

QSim: Simulador de arquitecturas Q

Susana Rosito

Tatiana Molinari

10 de noviembre de 2013

Índice

I Contexto	4
1. Sobre la materia Organización de Computadoras(susi)	4
2. Conceptos importantes	4
2.1. Enfoque de Von Neumann	4
2.2. Organización de la computadora	5
2.3. Ejecución de un programa	6
3. Arquitecturas Q	6
3.1. Características generales(susi)	6
3.2. Modos de direccionamiento	7
3.3. Repertorio de instrucciones	8
3.3.1. Instrucciones de 2 operandos	8
3.3.2. Instrucciones de 1 operando origen	9
3.3.3. Instrucciones de 1 operando destino	9
3.3.4. Instrucciones sin operandos	9
3.3.5. Instrucciones de salto condicional	10
3.4. Versiones de la arquitectura (SUSI)	11
3.4.1. Q1	11
3.4.2. Q2	11
3.4.3. Q3	11
3.4.4. Q4	12
3.4.5. Q5	12
3.4.6. Q6	12
4. Estado del arte	13
II Simulador QSim	13
5. Funcionalidad del simulador	13
5.1. Chequeo de sintaxis	14
5.2. Ensamblado	14
5.3. Cargado en memoria	14
5.4. Ejecución paso a paso	15
6. Implementación	16
6.1. Tecnología utilizada (tati)	16
6.2. (arq OO)	16
6.2.1. ALU	16
6.2.2. Bus de entrada y salida, memoria y puertos	17
6.2.3. CPU	17
6.2.4. Interprete	18
6.2.5. Modos de direccionamiento y W16	18
6.2.6. Instrucciones	19
6.2.7. Programa	20

6.2.8. Simulador	21
III Evaluación del desarrollo	21
7. Dificultades encontradas	22
7.1. Dificultades presentadas por el dominio	22
7.2. Dificultades de diseño	23
8. Casos de prueba	24
8.1. Chequeo de sintaxis en la distintas Qi	24
8.2. Ensamblado (Objetos \rightarrow Código Máquina)	25
8.3. Decodificación Código Máquina \rightarrow Objetos (Interprete)	25
8.4. Ejecución	25
8.4.1. Operaciones de ALU	26
8.4.2. Búsqueda de Operandos (para los distintos modos de direccionamiento)	26
8.4.3. Almacenamiento de Operandos	26
9. Ejemplos de uso	27

Resumen

Este trabajo...

Parte I

Contexto

1. Sobre la materia Organización de Computadoras(susi)

2. Conceptos importantes

2.1. Enfoque de Von Neumann

El matemático John Von Neumann en el año 1945 que por ese entonces se encontraba colaborando en el proyecto ENIAC, se comenzó a interesar por la problemática que significaba la la necesidad de reconfigurar la máquina para cada nueva tarea y tan sólo cuatro años más tarde propone y desarrolla una solución a este problema la cual consistía en poner la información sobre las operaciones a realizar en la misma memoria utilizada para los datos escribiendola en código binario al igual que los datos.

Este enfoque generaliza la organización de las computadoras distinguiendo en tres partes interconectadas: La CPU (con la unidad aritmético-lógica o ALU y la unidad de control) la memoria, y un módulo de entrada/salida. La interconexión es llevada a cabo por un un bus de sistema que proporciona un medio de transporte de los datos entre las distintas partes.

Con la propuesta de este modelo Von Neumann incorpora el concepto de **programa almacenado** *contiene un conjunto de instrucciones que podían ser almacenadas en memoria, o sea, un programa que detalla la computación del mismo*¹. Con esta idea, el programa se codifica de cierta manera para que pueda ser almacenado en memoria principal y posteriormente ejecutado quizás múltiples veces. De esta manera, la lógica del programa puede ser "recordada" y el programa toma un valor mayor, a diferencia de lo que ocurría hasta entonces, donde el programa se reflejaba en un conjunto de configuraciones de cables aplicadas a los equipos. Esto implica una separación entre el mecanismo de ejecución (el hardware) y la lógica de computo o instrucciones (el software).

Este tipo del diseño que permite un programa almacenado también da la posibilidad de que la ejecución de las instrucciones modifique el código máquina del mismo u otro programa. Por ejemplo un programa podría modificar o incrementar las referencias a las direcciones de memoria que tenga en algunas instrucciones y luego volver a ejecutar dichas instrucciones con el fin de procesar celdas diferentes de memoria cada vez. Esta característica es potente pero presenta un alto riesgo pues las modificaciones en los programas podía ser algo perjudicial, por accidente o por diseño.

¹no entiendo esta frase. la puse yo? Si mara, jajaja que hacemos? (tati)

2.2. Organización de la computadora

La CPU (Unidad Central de Procesamiento del inglés: Central Processing Unit), es el componente principal y el encargado ejecutar los programas y procesar los datos. La CPU contiene otros componentes de importancia tales como la **Unidad de Control**, el contador de programa (PC o *Program Counter*), el registro de instrucción (IR - *Instruction Register*), otros registros de uso específico y general y la Unidad Aritmético-Lógica (ALU).

La Unidad de Control dirige el ciclo de ejecución de cada instrucción, pidiendo la lectura de celdas de memoria donde esta alojada, decodificándola y ejecutándola luego en colaboración con los otros componentes del sistema: si es una operación lógica o aritmética le ordena a la ALU su ejecución, si es de movimiento de datos colabora con la memoria ó el módulo de Entrada/Salida.

El contador de programa (en inglés *Program Counter* o PC) es un registro que indica la posición de memoria donde estará la siguiente instrucción que debe ejecutarse. Luego de completar el ciclo de ejecución de una instrucción, el PC se incrementa en función de la cantidad de celdas que ocupa el código máquina de esta.

El registro de instrucción contiene el código máquina de la instrucción actual una vez que la misma es leída de memoria para luego decodificarla y ejecutarla.

El diseño de cada arquitectura ofrece un conjunto diferente de registros de uso general para ser usados en los programas. Estos registros son elementos de memoria de alta velocidad y poca capacidad que pueden ser utilizados como variables en los programas. Es importante marcar que pueden guardar tanto datos como direcciones de memoria.

La Unidad Aritmético-Lógica, recibe su nombre de las siglas en inglés de *Arithmetic and Logic Unit*. La ALU es un circuito digital que lleva a cabo operaciones aritméticas (suma, resta, multiplicación, división) y las operaciones lógicas como la negación, disyunción, conjunción, etc, entre dos cadenas binarias que son interpretadas como números o valores lógicos.

La memoria es un conjunto de celdas numeradas. La numeración de cada celda la identifica inequívocamente por lo cual a esta numeración se le llama dirección. En cada celda de la memoria se pueden almacenar datos o instrucciones en forma de cadenas binarias y este contenido puede leerse y modificarse. En la memoria es donde se alojan los programas que luego serán ejecutados.

El bus de sistema es el encargado de transferir los datos entre los componentes de la computadora. La unidad de control al pedir un contenido de una dirección de memoria lo hace a través del bus, y similarmente mismo cuando desea escribir en memoria.

2.3. Ejecución de un programa

La función de una computadora es la ejecución de programas. Los programas se encuentran almacenados en memoria y consisten en una sucesión de instrucciones que posee orden. La CPU es quien se encarga de ejecutar dichas instrucciones a través de un ciclo denominado ciclo instrucciones. Para ser almacenadas en memoria, las instrucciones deben codificarse en cadenas binarias (secuencias de ceros y unos) que no son legibles para las personas pero que la UC puede interpretar y traducir en acciones. Por eso para saber de qué instrucción se trata, y cuales son los valores o celdas de memoria que debería consultar o usar, la CPU utiliza la UC, que tomando bit por bit interpreta la cadena binaria y verificando los códigos únicos de cada instrucción o modos de direccionamiento, sabe qué instrucción y con qué valores debería ejecutar. La ejecución de instrucciones se divide en tres etapas importantes, fetch - decode - execute.

Al principio de cada ciclo de ejecución, durante el fetch de instrucción la CPU busca una instrucción que se encuentra en alguna parte de la memoria. Para saber donde esta dicha instrucción la CPU contiene un registro llamado contador de programa (PC), que tiene la dirección de la próxima instrucción a buscar. La CPU a través del bus lee la instrucción, y luego incrementa el valor contenido en PC; así podrá buscar la siguiente instrucción en la secuencia luego de terminar con la actual. La instrucción leída que está en la forma de cadena binaria se carga dentro de otro registro de la CPU, llamado registro de instrucción (IR).

Durante la decodificación la UC determina que significa la cadena binaria.

Finalmente al saber de qué instrucción se trata la CPU ejecuta la instrucción, es decir, realiza lo que la instrucción dice que debe hacer con sus argumentos, modificando la memoria o los registros como resultado final y el ciclo vuelve a comenzar hasta que el programa termine.

3. Arquitecturas Q

3.1. Características generales

- Tiene **8 registros** de uso general de 16 bits: **R0..R7**.
- La memoria tiene **direcciones de 16 bits**.
- Tiene un contador de programas (**program counter**) de 16 bits.
- **Stack Pointer** de 16 bits. Comienza en la dirección FFEF.
- Flags: **Z,N,C,V**(Zero, Negative, Carry, Overflow).
- Todas las instrucciones alteran los flags excepto **MOV, CALL, RET, JMP, Jxx**.
- De las instrucciones que alteran los Flags, todas dejan C y V en 0 a excepción de **ADD, SUB y CMP**.
- Las instrucciones que alteran Z y N: **ADD, SUB, CMP, DIV, MUL, AND, OR, NOT**. Las 3 primeras además calcula C y V.

Nota1: El caracter % denota el cociente de la división entera.

Nota2: El resultado de la operación **MUL** ocupa 32 bits, almacenándose los 16 bits menos significativos en el operando destino y los 16 bits mas significativos en el registro **R7**.

Nota3: Donde **rrr** es una codificación (en 3 bits) del número de registro.

Nota4: Con este formato, los campos **y** contienen valores constantes (si el modo respectivo es *inmediato*), contienen direcciones de memoria principal (si el modo es *directo*), o no existen (si el modo respectivo es *registro*).

Nota5: Saltos condicionales tiene como formato codop - desplazamiento donde los primeros 4 bits del campo codop corresponde a la cadea 1111. Las instrucciones en este formato son de la forma Jxx(salto relativo condicional). Si al evaluar la condicion de salto en los Flags el resultado es 1, el efecto es incrementar al PC con el valor de los 8 bits de desplazamiento, representado en complemento a 2 de 8 bits. En caso contrario la instruccion no hace nada.

3.2. Modos de direccionamiento

1. **Inmediato** Codigo de modo de direccionamiento: 000000 Este modo de direccionamiento es un valor que será utilizado como modo origen pero nunca como modo destino ya que en el no pueden guardarse datos.

Ejemplos: **0x0000** representa el modo de direccionamiento inmediato cuyo valor es cero. **0x000F** representa el modo de direccionamiento inmediato cuyo valor es 15.

2. **Directo** Codigo de modo de direccionamiento: 001000 Este modo de direccionamiento representa una direccion de memoria o de puertos que será utilizado como modo origen pero nunca como modo destino ya que en el no pueden guardarse datos.

Ejemplos: **[0x0000]** representa el modo de direccionamiento directo cuyo valor se encuentra en la celda de memoria que identifica unequivocamente a la dirección **0x0000**, es decir 0. **[0x000F]** representa el modo de direccionamiento directo cuyo valor se encuentra en la celda de memoria que identifica unequivocamente a la dirección **0x000F**, es decir 15.

3. **Indirecto** Codigo de modo de direccionamiento: 011000 Este modo de direccionamiento representa una direccion de memoria o de puertos que será utilizado como modo origen pero nunca como modo destino ya que en el no pueden guardarse datos.

Ejemplos: **[[0x0000]]** representa el modo de direccionamiento indirecto cuyo valor se encuentra en la celda de memoria cuya dirección esta guardada como dato en la celda de memoria que identifica unequivocamente a la dirección **0x0000**, es decir 0. **[[0x000F]]** representa el modo de direccionamiento indirecto cuyo valor se encuentra en la celda de memoria cuya dirección esta guardada como dato en la celda de memoria que identifica unequivocamente a la dirección **0x000F**, es decir 15.

4. **Registro** Código de modo de direccionamiento: 100rrr Donde rrr.^{es} una cadena binaria que puede tomar los valores desde 000 hasta 111 para identificar según el numero decimal que representen los diferentes ocho registros R0..R7.

Ejemplos: **R0** representa el modo de direccionamiento registro cuyo valor se encuentra en el registro **R0**. **R7** representa el modo de direccionamiento registro cuyo valor se encuentra en el registro **R7**.

5. **RegistroIndirecto** Código de modo de direccionamiento: 110rrr Donde rrr.^{es} una cadena binaria que puede tomar los valores desde 000 hasta 111 para identificar según el numero decimal que representen los diferentes ocho registros R0..R7.

Ejemplos: **[R0]** representa el modo de direccionamiento registro cuyo valor se encuentra en la direccion de memoria que esta guardada como dato el registro **R0**. **[R7]** representa el modo de direccionamiento registro cuyo valor se encuentra en la direccion de memoria que esta guardada como dato el registro **R7**.

3.3. Repertorio de instrucciones

3.3.1. Instrucciones de 2 operandos

El formato de instruccion de dos operandos es el siguiente:

CodOp + Modo destino + Modo origen + Destino + Origen (4 bits) (6 bits)
(6 bits) (16 bits) (16 bits)

Donde el operando destino no puede ser un modo de direccionamiento Inmediato, ya que en la direccion de memoria o registro que describa dicho operando es donde se guardará el resultado de la operación.

1. **MUL destino, origen** Código de operación: 0000 Esta instrucción describe la multiplicación entre los datos de los dos operandos. Esta operación es la única que cuyo resultado puede ser 32 bits, que son dos celdas de memoria en hexadecimal, por lo que los primeros 16 bits, es decir, la primer mitad, es guardada en el registro R7 y la segunda en el operando destino.
2. **ADD destino, origen** Código de operación: 0010 Esta instrucción describe la suma entre los datos de los dos operandos. El resultado de la ejecución de la suma es guardado en el operando destino.
3. **SUB destino, origen** Código de operación: 0011 Esta instrucción describe la resta entre los datos de los dos operandos. El resultado de la ejecución de dicha resta es guardado en el operando destino.
4. **DIV destino, origen** Código de operación: 0111 Esta instrucción describe la división entre el dato en el operando destino como dividendo y el dato en el operando origen como divisor. El resultado de la ejecución de la división es guardado en el operando destino.
5. **MOV destino, origen** Código de operación: 0001 Esta instrucción describe la copia de datos del dato alojado en el operando origen al operando

destino. El resultado de la ejecución del MOV es el dato guardado en el operando origen ahora también guardado en el operando destino.

6. **AND destino, origen** Código de operación: 0100 Esta instrucción describe la operación lógica "bit a bit" entre los datos de los dos operandos. El resultado de la ejecución de esta operación es guardado en el operando destino.
7. **CMP destino, origen** Código de operación: 0110 Esta instrucción describe la comparación bit a bit entre los datos de los dos operandos. El resultado de esta operación es solamente la actualización de flags en la cpu.
8. **OR destino, origen** Código de operación: 0101 Esta instrucción describe la operación lógica "bit a bit" entre los datos de los dos operandos. El resultado de la ejecución de esta operación es guardado en el operando destino.

3.3.2. Instrucciones de 1 operando origen

El formato de instrucción de un operando origen es el siguiente:

CodOp + relleno + Modo origen + Origen (4 bits) (000000) (6 bits) (16 bits)

1. **CALL origen** Código de operación: 1011 El efecto del CALL es guardar la dirección de memoria en la celda de la dirección que se encuentra guardada en el SP (Stack pointer) aumentar el SP y guardar en el PC (Program Counter) el dato que se encuentra guardado en el operando origen ya que describe el llamado a una subrutina que comienza en la celda de memoria cuya dirección esta guardada en el operando origen.
2. **JMP origen** Código de operación: 0110 El efecto del JMP es cambiar el PC (Program Counter) por el dato que esta guardado en el operando origen ya que esta operación describe el salto a otra parte de la memoria para continuar con la ejecución del programa.

3.3.3. Instrucciones de 1 operando destino

El formato de instrucción de un operando destino es el siguiente:

CodOp + Modo origen + relleno + Origen (4 bits) (6 bits) (000000) (16 bits)

1. **NOT destino** Código de operación: 1001 Esta instrucción describe la operación lógica "negación" bit a bit en el datos del operando destino. El resultado de la ejecución de esta operación es guardado en la misma celda o registro de donde es leído el dato inicialmente.

3.3.4. Instrucciones sin operandos

El formato de instrucción sin operandos es el siguiente:

CodOp + relleno (4 bits) (00000000000000)

1. **RET** Código de operación: 0110 El efecto del ret es cambiar el pc por el dato que esta guardado en la celda de memoria que se encuentra en el SP (Stack pointer) y decrementar el SP ya que describe la finalización de la ejecución de una subrutina y la ejecución del resto del programa.

3.3.5. Instrucciones de salto condicional

El formato de instrucción de salto condicional es el siguiente es el siguiente: prefijo + CodOp + desplazamiento (1111) (4 bits) (8 bits)

El efecto de cualquier salto condicional es aumentar el PC (Program Counter) en la cantidad de celdas que indique el desplazamiento si sólo la condición que cada salto condicional tiene da como resultado 1, lo cual es interpretado como verdadero.

1. **JE desplazamiento** Código de operación: 0001 La condición del salto es que el flag **Z** (Cero) sea 1, es decir la ultima operación matemática dió como resultado el número cero.
2. **JNE desplazamiento** Código de operación: 1001 La condición del salto es que el flag **Z** (Cero) sea 0, es decir la ultima operación matemática no dió como resultado el número cero.
3. **JLE desplazamiento** Código de operación: 0010 La condición del salto es el resultado de la siguiente operación lógica **Z OR (N XOR V)**, es decir la ultima operación matemática es menor o igual con signo.
4. **JG desplazamiento** Código de operación: 1010 La condición del salto es el resultado de la siguiente operación lógica **NOT (Z OR (N XOR V))**, es decir la ultima operación matemática es mayor con signo.
5. **JL desplazamiento** Código de operación: 0011 La condición del salto es el resultado de la siguiente operación lógica **N XOR V**, es decir la ultima operación matemática es menor con signo.
6. **JGE desplazamiento** Código de operación: 1011 La condición del salto es el resultado de la siguiente operación lógica **NOT (N XOR V)**, es decir la ultima operación matemática es mayor o igual con signo.
7. **JLEU desplazamiento** Código de operación: 0100 La condición del salto es el resultado de la siguiente operación lógica **C OR Z**, es decir la ultima operación matemática es menor o igual sin signo.
8. **JGU desplazamiento** Código de operación: 1100 La condición del salto es el resultado de la siguiente operación lógica **NOT (C OR Z)**, es decir la ultima operación matemática es mayor sin signo.
9. **JCS desplazamiento** Código de operación: 0101 La condición del salto es que el flag **C** sea 1, es decir la ultima operación matemática es menor sin signo.
10. **JNEG desplazamiento** Código de operación: 0101 La condición del salto es que el flag **N** sea 1, es decir si el último resultado de una operación dio negativo.

11. **JVS desplazamiento** Código de operación: 0111 La condición del salto es que el flag **V** sea 1, es decir si el último resultado de una operación dio overflow.

3.4. Versiones de la arquitectura (SUSI)

hacer un grafico copado donde se vea las capas de cebolla

Las versiones de la arquitectura Q están pensadas para incorporar funcionalidades de manera que la curva de aprendizaje sea adecuada para los alumnos, siendo paulatina e incremental, es decir, cada arquitectura Qi agrega más funcionalidad (ya sean instrucciones nuevas o modos de direccionamiento) a las arquitecturas Qi anteriores.

Incorporar lo que sigue a un gráfico de la estructura en cebolla²

3.4.1. Q1

Modos de direccionamiento

1. Inmediato
2. Registro

Instrucciones

1. MOV
2. SUB
3. DIV
4. ADD
5. MUL

3.4.2. Q2

Modos de direccionamiento

1. Modos de direccionamiento **Q1**
2. Directo

Instrucciones

1. Instrucciones **Q1**

3.4.3. Q3

Modos de direccionamiento

1. Modos de direccionamiento **Q2**

Instrucciones

1. Instrucciones **Q2**
2. CALL
3. RET

²

3.4.4. Q4**Modos de direccionamiento**

1. Modos de direccionamiento **Q3**

Instrucciones

1. Instrucciones **Q3**
2. CMP
3. JMP
4. JE
5. JNE
6. JLE
7. JG
8. JL
9. JGE
10. JLEU
11. JGU
12. JCS
13. JNEG
14. JVS

3.4.5. Q5**Modos de direccionamiento**

1. Modos de direccionamiento **Q4**
2. Indirecto
3. RegistroIndirecto

Instrucciones

1. Instrucciones **Q4**

3.4.6. Q6**Modos de direccionamiento**

1. Modos de direccionamiento **Q5**

Instrucciones

1. Instrucciones **Q5**
2. AND
3. OR
4. NOT

4. Estado del arte

En un principio la materia Organizacion y Arquitectura eligio el simulador intel 8085 para enseniar assembler.

- **Simulador intel 8085**

Es un simulador y ensamblador para el microprocesador Intel 8085, que ofrece una interfaz que a simple vista es dificil de entender, una parte para la ejecucion y otra parte para ver el estado de la memoria y registros. Las opiniones y experiencias con este simulador dentro de la materia no fueron muy buenas, fuera del gran esfuerzo que pusieron los profesores en que aprendieramos, dandonos mini tutoriales no resultaron como se esperaba. (Hmm si mi memoria no me falla se utilizo en un cuatrimestre [Mara si me equivoco corregime ehh!]). Todo lo que se trabajaba en ese simulador tambien se escribia a en hoja.

- En los primeros cuatrimestre la materia eligio incorporar la **Arquitectura QARQ** con su especificacion y ademas un lenguaje micro –(no me acuerdo el nombre). En estos tiempos se desistio en utilizar un simulador complejo pero no dejaron de buscar e investigar alguna alternativa que pueda convivir con los estudiantes. Todos los ejercicios se hacian en papel, la unica forma de verificar resultados es atraves de los profesores.

- Hace poco la materia o mejor dicho el equipo de Orga decidio dividir la especificacion de la arquitectura QARQ en varias partes, donde cada una recibe el nombre de **Arquitectura Q**.

Cada Arquitectura Qi agrega una nueva funcionalidad (instrucciones o modos de direccionamientos) que la anterior, son como capas de cebolla. Ademas incorporaron un trabajo practico donde tiene que programar un programa que monocromatice una image en la **Arquitectura IA - 32**. En el enunciado se explica bastante como funciona esta arquitectura. Sin embargo todos los ejercicios de las arquitecturas Qi se siguen haciendo en papel.

- **Hasta ahora!!!** La idea es incorporar el **Simulador Qsim** en la materia para que los chicos puedan entender el ciclo de instruccion en la etapa de ejecucion de un programa, poder ver que es lo que estan haciendo, chequear el efecto de los programas creados, jugar y divertirse aprendiendo :P (lo siento estoy re quemada). Pensamos que nuestro simulador va a encargar en la materia ya que cubre las necesidades basicas de un simulador. :D

Parte II

Simulador QSim

5. Funcionalidad del simulador

La funcionalidad del simulador puede caracterizarse mediante las siguientes partes importantes:

- Chequeo de sintaxis de los programas escritos en el lenguaje Q
- Ensamblado del código fuente de un programa en su correspondiente código máquina
- Cargado en memoria del código máquina
- Ejecución paso a paso de un programa cargado en memoria

5.1. Chequeo de sintaxis

El simulador provee al alumno de un editor de texto en el cual escribirá el programa en un lenguaje Q_i , que desea cargar en memoria y ejecutar. Una vez que el usuario haya terminado la escritura, al momento de cargar el programa, el simulador utilizará un parser para detectar errores de sintaxis, tales como la falta de una coma o un corchete, o la presencia de símbolos que no pertenecen al lenguaje (como por ejemplo signos de pregunta y símbolos matemáticos); o bien errores semánticos como la combinación incorrecta de elementos del lenguaje, por ejemplo: modos de direccionamiento mal ubicados. El parser solo revisará lo escrito por el alumno y de acuerdo a las gramática del lenguaje, mostrará alguno de los siguientes estados:

OK Este mensaje se obtiene cuando no hubo ningún error de sintaxis. Si se da este resultado, es posible continuar con el ensamblado y cargado en memoria.

SyntaxError Este mensaje de error se obtiene cuando en alguna línea del programa se detectó algún error de sintaxis o de semántica, como se describió arriba. Cuando ocurre este error se lo acompaña con una descripción lo mas detallada posible para que el alumno detecte donde ocurrió y pueda corregirlo. Un programa con errores no puede ser ensamblado y cargado en memoria.

5.2. Ensamblado

Una vez que el programa es sintácticamente válido es posible traducir el código fuente del programa en código máquina (representado en cadenas binarias). Para esto se respeta un formato de instrucción que indica cómo se codifica cada operación y los operandos.

[Mas detalle al respecto de este proceso en la sección de ...](#)³

5.3. Cargado en memoria

Una vez ensamblado, la representación binaria (o código máquina) del programa será cargado en memoria a partir de una ubicación (celda de memoria) que el alumno puede elegir. Esto permite visualizar el contenido de la memoria (con el programa cargado) y el estado de los registros de la CPU. La decodificación con desensamblado permite al alumno experimentar otros escenarios y efectos laterales, entre los cuales podemos enumerar:

³Poner vínculo a donde se ponen ejemplos de Q_i

- Si la ejecución paso a paso excede los límites del programa, pueden tomarse instrucciones de otra rutina y procesarse como una nueva instrucción.
- Si en cambio, se intenta ejecutar el contenido de una celda con datos (y no una instrucción) podrá ocurrir que se encuentre una instrucción inválida (combinación de modos, códigos, incorrecta) y el alumno verá el mensaje de error pertinente.

Durante la carga del programa en memoria puede producirse un `OutOfMemoryError?`, que significa que el programa no entra en la ubicación elegida en memoria ya que ocupa más celdas que las que se encuentran disponibles debajo de ella ya que como se mencionó en la sección 3.1, la memoria disponible tiene un tamaño limitado y por este motivo la asignación en memoria del código máquina puede exceder el espacio disponible a partir de la celda inicial anteriormente elegida. Si por el contrario, no se produce este error, el alumno podrá ver el programa cargado en memoria exitosamente.

5.4. Ejecución paso a paso

Se provee la funcionalidad de la ejecución paso a paso ya que se desea que el alumno pueda experimentar y así comprender los pasos del ciclo de ejecución. Además puede ejercitarse situaciones que se denominan "errores conceptuales de programación" Esto es a lo que llamamos Errores conceptuales, entre los cuales es posible mencionar:

- Tomar un dato de un sector de memoria equivocado.
- Que el programa sobrescriba su mismo código máquina.
- Permitir que la ejecución continúe una vez terminado el programar cargado en memoria.
-

El paso a paso que provee el simulador consiste en las siguientes etapas pertenecientes al ciclo de instrucción:

1. **Búsqueda de instrucción:** El alumno podrá visualizar el valor que contiene PC (Program counter) donde se encuentra la dirección de la celda en memoria que contiene la próxima instrucción a ejecutar (por ejemplo, en caso de ser la primera instrucción del programa recién cargado, el pc tendrá la dirección de memoria elegida por el alumno para iniciar el cargado del programa en memoria). El simulador, toma de la memoria el código máquina correspondiente a la instrucción que comienza en esa dirección tomada de PC (una instrucción puede ocupar más de una celda de memoria) y los guarda en el registro IR (*Instruction Register*). Será observable también para el alumno el incremento del registro PC, tantas como celdas ocupe la instrucción actual, lo que conceptualmente es, preparar el contexto de ejecución para tomar la siguiente instrucción.
2. **Decodificación:** En la decodificación el Interpretador se encarga de ensamblar el código máquina (abreviado en hexadecimal) que ya fue ubicado

en el registro IR para mostrar el código fuente de la instrucción actual con sus respectivos operandos. Si el programa escrito por el alumno es sintacticamente y conceptualmente correcto, este paso le permite comprobar que la instrucción actual es la que él mismo escribió y no otra, visualizándola en pantalla. En esta etapa se provee también la oportunidad de que el alumno aprecie otros conceptos, tales como los errores conceptuales mencionados antes.

3. **Ejecución** El execute ejecuta los efectos de la instrucción y muestra en pantalla los cambios en el estado de ejecución: memoria, puertos, registros y flags. Dentro de esta misma etapa se lleva a cabo el almacenamiento de resultados que, cuando sea necesario, guardará el valor resultante de la operación descrita por la instrucción en el operando destino. Esto cambiará el valor de una celda de memoria o de un registro y será visto en pantalla por el alumno.

6. Implementación

6.1. Tecnología utilizada (tati)

- **Scala** Elegimos el lenguaje Scala para realizar el simulador ya que, en la materia llamada Objetos III, lo utilizamos un pequeño lapso de tiempo y creímos que era una buena oportunidad para, en vez de elegir un lenguaje que hayamos utilizado más en la carrera como java, profundizar en la utilización de Scala y aprovechar las ventajas que este ofrecía al combinar el manejo de objetos y las características de un lenguaje funcional.
- **Arena** Utilizamos el framework Arena para realizar la interfaz de usuario del simulador porque es un framework de código abierto, que también pudimos utilizar en una de las materias y al poder ser combinado con Scala nos pareció una buena oportunidad de explotar lo que nos ofrecía para que este simulador en su totalidad sea de código abierto.
- **Eclipse** Se eligió utilizar el entorno de programación Eclipse ya que ambas trabajamos en diferentes sistemas operativos para los cuales Eclipse es funcional y puede tener los mismos plugins que hace que pueda soportar proyectos MVN y Scala, además de que, al ser el entorno de programación que más hemos usado nos sentíamos cómodos con esa elección.
- **Git** Elegimos git como repositorio externo para sincronizar nuestro código ya que promueve el código abierto (ACA QUIERO PONER LO DE QUE ESTA AHI ARRIBA Y NO LE CEDEMOS DERECHOS NI NADA A NADIE PERO NO ME SALE!!!)

6.2. (arq OO)

6.2.1. ALU

Como se observa en la figura 1 la ALU tiene toda la responsabilidad en la ejecución de operaciones matemáticas y lógicas, además del análisis sobre los flags luego de cada operación.

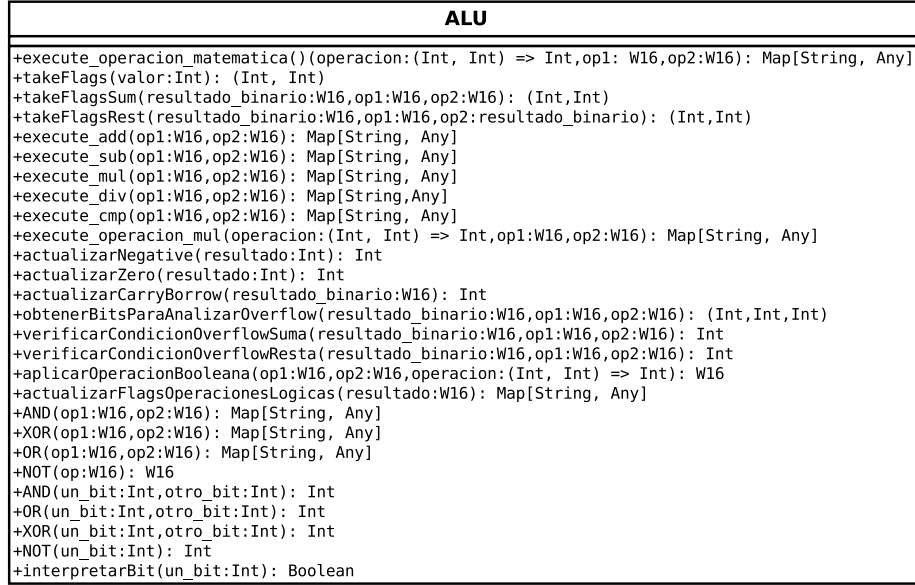
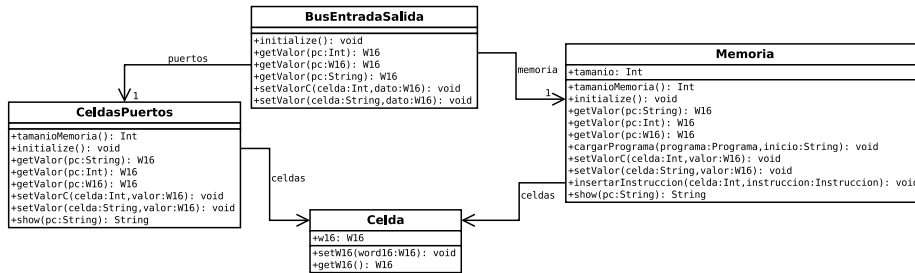


Figura 1: Diagrama de clase de la Unidad Aritmético-Lógica

6.2.2. Bus de entrada y salida, memoria y puertos

Como se observa en la figura 2 el Bus de entrada y salida tiene la responsabilidad de derivar según donde corresponda (Memoria o Puertos) la modificación de una celda o el leer un dato. Para ello conoce a una instancia de la clase Memoria y a otra de la clase CeldasPuertos. Ambas clases conocen muchas instancias de la clase Celda, y cada Celda a su vez conoce un dato: una instancia del W16, que representa a los datos guardados en memoria o en los puertos.

Figura 2: Diagrama de clase de la *Memoria* y *CeldasPuertos*

6.2.3. CPU

Como se observa en la figura 3, la CPU conoce a la ALU, contiene los registros IR y PC, y los flags (V,Z,C,N) y los ocho registros (R0...R7). La responsabilidad de la CPU es actualizar los flags, los registros, actualizar el PC y el IR, y ser la conexión con la ALU.

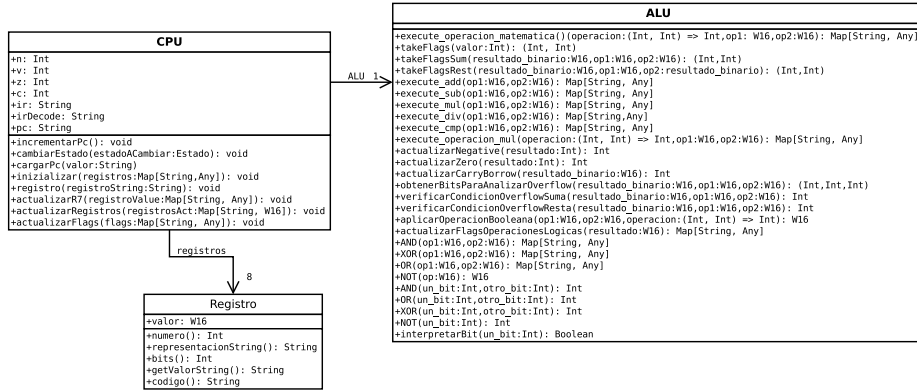


Figura 3: Diagrama de clase de la

6.2.4. Interprete

Como se observa en la figura 4 el Interprete es un singleton tiene la entera responsabilidad de recibir los datos leídos desde, una celda de memoria hasta tres celdas, y bit a bit, según los códigos de operación y de modo de direccionamiento saber de que instrucción se trata y cuales son sus operandos si es que los tiene y devolver el objeto que la representa. Se ocupa del decode de la instrucción.

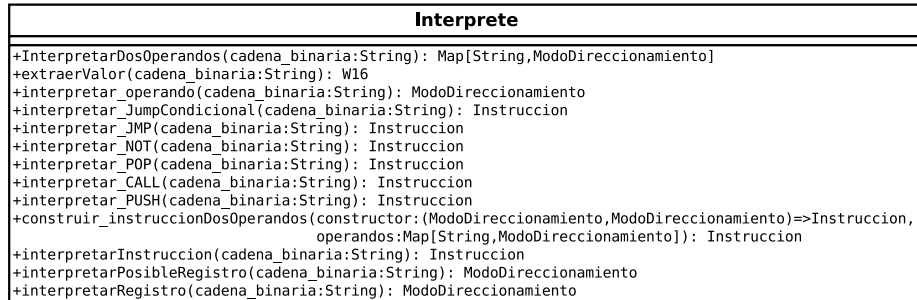


Figura 4: Diagrama de clase de la

6.2.5. Modos de direccionamiento y W16

Como se observa en la figura 5 los modos de direccionamiento extienden del trait ModoDireccionamiento, donde se encuentran declarados mensajes necesarios para majear todas las subclases poliformicamente. Entre los cuales se encuentran los mensajes:

- **representacionString** Que devuelve la representación en string como codigo fuente, por ejemplo la representación de un ADD(R0,R7), sería: ADD R0, R7
- **codigo** Retorna el string que representa al código único de modo de direccionamiento, por el ejemplo, el código de modo de direccionamiento del R7 es 100111.

- **getValorString** Retorna el string que representa al dato que posee el modo de direccionamiento. En el caso de un Inmediato que sea FF56, devolvera el string "FF56", y en el caso de cualquier registro, retornara el valor que represente su W16.

Los modos de direccionamiento diferentes a Inmediato y Registro, conocen en vez de un W16 otro modo de direccionamiento según corresponda:

- RegistroIndirecto conoce un a instancia de Registro.
- Directo conoce un a instancia de Inmediato.
- Indirecto conoce un a instancia de Directo.

Esto se implemento de esta manera para que el leer datos de memoria, puertos o registros, o guardarlos en los mismos sea más sencillo ya que se delega en el modo de direccionamiento que conoce.

La clase Etiqueta representa la etiqueta creada por el alumno cuando realiza el programa. Cuando el mismo es cargado en memoria, según cual sea la celda de inicio y cuanto ocupen las instrucciones, se calcula la dirección de memoria que representa la etiqueta y luego se descarta reemplazandola por un modo de direccionamiento Inmediato.

La clase W16 que también esta en la figura 5, representa el dato que es guardado en memoria. Tiene la responsabilidad de incrementarse, decrementarse, sumar una entero, devolver su representación binria y su valor en entero.

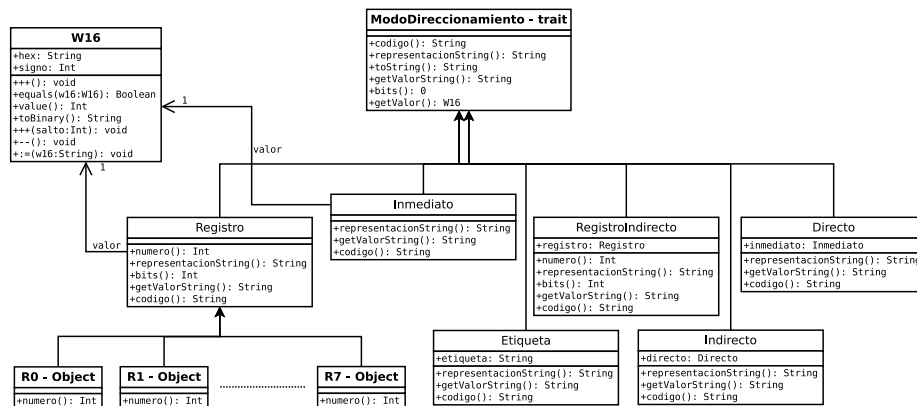


Figura 5: Diagrama de clase de la

6.2.6. Instrucciones

Como se observa en la figura 6 las Instrucciones estan modeladas en jerarquias que se dividen en:

- Un operando.
- Dos operandos.
- Sin operandos.

- JumpConducional.

Se realizo dicha jerarquia para permitir la fácil agregación de nuevas instrucciones siendo subclases de la clase que corresponda ya que comparten comportamiento tal como manera de mostrarse y de codificarse.

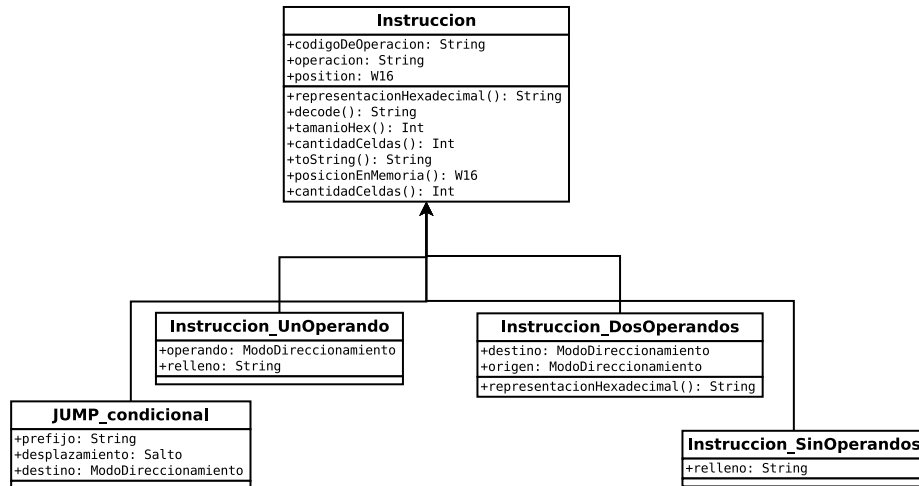


Figura 6: Diagrama de clase de la

Instrucciones sin operandos

Como se observa en la figura 7 la unica Instruccion sin operandos implementada es la Instruccion RET, a pesar de esto, se eligio hacer una jerarquia para que luego se facilite la inserción al modelo de nuevas instrucciones sin operandos.

Instrucciones de un operando

Como se observa en la figura 8 y concordando en lo mencionado sobre las Instrucciones anteriormente, puede haber dos tipos de instrucciones de un operando:

- Un operando origen.
- Un operando destino.

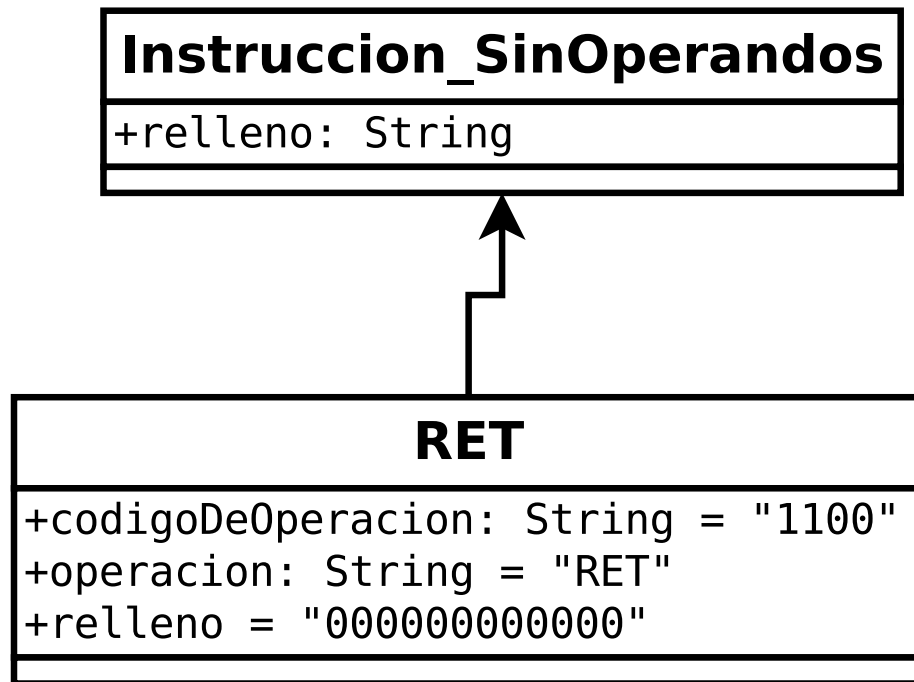
Ambas clases de instrucciones tienen un solo operando. La diferencia entre ellas es la logica de ejecución y la manera en la que se ensamblan y desensamblan ya que sus formato de instruccion difiere en donde esta colocado el relleno.

Instrucciones de dos operandos

Como se observa en la figura 9 la jerarquia de clases de instrucciones de dos operandos es la más amplia por tener mayor cantidad de instrucciones. Todas comparten la manera de decodificación e interpretación, además de la logica de impresión.

6.2.7. Programa

Como se observa en la figura 10 la clase Programa conoce un grupo de Instrucciones (las instrucciones que lo componen). Las instancias son creadas por el parser, luego se calculan las etiquetas (si es que las tiene) y finalmente

Figura 7: Diagrama de clase de la *sinOperandos*

cuando es cargado en memoria la instancia de programa es desechada ya que no vuelve a usarse en ningún momento de la ejecución.

6.2.8. Simulador

Como se observa en la figura 11 la clase Simulador conoce a una instancia de CPU, a una instancia de BusEntradaSalida y a una instancia de Instruccion que es la instruccion que se esta ejecutando en ese momento. La clase Simulador tiene la responsabilidad de obtener el codigo maquina a traves del Interprete de la siguiente instruccion (fetch y decode), calcular las etiquetas de un programa, cargar el programa en memoria, asi mismo, los datos en registros, ejecutar las instrucciones o delegar su ejecución a la ALU que conoce a traves de la CPU según corresponda (realizar el execute), y guardar datos en memoria o registros (Store). Es la clase principal del modelo y la encargada de coordinar la ejecución del programa paso a paso.

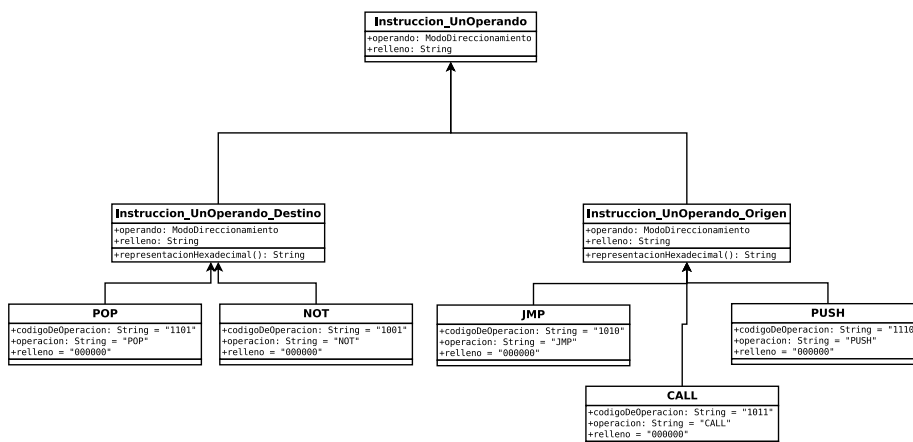


Figura 8: Diagrama de clase de la *UnOperando*

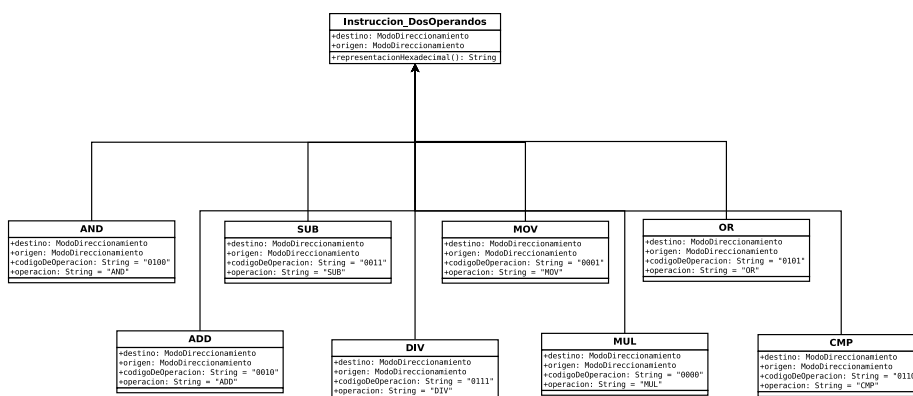


Figura 9: Diagrama de clase de la *posOperandos*

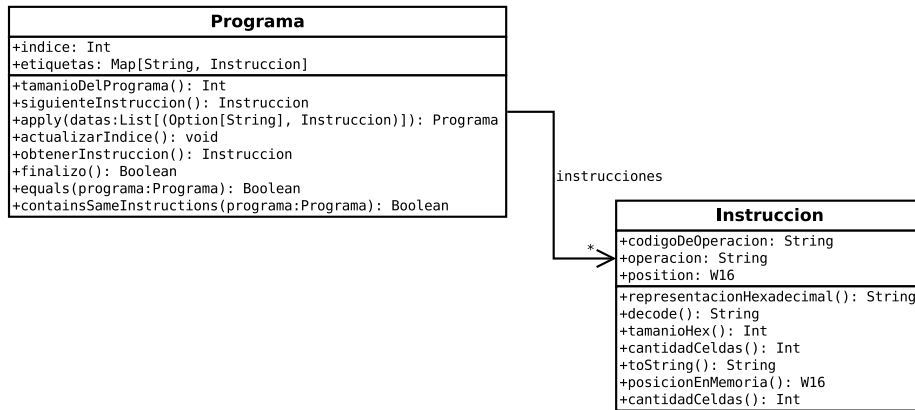


Figura 10: Diagrama de clase de la

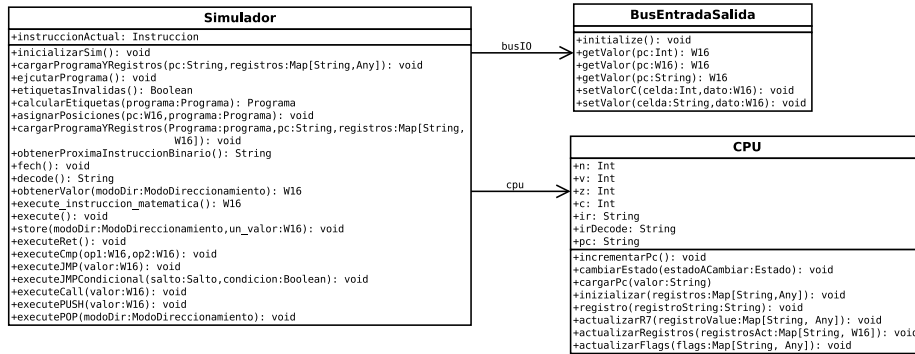


Figura 11: Diagrama de clase de la

7.2. Dificultades de diseño

En primera instancia se optó por implementar un modelo de objetos que utilizaba un objeto de la clase **Programa** a lo largo de toda la ejecución. Procesaba las instrucciones no leyendo de la matriz memoria, si no, pidiendo la siguiente instrucción al objeto instancia de la clase Programa, sobreviviendo así las distintas etapas una vez que fue creado y evitando la creación de un objeto cuya responsabilidad sea interpretar el código máquina alojado en la memoria, evitando la nueva creación de instancias de la clase **Instruccion** y haciendo mucho más sencillo al modelo ya que la memoria era sólo una clase que contenía datos que se reflejaban en pantalla y no se extraían datos de celdas en ningún momento. Luego entendimos que un programa no sólo podía modificar su entorno al ser ejecutado (otras celdas de memoria que no ocupen su código máquina, celdas de puertos, registros, etc) si no que también podría sobrescribir su código máquina (ya sea con ese propósito o sólo por un error conceptual), o bien, que el alumno debía tener la posibilidad de seguir ejecutando más allá del código máquina alojado en memoria o más, inevitablemente se necesitó corregir gran parte del modelo agregando una clase denominada **Intérprete**, cuya responsabilidad es interpretar la siguiente instrucción alojada en memoria para que luego sea ejecutada, otorgando más responsabilidad a la clase **Memoria** y descartando el

objeto instancia de **Programa** una vez que éste es cargado en memoria con éxito.

Tuvieron que ser solicitadas además extensiones al equipo de desarrolladores de Arena para poder realizar la interfaz con dicho framework.

- FileSelector para que los alumnos puedan cargar el archivo .Qsim en el cual se encuentra su programa
- CodeEditor (O actualmente llamado KeywordTextArea). Es el espacio donde se visualiza el programa QUARQ que realizaron los alumnos una vez que esta cargado.
- Bindings contra el background de componentes y celdas de una tabla, para permitir la visualizacion de cambios de colores en la memoria luego de las etapas de ejecución y de realizar algun cambio en la misma.
- TextBox multiLine (TextArea) con scroll para ser utilizado como consola de devolucion.
- Icono para la aplicacion, solo por cuestiones estéticas.

8. Casos de prueba

Descripcion de cada Caso de Prueba

8.1. Chequeo de sintaxis en la distintas Qi

Para chequear la sintaxis de cada **Arquitectura Q**, primero escribimos un programa Qi en un archivo con extension .qsim luego este archivo lo recibe el objeto Parser que se encargara de chequear la sintaxis. Para eso realizamos dos casos de prueba por cada Qi:

1. Chequear programa Qi valido

Agarramos el programa Qi valido y se lo pasamos al Parser, cuando termina de chequear devuelve un resultado, como es de esperar devuelve un objeto programa con la lista de instrucciones. Tomamos ese resultado y como ultimo paso comparamos si el programa resultado es igual al programa esperado.

(esto va en nota al pie)programa Qi valido = Es un programa en cualquier arquitectura Q escrito sintaticamente correcto.

2. Chequear programa Qi es invalido

En este caso escribimos un programa Qi con errores de sintaxis para ver como reacciona el parser.

Agarramos el programa Qi invalido y se lo pasamos al Parser, cuando termina de chequear devuelve un error, como es de esperar devuelve una Exception **SyntaxErrorException** porque dentro del archivo del programa hay una linea que tiene errores de sintaxis, el mensaje de la excepcion te muestra el numero de linea y la propia linea para que veas en donde te

equivocaste. Capturamos la Excepcion y tomamos como resultado el mensaje y como ultimo paso comparamos el mensaje resultado con el mensaje esperado.

8.2. Ensamblado (Objetos \rightarrow Código Máquina)

Para poder verificar que el Ensamblado se realizo correctamente tomamos el programa Qi mencionado anteriormente y como sabemos que el resultado del parser nos devuelve un objeto programa que contiene un conjunto de objetos instrucciones, el ensamblado simplemente lo realiza cada instruccion, esto quiere decir que cada instruccion sabe ensamblarse. De esta forma realizamos el ensamblado del programa y obtenemos una lista de codigo maquina como resultado y como ultimo paso comparamos la lista de codigo maquina resultado con la lista codigo maquina esperado.

8.3. Decodificacion Código Máquina \rightarrow Objetos (Interprete)

La Decodificacion se verifica correctamente a la inversa del Ensamblado tomando el ensamblado del programa Qi que como sabemos es la lista de codigo maquina, la cual se itera para pasar cada elemento al objeto Interprete que es el encargado de Decodificar. Como resultado puede devolver dos cosas:

1. Lista de de Instrucciones (decodificacion)

El resultado correcto es la lista de objetos instrucciones, por cada codigo maquina el interprete verifica los codigos de operacion para poder crear las instrucciones correctas y como ultimo paso de la verificacion comparamos la lista resultante con la lista esperada.

2. Error

El resultado puede ser un error porque puede pasar que el codigo de operacion o el codigo de algun modo de direccionamiento dentro del formato de cada instruccion sea invalido. La excepcion que puede tirar tiene el nombre de **CodigoInvalidoException**.

8.4. Ejecucion

Para verificar que la ejecucion de un programa Qi con el **Ciclo de Instruccion** se realice correctamente lo que hacemos es cargar el programa en la memoria a traves del objeto **Simulador**, de esta forma cuando se encuentra cargado esta listo para la ejecucion. Los pasos que se verifican son 3:

■ Fetch

Para verificar el fetch se toma la instruccion siguiente a ejecutar y se compara el valor que se guarda en el registro ir con el valor esperado. Ademas se verifica que el registro especial pc apunte a la siguiente instruccion a ejecutar.

■ Decode

Para verificar el decode, se toma el valor del registro ir para que lo reciba el interprete que dara como resultado el objeto instruccion correspondiente.

Este objeto instruccion sabe mostrarse quien es y por eso la responsabilidad de decodificar se la delega a la instruccion. Al obtener la decodificacion de cada instruccion podemos comparar la decodificacion resultante con la esperada.

■ **Execute**

Para verificar el efecto de cada instruccion se considero armar un caso de prueba por cada instruccion donde cada uno tiene un estado inicial y luego de realizar la ejecucion de dicha instruccion termina con un estado final modificado en el cual verificamos el resultado obtenido con el resultado esperado.

Las siguiente secciones son pasos que se realizan de acuerdo al efecto de la instruccion

8.4.1. Operaciones de ALU

La ALU es la encargada de ejecutar operaciones aritmeticas/logicas. Si la instruccion a ejecutar tiene un efecto aritmetico/logico se le delega la ejecucion de la operacion. Para poder verificar todas las operaciones creamos un caso de prueba por cada operacion. Tanto las operaciones aritmeticas como las logicas se analizan de la misma forma:

1. Se toma cada valor de los operandos(la busqueda se detalla en la siguiente instruccion)
2. Se realiza la operacion aritmetica/logica.
3. Se verifica el resultado obtenido con el resultado esperado.
4. Algunas operaciones (la mayoria) modifican los flags. la ALU tiene mucha relacion con los flags ya que por cada cuenta realizada actualiza los mismos. Estos tambien se se verifican comparando el resultado obtenido con el valor esperado.

8.4.2. Busqueda de Operandos (para los distintos modos de direccionamiento)

Para verificar la busqueda de operandos, tomamos una instruccion cualquiera y probamos todas las combinaciones de modos de direccionamiento para cada operando, teniendo en cuenta que son invalidas todas las combinaciones de 1/2 operando/s que tenga el operando destino el modo de direccionamiento Inmediato. Al tener la instruccion elegida procedemos a obtener el valor de cada operando, este valor se lo compara con el valor esperado.

8.4.3. Almacenamiento de Operandos

Para verificar el almacenamiento de operandos, salvo algunas instrucciones todas pasan por la etapa de store. Tomamos cada una de las instrucciones de un programa Qi y realizamos el ciclo de ejecucion mencionado anteriormente. Cuando estamos en la etapa de ejecucion, cada una tiene un operando destinado a guardar el resultado de su efecto. En esta instancia sabemos que dicho operando tiene el valor guardado, a este valor lo comparamos con el valor esperado.

9. Ejemplos de uso

Referencias

- [1] Williams Stallings, *Computer Organization and Architecture*, octava edición, Editorial Prentice Hall, 2010.
- [2] A. Tanenbaum, *Organización de Computadoras*, cuarta edición, Editorial Pearson.
- [3] Hennessy, Patterson. *Arquitectura de Computadores - Un enfoque cuantitativo*, primera edición, Editorial Mc Graw Hill.
- [4] Sitio oficial de la materia Organización de Computadoras: <http://orga.blog.unq.edu.ar> (2013)