

Taller de programación de sistemas. Diseño del ensamblador.

Práctica No. 1. Identificación de las partes de una línea de lenguaje ensamblador. Implementar el algoritmo para separar las partes de una línea de lenguaje ensamblador.

Descripción de la práctica:

1. El usuario podrá capturar con un editor de texto puro (por ejemplo, el “bloc de notas”) un código como el siguiente.

```
; comentario número 1 =TALLER=

ORG      %00001111
Et1  equ      $ffff
dos  LDAA      @4732
      SWI
      DS.b      %0011000011111100
;comentario numero 2 ;Programación!
;comentario numero 3
tres sWi
      End
```

2. Al capturar el ejemplo utilizar espacios en blanco y tabuladores de manera alternada para crear varias opciones de separación de palabras.
3. Grabar el archivo con el nombre que usted desee y dejarlo con extensión ASM
4. Cada una de las líneas capturadas es un ejemplo ficticio de un código en lenguaje ensamblador para la arquitectura HC12 de Motorola. Cada una de estas líneas puede ser de dos formas:
 - a. Línea de comentario.
 - b. Línea de ETIQUETA, CODIGO DE OPERACIÓN (CODOP) y OPERANDO.
5. Los comentarios tienen las siguientes reglas
 - a. Comienzan con el carácter de “;”.
 - b. Este carácter de “;” puede encontrarse en cualquier posición de la línea
 - c. Después del carácter de “;” pueden seguirle cualquier carácter alfanumérico. En caso de letras pueden ser indistintamente mayúsculas y minúsculas.
 - d. No existe un límite de caracteres por cada comentario.
 - e. El delimitador de la línea de comentario es el retorno de carro (“enter”).
6. Las líneas que están formadas por ETIQUETA, CODOP y OPERANDO tienen las siguientes reglas de escritura:
 - a. ETIQUETAS:
 - i. Comienzan con letra mayúsculas o minúsculas (son válidos los dos casos).
 - ii. Esta primera letra debe de estar sólo en la primera posición de la línea.
 - iii. Su longitud es de 8 caracteres máximo.
 - iv. Después de la primera letra le pueden seguir más letras, números (0 .. 9) o guiones bajos “_”.
 - v. Cualquier otro carácter representado es un error.
 - b. CODOPS (códigos de operación)
 - i. Comienzan con letra mayúscula o minúscula (son válidos los dos casos).
 - ii. El otro carácter válido es el “.” y sólo puede haber uno sólo en la palabra.
 - iii. Su longitud máxima es de 5 caracteres

- iv. Cualquier otro carácter es un error.
- c. OPERANDOS:
 - i. Pueden comenzar con cualquier carácter (en las siguientes prácticas se revisaran las excepciones).
 - ii. Pueden tener cualquier longitud
- 7. Este tipo de líneas pueden tener las siguientes combinaciones:
 - a. ETIQUETA, CODOP, OPERANDO
 - b. ETIQUETA, CODOP
 - c. CODOP, OPERANDO
 - d. CODOP
- 8. Para el primer caso: ETIQUETA, CODOP, OPERANDO, debemos de considerar que:
 - a. Si la línea comienza con ETIQUETA, el primer carácter de la línea debe de ser entonces una letra (mayúscula o minúscula).
 - b. Entre ETIQUETA y CODOP sólo podemos tener separadas estas dos palabras por espacios en blanco o tabuladores (considerar los dos casos).
 - c. Entre CODOP y OPERANDO sólo podemos tener separadas estas dos palabras por espacios en blanco o tabuladores (considerar los dos casos).
 - d. Después de la palabra OPERANDO se pueden tener espacios, pero se finaliza con el retorno de carro o “enter” este siempre será el delimitador de cada sentencia de lenguaje ensamblador.
- 9. Para el segundo caso: ETIQUETA, CODOP, debemos de considerar que:
 - a. Si la línea comienza con ETIQUETA, el primer carácter de la línea debe de ser entonces una letra (mayúscula o minúscula).
 - b. Entre ETIQUETA y CODOP sólo podemos tener separadas estas dos palabras por espacios en blanco o tabuladores (considerar los dos casos).
 - c. Después de CODOP se pueden tener espacios, pero se finaliza con el retorno de carro o “enter” este siempre será el delimitador de cada sentencia de lenguaje ensamblador.
- 10. Para el tercer caso: CODOP, OPERANDO, debemos de considerar que:
 - a. Antes de la palabra CODOP sólo podemos tener espacios en blanco o tabuladores (considerar los dos casos). Es decir, el primer carácter de la línea será o un espacio en blanco o un tabulador.
 - b. Entre las palabras CODOP y OPERANDO sólo podemos tener espacios en blanco o tabuladores, considerar los dos casos.
 - c. Después de la palabra OPERANDO se pueden tener espacios, pero se finaliza con el retorno de carro “enter” este siempre será considerado como el delimitador de cada sentencia de lenguaje ensamblador.
- 11. Para el cuarto caso: CODOP debemos de tener las siguientes consideraciones:
 - a. Antes de la palabra CODOP sólo podemos tener espacios en blanco y tabuladores considerar los dos casos. Es decir, el primer carácter de la línea será o un espacio en blanco o un tabulador.
 - b. Después de la palabra CODOP se pueden tener espacios, pero se finaliza con el retorno de carro o “enter” este debe de ser considerado siempre como el delimitador de una línea de lenguaje ensamblador.
- 12. Se debe de observar que siempre debe de haber por lo menos la palabra que conforme el CODOP ya que de otra forma sería un error.
- 13. La última línea del ejemplo siempre debe de tener la palabra END (ya sea escrita con mayúsculas, minúsculas o una combinación de mayúsculas y minúsculas). Es decir, al validar la línea que tiene la palabra END debemos de considerar que este es el final del archivo con el ejemplo de lenguaje ensamblador y detener la lectura del archivo
- 14. Al ser END una directiva esta debe ir siempre en la posición de CODOP, y puede tener una etiqueta.
- 15. Si no se encuentra la directiva END, indicarlo en el archivo de errores.

Para hacer el programa:

1. Hacer un programa que solicite la ruta del archivo y lo abra.
2. Leer línea por línea el archivo y llevar un conteo de las mismas, pudiendo:
 - a. Leer cada línea directamente del archivo, de una por una y procesar cada línea que se va leyendo; o bien,
 - b. Leer todo el archivo con un ciclo y conforme se lee el archivo almacenarlo en una estructura de datos temporal, posteriormente leer de esta estructura de datos.
 - c. En ambos casos existen ventajas y desventajas. Deberás de elegir la que sea más conveniente, recuerda que el programa debe de irse complementando con las siguientes prácticas.
3. Una vez que identificamos la primera línea se debe de separar esta en tokens para identificar las palabras e ignorar los espacios en blanco y tabuladores.
4. Una vez que se identifique la primera palabra de una línea se debe de determinar si esta fue ETIQUETA, CODOP u OPERANDO.
5. Se deben de utilizar tres variables que se deben de llamar, precisamente, como “etq”, “codop” y “oper” que deben de ser de tipo cadena ya que estas leyendo de un archivo de texto. Una vez que se identifiquen plenamente las palabras de una línea debes de imprimir en un archivo el valor de cada variable.
6. Si alguna variable no tuviera ningún valor debes de asignarle un valor “NULL”.
7. Se deberá crear un archivo con el mismo nombre del archivo recibido pero con extensión “INST”
8. En el archivo de instrucciones (INST) se deberán imprimir las líneas ya procesadas separadas por tabulaciones, indicando el número de línea y cada uno de sus tokens. (ver ejemplo mas adelante).
9. Para las validaciones de ETIQUETA, CODOP y OPERANDO, en los casos en que estas no sean válidas, generar un archivo con el mismo nombre del archivo ASM pero ponerle extensión ERR indicando el número de línea y el tipo de error. (Se debe indicar claramente el motivo del error).
10. Ni las líneas de comentarios ni las líneas con errores aparecen en el archivo de instrucciones.

Ejemplo de salida al archivo P1ASM.INST para el archivo P1ASM.ASM propuesto:

LINEA	ETQ	CODOP	OPER
2	NULL	ORG	%00001111
3	Et1	equ	\$ffff
4	dos	LDAA	@4732
5	NULL	SWI	NULL
6	NULL	DS.b	%0011000011111100
9	tres	sWi	NULL

Características del reporte.

1. Debe de imprimirse, máximo una página.
2. No debe de tener faltas de ortografía.
3. Debe de tener el nombre del alumno código, grupo y fecha de entrega.
4. Descripción del reporte:
 - a. Describir el algoritmo utilizado para leer el archivo. Describir el algoritmo para identificar las palabras.
 - b. No imprimir código.
 - c. Incluir una conclusión del desarrollo de la práctica