

RASM

roudoudou Assembler
v0.64

« Copyright © BERGÉ Édouard (roudoudou)

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation/source files of RASM, to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software. The Software is provided "as is", without warranty of any kind, express or implied, including but not limited to the warranties of merchantability, fitness for a particular purpose and noninfringement. In no event shall the authors or copyright holders be liable for any claim, damages or other liability, whether in an action of contract, tort or otherwise, arising from, out of or in connection with the software or the use or other dealings in the Software. »

Table of Contents

Introduction.....	5
Installation.....	5
Compilation.....	5
Rasm behaviour.....	6
Using command line.....	7
File option.....	8
Symbol option.....	8
Compatibility option.....	9
Code source format.....	11
generalities.....	11
Comments.....	11
Literal values.....	12
Authorized characters.....	12
Files.....	13
Labels.....	13
Expressions.....	14
Static variables or alias.....	15
Dynamic variables.....	15
Directives.....	16
EQU.....	16
ALIGN <boundary>.....	16
AMSDOS.....	16
BUILDCPR.....	16
BUILDSNA.....	16
BANK [<rom page number>,<ram page number>].....	17
BANKSET.....	17
IF, IFNOT, ELSE, ELSEIF, ENDIF.....	18
LZ48 / LZ49 / LZ4 / LZX7 / LZEXO.....	19
LZCLOSE.....	19
READ / INCLUDE 'file to read'.....	20
INCBIN 'file to read'[,<offset>,<size>,<extended offset>,<OFF>]]].....	20
INCL48 / INCL49 / INCLZ4 / INCZX7 / INCEXO 'file to read'.....	21
LIMIT <maximum adress>.....	21
ORG <code adress>[,<output adress>].....	21
PROTECT <start adress>,<end adress>.....	22
STR '<string>','<string2>',.....	22
STOP.....	22
PRINT.....	23
LIST / NOLIST / LET / BRK.....	23
RUN <adress>.....	23
WHILE / WEND.....	24
REPEAT <n> / UNTIL REPEAT / UNTIL <condition>.....	24
WRITE DIRECT <lower rom>,<upper rom>,<RAM>.....	25
SAVE 'binary file to write',<adress>,<size>.....	26
CHARSET 'string',<value> <byte>,<value> <start>,<end>,<value>.....	26

SWITCH, CASE, BREAK, DEFAULT, ENDSWITCH.....	27
Macros.....	28
Undocumented instructions.....	29
Limitations.....	29
RASM programming examples.....	30
Retrieve the bank of a label in order to connect a ROM.....	30

Introduction

18 years ago, I programmed an assembler/disassembler called Zasm/dizasm. In fact it was mainly a proof of concept of a mono-pass assembler. I used a little the disassembler with personal projects. Sources were published but the spread was so confidential that it was hard to retrieve them back on my HDD.

Since then, another assembler called Zasm exists, created by a spectrum developer team.

As the code of Rasm is from scratch (except a few concept), it's the pretext to choose a new name: Rasm.

Installation

Rasm a stand-alone executable, there is no installation.

Compilation

Linux compilation:

```
cc rasm_v063.c -O2 -lm -lrt -march=native  
mv a.out rasm  
strip rasm
```

Windows compilation:

```
cl.exe rasm_v063.c -O2
```

Rasm behaviour

Rasm is supposed to be easy to use. It will set default filename, guess what memory region to save or create a cartridge file is ROM bank were selected during assembly.

Rasm allow multiple ORG in the same memory set but it will not allow to write twice the same address in the same memory set. Each ORG, Rasm check there is no overwrite.

If you need to generate many code portion at the same address, you have two possibilities. Either you use the <output> parameter of ORG directive (to choose another output address), or you may create another memory set using directive WRITE DIRECT -1,-1,#C0

When you select ROM with directives BANK or WRITE DIRECT, this is always the same memory set for each ROM. If you need another memory set, do not use ROM.

When using SAVE directive, Rasm disable automatic binary output to file or cartridge. You may want to force CPR writing with directive BUILD CPR.

When possible, Rasm try to display comprehensive error messages to point out the problem.

Rasm is using a pre-processor to normalize code, remove useless spaces, comment, check quotes and legal characters, and finally transform some instructions into others. As an example, XOR,AND,OR,MOD will be converted in a single char like C syntax.

When a read file directive do not refer to an absolute path, le root path is the one from the current file.

Using command line

The mandatory syntaxe is: `rasm.exe <file to assemble> [options]`

Example: `rasm.exe myfile.asm -s`

File option

-o <output prefix> set the base name for default output files (binary, cartridge, symbols). The default value is “rasmoutput”.

`rasm.exe src.asm -o output -s` will produce the following filenames:
`output.bin, output.cpr, output.sym`

- ob <binary filename> set the full name for binary file output
- oc <cartridge filename> set the full name for cartridge file output
- os <symbol filename> set the full name for symbol file output

Symbol option

- s export symbols to Rasm format
- sp export symbols to Pasm format
- sw export symbols to Winape format
- ss export symbols in the snapshot

- sl additionnal option to the three previous one. It exports also local symbols.
- sv other additionnal option which exports variables.
- sq other additionnal option which exports alias EQU.
- sa option meaning -sl -sv -sq

- Arnold emulator can handle any export format
- Winape cannot handle label of hiw own format when they are too long. The debugger will crash.

- l <label file> import symbols before assembling

Compatibility option

-m maxam calculation style

Using this option change the behaviour of rasm. Compare operator is the simple =. All calculations are done with unsigned integer 16 bits values.

Developer options

- v simple verbose mode
- d pre-processing verbose mode
- a assembling verbose mode
- n display third-parties licences

Code source format

generalities

Assembleur is not COBOL, it is useless to indent sources with Rasm, except to be pretty. You may use two points to separate labels, or not. Rasm try to be free style. The direct consequence of this free writing style involves a difference with the assemblers of obsolete design. It is not possible (fortunately) to create a label that has the same name as a directive or Z80 instruction.

Files may be read to Unix or Windows format, an internal conversion is done.

Rasm is not case sensible. All chars will be converted upper case. Don't be surprise when Rasm display error messages.

Comments

Rasm comment starts with semicolon. All chars are ignored until next carriage return.

There is no multiline comment.

Literal values

Rasm knows the following literals:

- In decimal if the value begin with a digit.
- In binary if the value begin with a %.
- In hexadecimal if the value begin with a # or 0x.
- An ascii value if a single character is between quote.
- An internal value (variable or label) if the literal begin with a letter or an '@' for local labels.

Rasm does all its internal calculations in double floating-point precision. A correct rounding is performed at the end of the calculation chain for the integer needs.

Caution, the & character is reserved for the AND operator.

Authorized characters

Between quotations, all characters are allowed, at your own risk, for ASCII conversion to Amstrad. outside quotes, you can use all letters, all digits, point, arobas, parentheses, dollar, plus, minus, multiplied, divided, pipe, circumflex, percent, sharp, the paragraph, the rafters and the two types of quotes.

Files

Rasm automatically convert path into Unix style internally. As a consequence, if you use only relative paths (with filename that windows can read), the code may be compiled with Rasm Unix and Rasm Windows without modification.

The file management rule is quite simple. Each relative path root iare located to the root of the file where it was read.

Labels

Inside a loop (REPEAT / WHILE/UNTIL) or inside a macro it is possible to use local labels in the same way as with the integrated Winape assembler prefixing the label with the character '@'.

For each loop iteration, for each loop nesting, a suffix is added to the local label containing the hexadecimal value of the internal repetition counter. It is thus possible to call a local label to a repetition outside the loop, but this use is not advised.

It's possible to declare a label starting by a dot. This is an old declaration style. The call to this label must be done without the starting dot.

It's possible to use a label in a directive (ORG for example) if the label declaration is before the directive.

You can use the prefix {BANK} with a label (example: {BANK}mylabel) to get the number of the bank where the label is located, instead of the label adress.

Expressions

Rasm use a simple expression engine how can handle the following operators/functions:

- * multiplication
- / division
- + addition
- - subtraction
- & boolean operator AND
- | boolean operator OR
- ^ boolean operator OR exclusive
- § modulo (use MOD if your keyboard does not support this char)
- AND, OR, XOR, MOD inline maxam style
- sin() sinus
- cos() cosinus
- asin() arc-sinus
- acos() arc-cosinus
- atan() arc-tangente
- int() double to int conversion
- floor() truncate value to lower integer
- abs() absolute value
- ln() natural logarithm
- log10() logarithm base 10
- exp() exponential
- sqrt() square root

Comparison operators:

- == equality (only one = in maxam compatibility mode)
- != different from
- <= less than or equal to
- >= greater than or equal to
- < less than
- > greater than

Static variables or alias

It's possible to create aliases with EQU directive. Thoses aliases can't be modified.

Dynamic variables

Rasm may use an unlimited number of variables for internal calculations.

Syntax: `mavariablename=5` or `LET mavariablename=5`

Thoses variables may be used as loop counter, offset or value in a loop or the main code.

Example:

```
dep=0
repeat 16
ld (ix+dep),a
dep=dep+8
rend
```

```
ang=0
repeat 256
defb 127*sin(ang)
ang=ang+360/256
rend
```

Directives

EQU

Create an alias

Example:

```
myvar EQU 5  
myothervar EQU myvar*2
```

ALIGN <boundary>

Align code to a given boundary

Examples:

```
ALIGN 2 to align to even adress  
ALIGN 256 to align to the high byte adress
```

AMSDOS

Add an Amsdos header to the automatic binary file. Note: There is no Amsdos header added with SAVE directive.

BUILD CPR

Force cartridge output when using SAVE directive.

BUILD SNA

Select snapshot output.

BANK [<rom page number>,<ram page number>]

Select a ROM for assembling. Even if a ROM size is 16K, you can start to write anywhere in the ROM. Rasm will use the first outputed byte as start adress when creating CPR file. ROM number range is from 0 to 31 and RAM number range is from 0 to 35.

Without parameter, the directive opens a new memory workspace like the ugly WRITE DIRECT -1,-1,#C0

BANKSET

Select 4 pages at a time. This option implies snapshot generation. The bankset 0 stand for pages 0,1,2,3 then the bankset 1 stand for pages 4,5,6,7 ...

If you mix BANK and BANKSET, do not use pages of a bankset. A control mechanism will check you are not trying to do so.

BANKSET usage enable snapshot generation.

IF, IFNOT, ELSE, ELSEIF, ENDIF

Conditionnal assembly

Example:

```
CODE_PRODUCTION=1
[...]
IF CODE_PRODUCTION
OR #80
ELSE
PRINT 'Test version'
ENDIF
```

LZ48 / LZ49 / LZ4 / LZX7 / LZEXO

Open a LZ48, LZ49, LZ4, ZX7 or Exomizer crunched code part. The output code will be crunched after assembly and following code will be relocated.

It's not possible to call a label located after the crunched code, from the crunched code as we cannot know how big will be the crunched code. There will be an error if you try so.

Limitations:

- It's not possible to assemble a code bigger than 64ko before crunch.
- It's not possible to embed portions of crunched code.

LZCLOSE

Close a crunched part opened with LZ48, LZ49, LZ4, LZX7 or LZEXO

READ / INCLUDE 'file to read'

Read a text file and include it into the current source code. Relative path starts from the absolute location of the current file. An absolute path do not care about this.

There is not recursion limit. Be aware of what you do.

INCBIN 'file to read',[offset],[size],[extended offset],[OFF]]]

Include a binary file. Read data will be directly outputed. Parameters are like Winape syntax except first offset is not limited to 64K. Extended offset is here for compatibility only.

The offset parameter may be a negative value. Then the offset is relative to the end of the file.

The size parameter may be a negative value. Then this value is relative to the filesize, in order to read less bytes. If you want to stop 10 bytes before the end, juste put -10 in the filesize.

A zero size will load all the file.

OFF parameter: If you want to load a file in order to initialise memory then assembling code over it, you may disable the overwrite check.

Example:

```
org #4000
incbin'makeraw.bin',0,0,0,OFF ; read the file in #4000
org #4001
defb #BB ; overwrite second byte of the file without error
```

INCL48 / INCL49 / INCLZ4 / INCZX7 / INCEXO 'file to read'

Read a binary file, crunch it with LZ48, LZ49, LZ4, LZX7 or Exomizer and directly output in the memory.

LIMIT <maximum adress>

Set a lower limit (other than 64K) for outputed code. If you want to protect a zone, you may use PROTECT directive.

ORG <code adress>[,<output adress>]

Assemble at a given adress. You may set an output adress to generate the code at a different memory adress.

PROTECT <start adress>,<end adress>

Protect memory region in the current memory set. Any write to this region will trigger an error and a message.

STR '<string>','<string2>',...

Similar to DEFB '<string>', the last character of each string has bit 7 set to 1 (OR #80 on the last byte of each string).

STOP

Stop assembling.

PRINT

Write text/variables/formulas during assembling. It's possible to set display format with tag prefixes.

{hex} Display in hexa. If the value is less than #100 then the display is forced to 2 digits. If the value is less than #10000 then the display is forced to 4 digits. With upper value there won't be any extra-zeros.

{hex2}, **{hex4}**, **{hex8}** display is forced to 2, 4 or 8 digits, with any value.

{bin} display value as binary. If the value is less than #100 then the display is forced to 8 bits. If the value is less than #10000 then the display is forced to 16 bits. With upper value there won't be any extra-zeros. A preprocessing remove the 16 upper bits of the 32 bits value if all bits are set to one (aka 16 bits negative value).

{bin8}, **{bin16}**, **{bin32}** display is forced to 8, 16 or 32 digits, with any value.

{int} display rounded integer value.

Without prefix the value is displayed as a floating-point value.

LIST / NOLIST / LET / BRK

Directives are ignored, only for compatibility with Winape.

RUN <adress>

Set the run adress when using snapshot output.

WHILE / WEND

Repeat a block of instructions everytime the condition is true.

Example:

```
cpt=10
while cpt>0
ldi
cpt=cpt-1
wend
```

REPEAT <n> / UNTIL | REPEAT / UNTIL <condition>

Repeat a block of instructions. You may set a fixed number of repetition or set a conditionnal repeat.

Examples:

```
repeat 16
ldi
rend
```

```
cpt=10
repeat
ldi
cpt=cpt-1
until cpt>0
```


WRITE DIRECT <lower rom>,<upper rom>,<RAM>

If you set a ROM number (lower from 0 to 7, upper from 0 to 31), the directive will have the very same effect as BANK directive.

If you set the RAM (disabling ROM with -1 both lower and upper ROM), Rasm create a new memory set for assembling. Using multiple WRITE DIRECT you can assemble many code at the same adress, but different memory sets.

Example:

```
;default memory space, ORG at zero
defs 65536,0
WRITE DIRECT -1,-1,#C0
defs 65536,0
; new memory space, no matter the RAM bank
; in this example twice #C0
WRITE DIRECT -1,-1,#C0
; no error since it's a third memory space
defs 65536,0
```

Rasm created three memory spaces. The default memory space and two additionnal set for each WRITE DIRECT call. There is no limit for memory space, except your system memory.

When a new memory space is created, you cannot get back to any previous one.

SAVE 'binary file to write',<adress>,<size>

Write a binary file starting from adress of the current memory set to adress+size. Note that the saves are done only if there is no error at the end of assembly.

CHARSET 'string',<value> | <byte>,<value> | <start>,<end>,<value>

Allows the value of characters or characters in strings to be assigned alternative values. There are 4 forms of this directive.

- 'string',<value> First character of the string will have a new ascii <value>. Next char will have <value>+1 etc. until the end of the string.
- <byte>,<value> set a new <value> for the character <byte>
- <start>,<end>,<value> Set from the character <start> until <end> an incremental value starting from <value>.
- “no parameter” will set default values to all characters.

This directive is compatible with Winape.

SWITCH, CASE, BREAK, DEFAULT, ENDSWITCH

The syntax of those directives is like C syntax except you can write more than one case with the same value, which gives more flexibility.

In this example, calling the macro with 5 will assemble all 'yes' and 'yes again' strings:

```
macro yesyes myvar
switch {myvar}
    nop ; won't be assembled because outside case
    case 3
        defb 'no'
    case 5
        defb 'yes'
    case 7
        defb 'again yes'
        break
    case 8
        defb 'no'
    case 5
        defb 'again yes'
        break
    default
        defb 'no'
endswitch
mend
```

Macros

Rasm handle macros with curly brackets (Winape compatible). It is possible to do conditionnal assembling inside macros because each macro call insert a brand new code with new substitutions of parameters. Then the code is interpreted like if there were no macro.

Example for a long distance LD (IX+offset),register write:

```
macro LDIXREG register,dep
if {dep}<-128
    push bc
    ld bc,{dep}
    add ix,bc
    ld (ix+0),{register}
    pop bc
else
if {dep}>128
    push bc
    ld bc,{dep}
    add ix,bc
    ld (ix+0),{register}
    pop bc
else
    ld (ix+{dep}},{register}
endif
endif
mend
```

Note: It's possible to use ENDM as MEND alias

Undocumented instructions

All documented and undocumented instructions are supported.

Index registers IX and IY 8-bits form may be used with `ixl`, `lx` ou `xl`, etc.

Complex instructions syntax:

```
res 0,(ix+0),a  
bit 0,(ix+0),a  
sll 0,(ix+0),a  
rl  0,(ix+0),a  
rr  0,(ix+0),a
```

Input/Output undocumented instructions syntax:

```
out (<n>),a ; <n> is a 8 bits value  
in a,(<n>)  
in 0,(c) or in f,(c)
```

Limitations

- It's not possible to use instruction or directive as label.
- It's not possible to declare label, alias or variable with the same name.
- Rasm is not case sensitive.
- Crunched section cannot exceed 64K before crunch.

RASM programming examples

Retrieve the bank of a label in order to connect a ROM

Using the {BANK} prefix allow to move label in another bank without changing the code.

Bank 0

```
ld a,{bank}maroutine ; will assemble LD A,1
```

```
call connect_bank    ; define your own!
```

```
jp maroutine
```

bank 1

```
defb 'coucou'
```

```
maroutine
```

```
jr $
```