

**SISTEMAS OPERATIVOS**

**CURSO 2017-2018**

**SIMULADOR DE UN SISTEMA  
INFORMÁTICO PRIMITIVO**

**MANUAL - V0**

---

## Introducción

Construir un sistema operativo (SO) sobre hardware real es una tarea muy compleja, por muy simple que sea el SO a construir. La razón de ello es que el hardware es complejo y diverso y, por tanto, toda pieza de software que aspire a controlarlo y gestionarlo adecuada y eficientemente, tiene que ser, a su vez, muy compleja. A pesar de ello, no renunciamos a poder estudiar los detalles e interioridades de un SO desde el punto de vista de su diseño y programación y, si no podemos hacerlo de una manera muy realista, podemos simplificar (quedándonos únicamente con los aspectos relevantes) y simular.

Los primeros sistemas informáticos (SI) eran muy primitivos. Cada uno de sus componentes, tanto hardware como software, era muy simple, en comparación con los que existen en la actualidad. Desde el punto de vista de la simulación, nuestro SI primitivo estaría compuesto de:

- Un procesador (primitivo).
- Una memoria principal.
- Los buses del sistema.
- Un SO (primitivo).

Explicándolo muy rápidamente, el SO se encargaría de cargar en la memoria principal un programa ejecutable, que sería posteriormente ejecutado por el procesador. Veamos a continuación el detalle de los componentes anteriores.

## DISEÑO

### *El procesador*

El procesador es muy elemental. Desde el punto de vista de estas prácticas no nos interesa hacer una disección tradicional del mismo (registros, Unidad de Control, Unidad Aritmético y Lógica) sino que realizaremos un diseño basado más en aspectos funcionales del mismo. Por tanto, el procesador estará compuesto de:

- Un conjunto de registros:
  - Registro acumulador, que se usará, en general, en operaciones aritméticas y de manipulación de la memoria principal.
  - Registro contador de programa (Program counter - PC).
  - Registro de instrucción (Instruction register - IR).
  - Registro de palabra de estado del procesador (Processor state word - PSW).
  - Registro de direcciones de memoria (Memory address register - MAR).
  - Registro intermedio de memoria (Memory buffer register - MBR).
- La funcionalidad necesaria para la ejecución iterativa del ciclo de instrucción (etapas del ciclo de instrucción):
  - Búsqueda de instrucción.
  - Decodificación de instrucción.
  - Ejecución de la instrucción.
  - Almacenamiento del resultado.
  - Tratamiento hardware de interrupciones.

Evidentemente, todo procesador está construido para procesar un cierto conjunto de instrucciones

definidas a priori (su juego de instrucciones). El juego de instrucciones del procesador es muy simple:

- `ADD op1 op2`: suma los valores enteros `op1` y `op2` y deja el resultado dentro del registro acumulador.
- `SUB op1 op2`: resta el valor `op2` de `op1` y deja el resultado dentro del registro acumulador.
- `NOP`: típica instrucción *NO-OPERATION* de los procesadores, que consume un ciclo de instrucción sin realizar acción alguna.
- `JUMP dirRelativa`: modifica el valor del registro PC sumándole el valor indicado en `dirRelativa`.
- `ZJUMP dirRelativa`: funciona igual que `JUMP` si el valor actual del registro acumulador es 0.
- `READ dir`: lee el contenido de la posición de memoria principal indicada por `dir`, depositando el valor leído en el registro acumulador.
- `WRITE dir`: almacena el valor contenido el registro acumulador en la posición de memoria principal indicada por `dir`.
- `INC op1`: incrementa el valor actual del acumulador con el valor indicado por `op1`.
- `HALT`: provoca la detención, definitiva, del funcionamiento del procesador.

Por ejemplo, el programa siguiente recorre los números enteros desde el 10 hasta el 0, momento en el que se detiene el simulador:

```
ADD 10 0
INC -1
ZJUMP 2 // Si el acumulador==0 en la última operación, saltar dos posiciones hacia adelante
JUMP -2 // Saltar dos posiciones de memoria hacia atrás
HALT
```

## ***La memoria principal***

Desde el punto de vista del diseño, la memoria principal puede ser considerada como un vector (*array*) de celdas de memoria, todas ellas con la misma capacidad de almacenamiento. Las operaciones básicas que se pueden hacer sobre la memoria son dos:

- Almacenar un cierto valor en una posición (celda) de memoria dada.
- Recuperar el valor almacenado en una cierta posición (celda) de memoria.

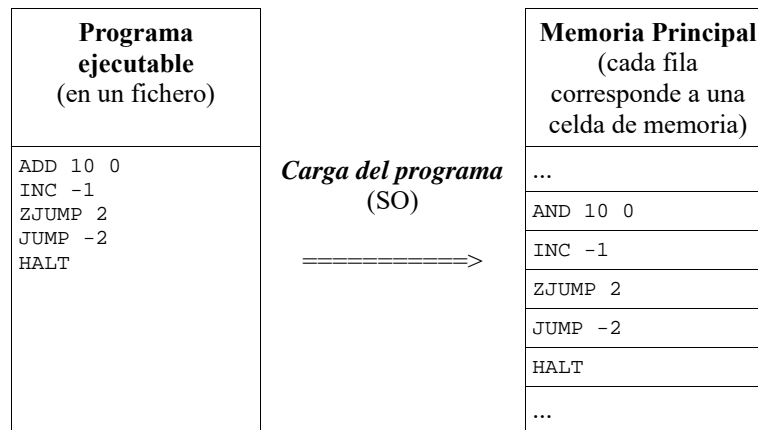
En el apartado anterior vimos que el procesador dispone de instrucciones para realizar las acciones anteriores. No podemos perder de vista, tampoco, que el procesador también accede a la memoria, por ejemplo, para obtener una instrucción o para almacenar el resultado de su ejecución (si es el caso).

## ***Los buses del sistema***

Son las típicas “vías de comunicación” entre diferentes componentes hardware. En el SI primitivo comunicarán únicamente al procesador con la memoria principal.

## ***El sistema operativo***

El SO de este SI primitivo se restringe a ser capaz de cargar en memoria principal un programa ejecutable almacenado en un fichero. El procesador sólo es capaz de ejecutar instrucciones almacenadas en memoria principal, por lo que la carga realizada por el SO es indispensable.



## IMPLEMENTACIÓN

### *Consideraciones generales sobre la implementación*

El simulador de SI está escrito en lenguaje C. En general, se han repartido en ficheros de código fuente diferentes los diferentes componentes del SI que resulta razonable e interesante separar. Además, cada componente del SI está implementado utilizando dos ficheros:

- Un fichero cuyo nombre termina en “.c”, que define las propiedades básicas de ese componente e implementa su funcionalidad.
- Un fichero cuyo nombre termina en “.h”, que define tipos de datos propios de ese componente y los prototipos de las funciones implementadas en el correspondiente fichero “.c”, y que deben ser conocidos desde otros ficheros.

### *El procesador*

Está implementado en los ficheros `Processor.h` y `Processor.c`. Define el comportamiento de un procesador elemental, ya comentado anteriormente. El conjunto de propiedades de este componente hardware se limita a sus registros:

```
int registerPC_CPU; // Program counter
int registerAccumulator_CPU; // Accumulator
MEMORYCELL registerIR_CPU; // Instruction register
unsigned int registerPSW_CPU; // Processor state word
int registerMAR_CPU; // Memory Address Register
MEMORYCELL registerMBR_CPU; // Memory Buffer Register
```

En cuanto a su funcionalidad, define cuatro funciones:

- Una función para dar valores iniciales a los registros del procesador:

```
void Processor_InitializeRegisters(int regPC, int regAcum, unsigned int regPSW) {
    registerPC_CPU=regPC;
    registerAccumulator_CPU=regAcum;
    registerPSW_CPU=regPSW;
}
```

- Una función que simula el comportamiento iterativo de todo procesador: la ejecución de su ciclo de instrucción. La ejecución iterativa del ciclo de instrucción llega a su fin cuando en el registro PSW se anota el valor `POWEROFF`. Dicha anotación tiene lugar cuando se ejecuta la instrucción `HALT`. Las etapas habituales del ciclo de instrucción se han separado en tres funciones:

```

void Processor_InstructionCycleLoop() {
    while (registerPSW_CPU!=POWEROFF) {
        Processor_FetchInstruction();
        Processor_DecompileAndExecuteInstruction();
        Processor_ManageInterrupts();
    }
}

```

- Una función que simula la etapa de búsqueda de instrucción:

```

void Processor_FetchInstruction() {

    // The instruction must be located at the memory address pointed by the PC register
    registerMAR_CPU=registerPC_CPU;
    // Send to the main memory controller the address in which the reading has to take
    place: use the address bus for this
    Buses_write_AddressBus_From_To(CPU, MAINMEMORY);
    // Tell the main memory controller to read
    MainMemory_readMemory();
    // All the read data is stored in the MBR register. Because it is an instruction
    // we have to copy it to the IR register
    memcpy((void *) (&registerIR_CPU), (void *) (&registerMBR_CPU), sizeof(MEMORYCELL));
    ...
}

```

- Una función que simula las etapas de decodificación de instrucción, búsqueda de operandos, ejecución de instrucción y almacenamiento del resultado (lo que sigue es sólo un extracto):

```

void Processor_DecompileAndExecuteInstruction() {

    switch (registerIR_CPU.operationCode) {

        // Instruction ADD
        case 'a': registerAccumulator_CPU=
                    registerIR_CPU.operand1 + registerIR_CPU.operand2;
                    registerPC_CPU++;
                    break;

        // Instruction SUB
        case 's': registerAccumulator_CPU=
                    registerIR_CPU.operand1 - registerIR_CPU.operand2;
                    registerPC_CPU++;
                    break;

        ...
    }
}

```

## La memoria principal

Está implementada en los ficheros `MainMemory.h` y `MainMemory.c`. Define el comportamiento básico de la memoria principal: leer una celda de memoria y escribir en una celda de memoria.

La memoria principal está definida como un array de `MAINMEMORYSIZE` celdas de memoria:

```

// Main memory size (number of memory cells)
#define MAINMEMORYSIZE 256

// A memory cell is capable of storing a structure of the
// MEMORYCELL TYPE
typedef struct {
    char operationCode;
    int operand1;
    int operand2;
} MEMORYCELL;

```

```
MEMORYCELL mainMemory[MAINMEMORYSIZE];

int registerMAR_MainMemory;

MEMORYCELL registerMBR_MainMemory;
```

A la vista de las definiciones previas, se puede comprobar que en una celda de memoria se puede llegar a almacenar una instrucción compuesta por un código de operación y dos operandos. Cuando describamos la implementación del SO entenderemos la razón que está detrás de esta decisión.

Además, la memoria principal tiene dos registros:

- Registro de direcciones de memoria (Memory Address Register - MAR): contiene la dirección de memoria que indica dónde va a tener lugar la siguiente operación de acceso a la misma.
- Registro intermedio de memoria (Memory Buffer Register - MBR): contiene la información a ser escrita o la información ya leída en una operación de acceso a la memoria.

En cuanto a su funcionalidad, define dos funciones:

- Una función para leer el contenido de una celda de memoria principal:

```
void MainMemory_readMemory() {
    // registerMBR_MainMemory.operationCode= mainMemory[registerMAR_MainMemory].operationCode;
    // registerMBR_MainMemory.operand1      = mainMemory[registerMAR_MainMemory].operand1;
    // registerMBR_MainMemory.operand2      = mainMemory[registerMAR_MainMemory].operand2;
    memcpy((void *)(&registerMBR_MainMemory), (void *)(&mainMemory[registerMAR_MainMemory]),
           sizeof(MEMORYCELL));
    Buses_write_DataBus_From_To(MAINMEMORY, CPU);
}
```

- Una función para escribir en una celda de memoria principal:

```
void MainMemory_writeMemory() {
    Buses_write_DataBus_From_To(CPU, MAINMEMORY);
    // mainMemory[registerMAR_MainMemory].operationCode = registerMBR_MainMemory.operationCode;
    // mainMemory[registerMAR_MainMemory].operand1      = registerMBR_MainMemory.operand1      ;
    // mainMemory[registerMAR_MainMemory].operand2      = registerMBR_MainMemory.operand2      ;
    memcpy((void *) (&mainMemory[registerMAR_MainMemory]), (void *) (&registerMBR_MainMemory),
           sizeof(MEMORYCELL));
}
```

## Los buses del sistema

Están implementados en los ficheros `Buses.h` y `Buses.c`. Su funcionalidad se limita a simular el funcionamiento de los buses de datos y direcciones: leer el contenido del registro apropiado del componente hardware que envía la información a través del bus y escribir en el registro apropiado del componente hardware destinatario de la información.

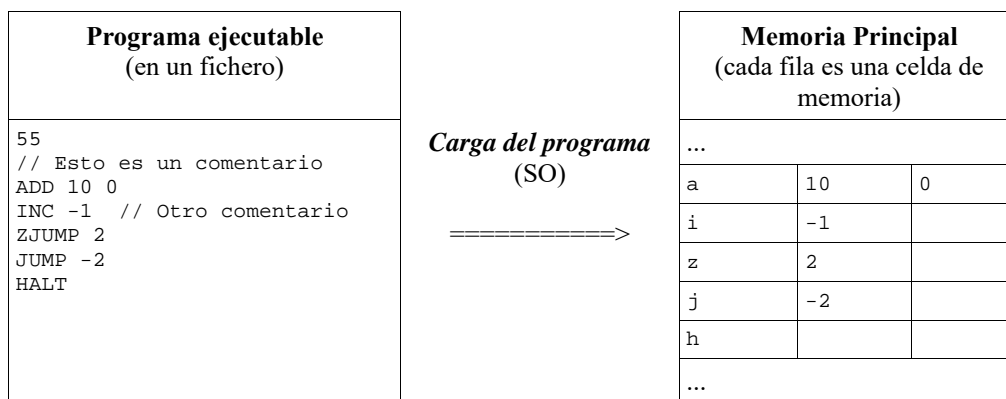
```
void Buses_write_AddressBus_From_To(int fromRegister, int toRegister) {
    ...
    data=Processor_GetMAR(); // si fromRegister es CPU
    ...
    MainMemory_SetMAR(data); // y toRegister es MAINMEMORY
    ...
}

void Buses_write_DataBus_From_To(int fromRegister, int toRegister) {
    ...
    MainMemory_GetMBR(data); // si fromRegister es MAINMEMORY
    ...
    Processor_SetMBR(data);  // si toRegister es CPU
    ...
}
```

## ***El sistema operativo***

Está implementado en los ficheros `OperatingSystem.h` y `OperatingSystem.c`. Su funcionalidad se limita a la carga de un programa ejecutable en memoria principal, teniendo en cuenta que:

- El programa ejecutable contiene un primer valor entero, que indica el número de posiciones de memoria principal que necesita el programa.
- A continuación, viene un conjunto de líneas que contienen las instrucciones del programa:
  - Cada instrucción ocupa una línea del fichero.
  - El programa ejecutable puede contener comentarios.
- Durante la carga del programa, el SO acomoda la información leída a un formato que podrá tratar con mayor facilidad el procesador:
  - Los códigos de operación se reducen a un sólo carácter.
  - En cada celda de memoria quedan separados código de operación, primer operando (si existe) y segundo operando (si existe).



## El sistema informático

Es un componente necesario desde el punto de vista de la implementación y juega, fundamentalmente, el papel de “contenedor” del resto de componentes. Está implementado en los ficheros `ComputerSystem.h` y `ComputerSystem.c`. Sus funciones principales consisten en arrancar y apagar el sistema y proveer una función para la visualización de mensajes de funcionamiento interno del simulador.

- La función que arranca el sistema se encarga de inicializar los registros del procesador, , cargar los mensajes de depuración, solicitar al SO que cargue en memoria el programa indicado y, finalmente, indicar al procesador que se disponga a ejecutar, de manera indefinida, su ciclo de instrucción:

```
void ComputerSystem_PowerOn(int argc, char *argv[]) {
...
    // Initial values for the processor registers. In the old days, the
    // CS operator had to initialize them in a similar way
    const int initialValueForPCRegister=230;
    const int initialValueForAccumulatorRegister=0;
    const int initialValueForPSWRegister=128;

    // Load debug messages
    nm=Messages_Load_Messages();
    printf("%d Messages Loaded\n",nm);

    // To remember the simulator sections to be message-debugged and if messages must be coloured
    debugLevel = argv[1];
    ...
    // Initialize processor registers
    Processor_InitializeRegisters(initialValueForPCRegister, initialValueForAccumulatorRegister,
                                initialValueForPSWRegister);

    // If PROGRAM_TO_BE_EXECUTED exists, is executed
    programFile= fopen(PROGRAM_TO_BE_EXECUTED, "r");

    // Check if programFile exists, if not, poweroff system
    if (programFile==NULL)
        ComputerSystem_PowerOff();

    // Load the program in main memory, beginning at the address given by the second argument
    OperatingSystem_LoadProgram(programFile, initialValueForPCRegister);

    // Tell the processor to begin its instruction cycle
    Processor_InstructionCycleLoop();
}
```

- El apagado del sistema, por su parte, es simple: consiste únicamente en hacer finalizar el programa C.

```
void ComputerSystem_PowerOff() {
    exit(0);
}
```

- Finalmente, el SI proporciona una función cuyo fin es dar la posibilidad al programador de realizar una visualización ordenada del funcionamiento interno del simulador. No es necesario comprender la implementación de esta función, pero sí cómo invocarla y crear los mensajes.

La parte constante de los mensajes, los tipos de las partes variables (caracteres de conversión) y los códigos de colores “@?” se especifican en el fichero “`messages.txt`”.

Cada línea tiene el número de mensaje, seguido del formato del mensaje (separado por una coma).

Los posibles caracteres de conversión son: %s, %d, %f y %c, que se sustituirían por los parámetros de tipo: *cadena de caracteres*, *número decimal*, *número real* y *carácter* respectivamente.

Los posibles códigos de colores son: @R (rojo), @G (verde), @B (azul), @M (magenta), @C (cian),



@W (blanco) y @@ (monocromo). Se sigue usando el color definido hasta que aparezca otro código de color o se termine el mensaje.

```
1,%c %d %d
3,(PC: @R%d@@, Accumulator: @R%d@@)\n
```

- o Por ejemplo, la invocación:

```
ComputerSystem_DebugMessage(3,HARDWARE,100,1617);
```

Mostraría en pantalla:

```
(PC: 100, Accumulator: 1617)
```

- o Los parámetros de la función son, por este orden:
  - El número del mensaje
  - La sección del simulador en la que se circunscribe el mensaje (HARDWARE, en el ejemplo). Es un parámetro de tipo `char`. Se pueden encontrar diferentes secciones predefinidas en el fichero `ComputerSystem.h`.
  - Finalmente vienen los argumentos que se corresponden con el mensaje propiamente dicho (100 y 1617, en el ejemplo). Los tipos de estos argumentos tienen que ser compatibles con los caracteres de conversión del mensaje del fichero de mensajes.

## Compilación del simulador

La tarea de compilación se ha simplificado al máximo, con la utilización de un fichero `Makefile`. Basta, por tanto, ejecutar la orden `make` para que se genere el código ejecutable, si es necesario (el fichero `Makefile` contiene una descripción de las dependencias que existen entre los distintos ficheros de código fuente, de manera que sólo se recompilará aquella parte del proyecto software que haya sufrido cambios).

```
PROGRAM = Simulator

# Detalles de Compilation
SHELL = /bin/sh
CC = cc
STDCFLAGS = -g -c -Wall
INCLUDES =
LIBRARIES =

${PROGRAM}: Simulator.o ComputerSystem.o MainMemory.o OperatingSystem.o /
    Processor.o Buses.o Messages.o
    $(CC) -o ${PROGRAM} Simulator.o ComputerSystem.o MainMemory.o /
        OperatingSystem.o Processor.o Buses.o Messages.o $( LIBRARIES)

Simulator.o: Simulator.c Simulator.h ComputerSystem.h
    $(CC) $(STDCFLAGS) $(INCLUDES) Simulator.c
...
```

## Ejecución del simulador

El simulador se invoca, desde línea de comandos, tal y como se puede ver en el siguiente ejemplo:

```
$ ./Simulator A
```

Donde `Simulator` es el nombre del programa ejecutable resultante de la compilación y el argumento `A` indica las secciones del simulador de las que el usuario está interesado obtener mensajes por pantalla. Si ejecutamos la orden anterior, obtendríamos un resultado similar al siguiente:

Fichero del programa	Resultado (en pantalla) de ejecutar la orden:
----------------------	-----------------------------------------------

55	<code>./Simulator A</code>
<code>// Esto es un comentario</code>	<code>a 10 0 (PC: 231, Accumulator: 10)</code>
<code>ADD 10 0</code>	<code>i -1 0 (PC: 232, Accumulator: 9)</code>
<code>INC -1 // Otro comentario</code>	<code>z 2 0 (PC: 233, Accumulator: 9)</code>
<code>ZJUMP 2</code>	<code>j -2 0 (PC: 231, Accumulator: 9)</code>
<code>JUMP -2</code>	<code>i -1 0 (PC: 232, Accumulator: 8)</code>
<code>HALT</code>	<code>z 2 0 (PC: 233, Accumulator: 8)</code>
	<code>j -2 0 (PC: 231, Accumulator: 8)</code>
	<code>...</code>
	<code>i -1 0 (PC: 232, Accumulator: 0)</code>
	<code>z 2 0 (PC: 234, Accumulator: 0)</code>
	<code>h 0 0 (PC: 234, Accumulator: 0)</code>

El texto que aparece en pantalla se corresponde con el resultado de la ejecución (repetida) de la función `ComputerSystem_DebugMessage (HARDWARE, ...)` presente en el fichero `Processor.c`. El argumento `A` de la línea de comandos anterior especifica que se desean mostrar todos (*All*) los mensajes de todas las secciones del simulador. Igual resultado hubiésemos obtenido con la siguiente orden:

```
$ ./Simulator H
```

donde el argumento `H` se corresponde con la sección `HARDWARE` (ver `ComputerSystem.h`).

Si no quisiésemos que apareciese en pantalla ningún mensaje (excepto los de `ERROR`), podríamos utilizar el carácter `N` (*None*, ninguno).

```
$ ./Simulator N
```

Tampoco aparecería en pantalla ningún mensaje si desde línea de comandos especificásemos una o varias secciones del simulador para las que no hay llamadas a `ComputerSystem_DebugMessage(...)` dentro del código fuente del simulador. Por ejemplo, si ejecutamos la orden:

```
$ ./Simulator MF
```

No aparecerá mensaje alguno en la pantalla, dado que no hay llamadas a la función indicada en el código actual del simulador para las secciones `SYSMEM` y `SYSFILE`.

Si las letras correspondientes a las secciones están en mayúsculas, se muestran los mensajes con colores; si son en minúsculas, sólo en un color.

Los mensajes de la sección `ERROR (E)` se muestran siempre, aunque no se ponga la sección.

