE command.

For IF command, byte code 01H-08H are reserved and 81H-88H are reserved for end of IF line. ELSE command, byte code 09H-0FH are reserved and 89H-8FH are reserved for end of ELSE line.

For REPEAT command, interpreter is designed to have 8 levels. Number of repeating can be 0-255. Byte code 21H-28H is reserved for REPEAT command.



User Code	BYTECODE
LABEL:	;00,00,06,64,00,00,05,00,01,
RX232	;00,01,02,00,14,
TX232 (USERTEMP)	;00,02,03,00,04,13,
GOTO LABEL	;00,03,05,65,00,00,05,00,

Figure 14 Example program#2.

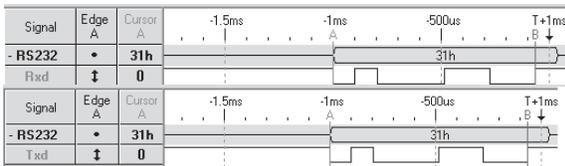


Figure 15 TX232 and RX232.

- 00, 01 : Line address
 - 04 : SUM
 - 00 : Condition1 (No condition command)
 - 01,5A : Push data 5AH to data stack
 - 4A : Send data in data stack to I²C
- The third line 00, 02, 02, 00, 4F
- 00, 02 : Line address
 - 02 : SUM
 - 00 : Condition1 (No condition command)
 - 4F : Write stop I²C signal

The second example is depicted in figure 14. The program receives data from RS232 channel and sends the same data back to host (Personal computer). Figure 15 shows RS232 signals.

The first line 00, 00, 06, 64, 00, 00, 05, 00, 01

- 00, 00 : Line address
 - 06 : SUM
 - 64 : Condition1 (Label command)
 - 00 : No Condition2
 - 00,05 : Label stack location
 - 00,01 : Line address for go to command
- The second line 00, 01, 02, 00, 14

- 00, 01 : Line address
- 02 : SUM
- 00 : Condition1 (No condition command)

User Code	Byte Code
LABEL:	;00,00,06,64,00,00,05,00,01,
OUTPUT(0x01)	;00,01,04,00,01,01,4C,
WAITTIME(4)	;00,02,04,00,01,04,10,
OUTPUT(0x02)	;00,03,04,00,01,02,4C,
WAITTIME(4)	;00,04,04,00,01,04,10,
OUTPUT(0x04)	;00,05,04,00,01,04,4C,
WAITTIME(4)	;00,06,04,00,01,04,10,
OUTPUT(0x08)	;00,07,04,00,01,08,4C,
WAITTIME(4)	;00,08,04,00,01,04,10,
OUTPUT(0x10)	;00,09,04,00,01,10,4C,
WAITTIME(4)	;00,0A,04,00,01,04,10,
OUTPUT(0x20)	;00,0B,04,00,01,20,4C,
WAITTIME(4)	;00,0C,04,00,01,04,10,
OUTPUT(0x40)	;00,0D,04,00,01,40,4C,
WAITTIME(4)	;00,0E,04,00,01,04,10,
OUTPUT(0x80)	;00,0F,04,00,01,80,4C,
WAITTIME(4)	;00,10,04,00,01,04,10,
GOTO LABEL	;00,11,05,65,00,00,05,00,

Figure 16 Example program#3.

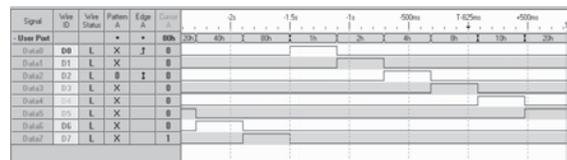


Figure 17 Parallel Port.

- 14 : Read data from RS232
- The third line 00, 02, 02, 00, 04, 13
- 00, 02 : Line address
 - 02 : SUM
 - 00 : Condition1 (No condition command)
 - 04 : Push RS232 data to data stack
 - 13 : Write data from data stack to RS232
- The fourth line 00, 03, 05, 65, 00, 00, 05, 00
- 00, 03 : Line address
 - 05 : SUM
 - 65 : Condition1 (Go to command)
 - 00 : No Condition2
 - 00, 05 : Reload Label at stack location 05
 - 00 : Reserved for further used

The third example is depicted in figure 16, the program will send data to parallel port with data 1, 2, 4, 8, 10H, 20H, 40H and 80H with 400mS delay time. Testing result is depicted in figure 17.

- The first line 00, 00, 06, 64, 00, 00, 05, 00, 01
- 00, 00 : Line address
 - 06 : SUM

```
[0] lcdinit
[1-2] lcd("12345678 abcdefgh")
[3] loop:
[4] goto loop;

[0] 00,00,02,00,A2,
[1] 00,01,19,00,01,31,A0,01,32,A0,01,33,A0,01,34,A0,01,35,A0,01,36,A0,01,37,A0,01,38,A0,
[2] 00,01,1A,00,A5,01,41,A0,01,62,A0,01,63,A0,01,64,A0,01,65,A0,01,66,A0,01,67,A0,01,68,A0,
[3] 00,03,06,64,00,00,05,00,01,
[4] 00,04,05,65,00,00,05,00
```

Figure 18 LCD program.

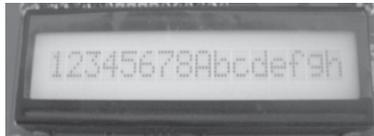


Figure 19 LCD output.

- 64 : Condition1 (Label command)
- 00 : No Condition2
- 00, 05 : Label stack location

The second line 00, 01, 04, 00, 01, 01, 4C

- 00, 01 : Line address
- 04 : SUM
- 00 : Condition1 (No condition command)
- 01, 01 : Push data 01 to data stack
- 4C : Write data in data stack to parallel port

The third line 00, 02, 04, 00, 01,04, 10

- 00, 02 : Line address
- 04 : SUM
- 00 : Condition1 (No condition command)
- 01, 04 : Push data 01 to data stack
- 10 : Delay time with data in data stack

multiply by 100mS

The fourth line 00, 03, 04, 00, 01, 02, 4C

- 00, 03 : Line address
- 04 : SUM
- 00 : Condition1 (No condition command)
- 01, 02 : Push data 02 to data stack
- 4C : Write data in data stack to parallel port

The rest program will be repeated the same as line1 and line2 but only change data out.

The last example is LCD command, figure 18 is program and figure 19 is testing result.

The first line 00, 00, 02, 00, A2

- 00, 00 : Line address
- 02 : SUM
- 00 : Condition1 (No condition command)
- A2 : Initial LCD byte code command

The second line 00, 01, 19, 00, 01, 31, A0, 01, 32, A0, 01, 33, A0, 01, 34, A0, 01, 35, A0, 01, 36, A0, 01, 37, A0, 01, 38, A0,

- 00, 01 : Line address
- 19 : SUM
- 00 : Condition1 (No condition command)
- 01, 31 : Push data 31(character "1") to data stack
- A0 : Write data in data stack to LCD
- 01, 32 : Push data 32(character "2") to data stack
- A0 : Write data in data stack to LCD
- 01, 33 : Push data 33(character "3") to data stack
- A0 : Write data in data stack to LCD
- 01, 34 : Push data 34(character "4") to data stack
- A0 : Write data in data stack to LCD

The rest program will write data "5678 abcdefgh" to LCD.

4.1 Result Analysis

4.1.1 Memory Usage

This paper we have installed byte code interpreter into 8051(CISC) microcontroller. The interpreter is written by C language with KEIL compiler. Table 1 is comparison for memory usage.

Table 1 Memory usage

MCU	IDE	Code	Data
AT89C5131	KEIL	10.97KB	434B
PIC18F4620	MPLAB	14.96KB	450B



4.1.2 Speed Improvement

From the previous paper, the byte code firmware for Microcontroller Device [11], user program was stored in external memory EEPROM with I²C protocol. It is the bottle neck of the system because it took 10 clocks to fetch data from EEPROM 1 Byte. In this paper we have improved the problem by use built in Flash memory of the microcontroller, it took only 2 clocks to fetch data 1 Byte.

5. Conclusion

According to testing results, even we have changed system from PIC to 8051 microcontroller, the interpreter still can execute byte code commands correctly. It is proved that the hardware is independent. The execution byte code command is improved by changing external memory to internal Flash memory. Further to reduce interpreter code size and speed improvement, we can use Assembly language instead of C language. Finally to check byte code syntax, the compiler should be developed as well.

Reference

- [1] S. Wilson and J. Kesselman, *Java platform performance strategies and tactics*, Addison-Wesley, Boston, 2000.
- [2] T. Lindholm and F. Yellin, *The Java virtual machine specification*, Addison-Wesley, Reading, Mass., 1997.
- [3] Jin Sato, *LEGO MINDSTORMS: The Master's Technique*, O'Reilly Media, Inc., 2008.
- [4] F. Matin, B. Mikhak and B. Silverman, "MetaCricket: A designer's kit for making computational devices," *IBM Systems Journal*, vol. 39, no. 3&4, 2000.
- [5] Arnan (Roger) Sipitakiat, "GOGO BOARD," Available from: <http://www.gogoboard.org>
- [6] Al Williams, "Microcontroller Projects Using The Basic Stamp," *CMP Books*, 2002.
- [7] Sun Microsystems, "picoJava-I: picoJava-I Core Microprocessor Architecture," *Sun Microsystems white paper*, October 1996.
- [8] Sun Microsystems, "picoJava-II: Java Processor Core," *Sun Microsystems data sheet*, April 1998.
- [9] Graham Mathias and Kenneth B. Kent "An Embedded Java Virtual Machine Using Network-on-Chip Design," *IEEE International Workshop on Rapid System Prototyping*, 2006.
- [10] Kenneth B. Kent and Micaela Serra, "Hardware/Software Co-Design of a Java Virtual Machine," *IEEE*, 2000.
- [11] Narakorn Jeenjun, "The Byte code Firmware Design for Microcontroller Device," *KKU Journal*, vol. 34, pp. 535-546, 2007.
- [12] ETT-JR51USB Board Available from: www.ett.co.th

8051 microcontroller is a 8 bit microcontroller which has 128 bytes of on chip RAM, 4K bytes of on chip ROM, two timers, one serial port and four 8bit ports. 8052 microcontroller is an extension of microcontroller. The table below shows the comparison of 8051 family members. The code is split into small meaningful chunks and explained below. For 16*2 LCD interfacing with 8051 microcontroller, we have to define pins on which 16*2 lcd is connected to 8051 microcontroller. RS pin of 16*2 lcd is connected to P3.7, RW pin of 16*2 lcd is connected to P3.6 and E pin of 16*2 lcd is connected to P3.5. Data pins are connected to port 1 of 8051 microcontroller. `sbit rs=P3^7;`
`sbit rw=P3^6;`
`sbit en=P3^5;` Next we have to define some functions which are used in the program. Microcontroller Basics. Interpreters and compilers are two forms of higher-level languages. An interpreter translates a program into machine code each time the program runs, while a compiler translates only once, creating a new, executable file that the computer runs directly, without re-trans-lating. As a rule, interpreters are very convenient for shorter programs where execution speed isn't critical. It controls an external latch that stores the lower address byte during memory accesses. When the 8052 reads or writes to external memory, it places the lower address byte on AD0-AD7 and strobes ALE, which causes the external latch to save the lower address byte for the rest of the read or write cycle. 8051 Microcontroller port programming. There are four ports P0, P1, P2 and P3 each use 8 pins, making them 8-bit ports. All the ports upon RESET are configured as output, ready to be used as output ports. To use any of these ports as an input port, it must be programmed. For example, the following code will send out continuously to port 2 the alternating values 55h and AAH. That is all the bits of port 2 toggle continuously. `MOV A,#55H ; A = 55 hex`
`BACK: MOV P2,A ;send it to Port 2`
`ACALL DELAY ;call delay routine`
`CPL A ;make A=0`
`SJMP BACK.` Since an 8031 is capable of accessing 64K bytes of external memory, it needs a path for the 16 bits of the address. While P0 provides the lower 8 bits via A0-A7, it is the job of P2 to provide bits A8-A15 of the address.