

Programmazione Orientata agli Oggetti

Appunti delle Lezioni di Programmazione Orientata agli Oggetti

Anno Accademico: 2023/24

Giacomo Sturm

*Dipartimento di Ingegneria Civile, Informatica e delle Tecnologie Aeronautiche
Università degli Studi “Roma Tre”*

Sorgente del file LaTeX disponibile al seguente link:

<https://github.com/00Darxk/Programmazione-Orientata-agli-Oggetti>

Indice

1	Introduzione al Linguaggio Java	1
1.1	Rapporto con il Linguaggio C	1
1.2	Paradigma Orientato agli Oggetti	2
1.3	Oggetti e Riferimenti	3
1.4	Gestione della Memoria	4
1.5	Costruttori, Stringhe e Array	5
2	Qualità del Codice	8
2.1	Librerie	8
2.2	Coesione e Accoppiamento	9
2.3	Tecniche di Testing	10

1 Introduzione al Linguaggio Java

Java venne introdotto nel 1995, progettato per programmare dispositivi embedded, ma velocemente si utilizzò per scopi diversi da quelli descritti dai suoi creatori. Il team di Java sviluppò un sistema di embedded per i browser, le applet, delle piccole applicazioni che potevano essere eseguite all'interno del browser. Questo fu uno dei primi motivi per il successo di Java, un'altro aspetto importante per la diffusione di Java fu la sua caratteristica di semplificare aspetti di C e C++. Utilizzando un linguaggio più semplice diminuiva il costo associato ad un programmatore, per cui l'adozione di Java venne trainata dalle aziende alla ricerca di abbassare gli stipendi dei suoi dipendenti. In seguito venne utilizzato per creare la macchina virtuale su cui operano tutti i dispositivi Android.

Java è stato ideato con la capacità di mantenere la retro-compatibilità, nonostante le spinte di innovazione tecnologica che il linguaggio ha sostenuto negli anni. Nonostante la sua età Java è uno dei linguaggi di programmazione più usati dalle aziende. La diffusione più importante di Java si ebbe tra il '95 ed il '98.

Java fu il primo progetto industriale che introdusse un nuovo concetto di portabilità su scala industriale, un processore virtuale, poiché non viene progettato per un calcolatore fisico, ma la piattaforma Java Virtual Machine (JVM), per cui uno stesso programma Java può essere eseguito su tutte le piattaforme senza dover ricompilare nuovamente il codice per quella specifica piattaforma. Questa rappresenta una rivoluzione in ambito industriale, il codice oggetto Java "bytecode" prodotto dalla compilazione, come un file di estensione `.class`. Questo bytecode può quindi essere eseguito da qualsiasi piattaforma che dispone di un'implementazione della JVM. Invece programmi creati da un linguaggio più vecchio come C, producono un codice oggetto che presenta istruzioni macchina uniche alla piattaforma in cui è stato compilato. Astruendo il processore fisico si è reso possibile la creazione di programmi che non devono essere ricompilati per ogni piattaforma su cui deve essere eseguito.

Gli applet realizzati tramite Java, venivano eseguiti dalla JVM, e permise la creazione di primi siti web, rivoluzionando il paradigma di programmazione dell'epoca. La sintassi e la semantica del linguaggio sono descritte in un documento noto come Java Language Specification, per risolvere eventuali ambiguità. Inoltre fu uno dei primi linguaggi ad aver introdotto librerie incluse insieme al compilatore, sulla piattaforma Java.

1.1 Rapporto con il Linguaggio C

Il linguaggio C viene descritto come un linguaggio ad un medio livello di programmazione, di alto livello rispetto all'assembly, e permette di realizzare programmi ad alte prestazioni, utilizzato per scrivere sistemi operativi. Java invece venne introdotto in un ambiente diverso rispetto al linguaggio C, per cui si indica come un linguaggio di alto livello di astrazione, non è un linguaggio che cerca le migliori prestazioni, ma cerca di creare programmi semplici e portabili. Nel periodo in cui venne introdotto Java l'andamento delle prestazioni dei calcolatori permise di realizzare linguaggi che non avessero come scopo principale il risparmiare risorse del sistema. Il linguaggio C++ permette di avere prestazioni simili al linguaggio C, ma introducendo paradigmi realizzati in Java. Ma si preferisce per realizzare programmi ad alte prestazioni il C, rispetto al C++, poiché fornisce troppe scelte, complicando l'implementazione e l'uso del linguaggio.

In Java una dichiarazione di una variabile comprende la sua inizializzazione ad un valore nullo. Le dichiarazioni e assegnazioni in Java si ottengono mediante la stessa sintassi del linguaggio C. Si introduce un tipo **boolean** utilizzato per rappresentare i due possibili valori booleani **true** e **false**. Le descrizioni di tutti i tipi utilizzati in Java vengono definite nel JLS. Entrambi sono due linguaggi staticamente tipati, ovvero a tempo di compilazione deve essere noto il tipo di ogni dato utilizzato nel programma. Java introduce il tipo **String** utilizzato per rappresentare stringhe, non è un dato primitivo, e permette l'uso molto più semplice delle stringhe rispetto a C. Rappresenta un tipo particolare di classe, in appoggio al compilatore per favorire la loro gestione da parte del compilatore. Le stringhe vengono inizializzate a letterali stringhe, contenuti tra due doppi apici " ". Da Java 13 è possibile scrivere stringhe letterali multi-linea, utilizzando tripli doppi apici """.

La sintassi di controlli di flusso è esattamente la stessa del linguaggio C, ma le condizioni sono diverse a causa del tipo booleano.

Una grande differenza tra i due linguaggi consiste nella diversa diagnostica a tempo di compilazione e di esecuzione degli errori, poiché in Java i gli errori ed i messaggi di errore forniscono informazioni molto utili per la sua risoluzione. Gli errori a tempo di compilazione sono altrettanto efficaci in Java quanto in C, ma la differenza principale consiste negli errori a tempo di esecuzione tra i due linguaggi, poiché in C, in molte piattaforme, viene fornito un errore estremamente generico, in base alla versione di C, del compilatore, e da altre condizioni.

1.2 Paradigma Orientato agli Oggetti

Il paradigma procedurale, usato nel linguaggio C, separa in maniera netta tra il codice e le operazioni, ovvero la memoria e ed il processore, seguendo l'architettura definita da von Neumann. Anche se rappresenta il modo più naturale per scrivere programmi, nell'industria è stato ampiamente superato a favore del paradigma orientato agli oggetti. Questo paradigma segue la filosofia secondo cui le operazioni e lo stato sono connesse tra di loro, accomunate utilizzando "oggetti". Si è quindi rimossa la divisione imposta sui vecchi linguaggi di programmazione, che seguono l'architettura reale di un calcolatore. Per cui un problema è più facilmente modellabile se viene diviso in una pluralità di oggetti e classi di oggetti, contenenti lo stato e le operazione eseguibili su quell'oggetto. Questi oggetti si scambiano informazioni tra di loro, invocando particolari comandi; in questo modo gli oggetti conoscono ed interagiscono con altri esemplari, anche dello stesso tipo, o classe. Per conoscersi, questi oggetti contengono riferimenti agli altri oggetti.

Questo modo di interpretare un problema è vicino al nostro modo di pensare, il che aiuta nella creazione di programmi, utilizzando linguaggi di programmazione che seguono questo paradigma. Utilizzando un linguaggio come C, progettato per il paradigma procedurale, è possibile programmare seguendo il paradigma orientato agli oggetti, ma non essendo stato creato con questo scopo, la realizzazione di programmi è notevolmente più complesso.

L'esecuzione di un programma è uno scambio di messaggi tra questi oggetti connessi tra di loro tramite riferimenti interni, creando una rete di oggetti, modificando lo stato degli oggetti coinvolti nelle operazioni. Si può rappresentare la rete di oggetti tramite un diagramma di oggetti, indicati come rettangoli contenenti diversi campi, ed eventuali riferimenti ad altri oggetti, rappresentati come archi orientati verso altri oggetti.

Ogni oggetto contiene lo stato, un comportamento, le operazioni che offre, ed un'identità, per differenziare diversi esemplari della stessa classe di oggetti. Si è dimostrato nel tempo come questo paradigma permetta di scrivere e mantenere il codice molto più efficiente rispetto ad altri paradigmi. Gli oggetti vengono definiti tramite classi, rappresentate come file diversi del programma, contenenti lo stato, e la definizione dei metodi. Ogni oggetto creato da una classe possiede un'identità unica, e presenta lo stesso stato e le operazioni della classe di appartenenza. Le classi sono quindi delle istruzioni di montaggio, o fabbriche che producono oggetti.

Per studiare questo paradigma, si utilizzeranno esempi di forme geometriche, come studio di caso più semplice e limitato, ed uno studio di caso più esteso, mirato al refactoring, chiamato “diadia”.

In Java ogni file contenente classi deve essere salvato con lo stesso nome della classe contenente, con estensione “.java”. Le operazioni di oggetti in Java si indicano come metodi, invece di funzioni come nel linguaggio C, poiché rappresentano operazioni nel contesto di una classe, mentre le funzioni possono operare al di fuori di una classe.

1.3 Oggetti e Riferimenti

Si considera un esempio semplice di una classe chiamata **Punto**, che rappresenta un punto in un piano di coordinate cartesiane bidimensionali:

```
public class Punto{
    // stato dell'oggetto
    private int x;
    private int y;

    // metodi propri della classe
    public void setX(int posX){
        this.x = posX;
    }
    public void setY(int posY){
        this.y = posY;
    }
    public int getX(){
        return x;
    }
    public int getY(){
        return y;
    }
}
```

Le variabili della classe **Punto** **x** e **y**, vengono chiamate variabili d'istanza. La creazione di un oggetto di una classe si ottiene tramite l'operatore **new**, secondo la seguente sintassi:

```
// creazione ed assegnazione di un nuovo oggetto Punto tramite riferimento
Punto origine = new Punto();
```

Questa riga di codice invoca un costruttore tramite l'operatore **new**, restituisce un riferimento ad un oggetto del tipo **Punto** appena creato, e viene conservato con un'assegnazione ad una variabile locale, in questo caso chiamato **origine**.

Tramite questo riferimenti è possibile accedere ai metodi, ed eventualmente allo stato dell'oggetto, tramite la notazione puntata, che segue la sintassi **<riferimento-oggetto>.<metodo>(<parametri-attuali>;**

```
origine.setX(0); // assegna lo stato alla variabile x
origine.setY(0); // assegna lo stato alla variabile y
```

Questa notazione è molto efficace nell'esprimere la vicinanza tra i dati e le operazioni, per cui viene ampiamente usata in tutti i linguaggi OO (Orientati agli Oggetti). Ogni nuovo oggetto creato viene salvato in memoria, ed è possibile creare nuovi oggetti fino all'esaurimento della memoria virtuale. Questa classe viene quindi contenuta in un file **Punto.java**.

Una variabile contenente un riferimento ad un oggetto contiene una sequenza di caratteri che rappresenta l'indirizzo in memoria virtuale dell'oggetto, anche se rappresenta una semplificazione, è utile per analizzare il suo comportamento. Mentre per accedere ai valori contenuti, ed assegnare valori alle variabili interne all'oggetto, bisogna utilizzare metodi "getter" e "setter", descritti nella classe dell'oggetto.

Utilizzando due riferimenti allo stesso oggetto per effettuare azioni sullo stesso oggetto si chiama effetto collaterale.

Quando viene chiamato un metodo passando una variabile, viene passato per valore.

In Java il riferimento nullo è un unicamente il letterale di tipo riferimento all'oggetto **null**. Ogni riferimento ad oggetto dichiarato, senza essere inizializzato viene assegnato al riferimento nullo. Indica l'assenza di un riferimento reale ad un oggetto esistente, al contrario del riferimento nullo in C, non corrisponde all'intero 0. Invocando un metodo su un riferimento ad un oggetto nullo si verifica un errore, a tempo di esecuzione, si indica in Java come eccezione: "runtime-exception"; in questo caso: **NullPointerException**.

1.4 Gestione della Memoria

In Java la creazione di oggetti, e quindi la loro allocazione in memoria, avviene tramite l'operatore **new**. Questo operatore richiede la specifica della classe di cui si vuole creare una nuova istanza, e richiede uno dei costruttori della classe. Un costruttore è un metodo che presenta esattamente lo stesso nome della classe e non restituisce niente, e generalmente inizializza tutte le variabili di istanza della classe.

```
public class Punto{
    private int x;
    private int y;
    // costruttore "no-args" della classe "Punto"
    public Punto(){
        this.x = 0;
        this.y = 0;
    }
}
```

Questo tipo di costruttori non prendono parametri, ma è possibile definire costruttori che accettano degli input. Ogni classe presenta sempre almeno un costruttore, e costruisce lo stato iniziale dell'oggetto, per cui è possibile inizializzare l'oggetto ad un certo stato fornito alla chiamata del costruttore:

```
public class Punto{
    private int x;
    private int y;
    // costruttore della classe "Punto"
    public Punto(int x, int y){
        this.x = x;
        this.y = y;
    }
}
```

Al termine dell'esecuzione del costruttore, l'oggetto non è stato completamente inizializzato

Poiché una classe può contenere più di un costruttore, questi vengono distinti in base agli argomenti che prendono. Se invece non viene dichiarato alcuno costruttore all'interno di una classe, è possibile creare un nuovo oggetto usando la stessa terminologia, e tutti i campi numerici vengono inizializzati a zero, i riferimenti a null, e le stringhe alla stringa vuota. Per cui il compilatore crea automaticamente un compilatore "no-args" con le inizializzazioni di default, ma se viene dichiarato un costruttore con parametri all'interno di una classe, e si prova a creare un oggetto della classe usando un costruttore "no-args" il compilatore produce un errore. Quindi il compilatore genera un costruttore implicito solamente se non è dichiarato alcun costruttore dentro una classe.

Poiché gli oggetti si passano come riferimenti, l'operatore == confronta solamente i riferimenti all'oggetto, e bisogna utilizzare un altro metodo per poter confrontare i valori interni degli oggetti tra di loro. Per determinare l'equivalenza tra due oggetti distinti si usa il metodo `equals()`, viene offerto sempre da tutte le classi, se invece non viene espresso esplicitamente ha la stessa semantica dell'operatore == sui riferimenti. Poiché è un metodo interno ad una classe, in generale non è un metodo simmetrico: `a.equals(b) ≠ b.equals(a)`.

La deallocazione della memoria avviene automaticamente da un metodo chiamato Garbage Collector, per cui non è necessario utilizzare esplicitamente una funzione `free()` come in C. Se questo Garbage Collector si accorge che degli oggetti rimangono appesi, li libera dalla memoria.

1.5 Costruttori, Stringhe e Array

L'overloading è una caratteristica di un linguaggio di programmazione, come Java, che permette di definire più procedure con lo stesso nome, ospitati all'interno della stessa classe. Questi metodi, diversi, si distinguono solamente per la lista dei parametri formali, per il numero, per il tipo di uno o più, oppure per il loro ordine.

```
public class Sommatore{
    public int add(int a, int b){
        return a + b;
    }
}
```

```
}  
public int add(int a, int b, int c){  
    return a + b + c;  
}  
public double add(int a, double b){  
    return a + b;  
}  
public double add(double a, int b){  
    return a + b;  
}  
}
```

Nel linguaggio C, questo non è possibile. Se c'è una corrispondenza perfetta tra i parametri passati al metodo, non si riscontrano problemi, invece se non è presente una corrispondenza perfetta tra i tipi passati al metodo, viene eseguito un algoritmo di risoluzione, prima di fermare la compilazione che cerca di capire se è possibile effettuare la chiamata tramite semplici conversioni di tipo, oppure conversioni di tipo più "conservative". La promozione di tipo può essere implicita alla chiamata di un metodo, per cui potrebbe essere complesso determinare quale metodo è stato chiamato.

Alla chiamata di un metodo sovraccaricato è il compilatore a scegliere quale dei metodi da applicare, la scelta è definitiva e viene scritta nell'eseguibile `.class`. Versioni sovraccariche di un metodo potrebbero differire anche per il metodo di ritorno, ma il tipo di ritorno non può essere usato per distinguere due metodi.

L'overloading è presente anche sull'operatore `+`, applicandolo su diversi dati, tramite somme algebriche di interi, numerali a virgola mobile; inoltre viene usato per concatenare le stringhe. In C++ ed in Scala è possibile sovraccaricare gli altri operatori, mentre non è possibile in Java, dove l'unico operatore sovraccaricato è `+`, nella maniera descritta.

Costruttori definiti ripetendo codice, vengono sconsigliati, anche per i metodi sovraccaricati, per cui si tende ad eleggere un costruttore al ruolo di costruttore "primario", il più generico possibile. In seguito gli altri costruttori vengono definiti in base a questo costruttore, invocando questo costruttore, con il metodo `this()`, specificando i parametri che si vuole modificare rispetto al costruttore primario.

```
public class Rettangolo{  
    private int altezza;  
    private int base;  
    private Punto vertice;  
  
    public Rettangolo(Punto v, int b, int h){  
        this.vertice = v;  
        this.base = b;  
        this.altezza = h;  
    }  
}
```



```
public Rettangolo(int b, int h){
    this(new Punto(0,0), b, h);
}

public Rettangolo(){
    this(new Punto(0,0), 0, 0);
}
}
```

In Java esiste la classe **String** per poter rappresentare sequenze di caratteri immutabili, rappresenta un riferimento ad un oggetto istanza della classe **String**. Per rendere il linguaggio più semplice questa classe possiede dei letterali appositi, ed è l'unico oggetto che può essere creato senza una **new** esplicita. L'operatore **==** su due stringhe verifica l'uguaglianza dei riferimenti, per cui per valutare l'equivalenza di contenuto tra due stringhe bisogna utilizzare il metodo **equals()**.

Poiché rappresenta una sequenza di caratteri immutabili, per modificare o aggiornare una stringa è necessario creare una nuova stringa, altrimenti è possibile concatenare due stringhe utilizzando l'operatore **+**, che rappresenta un nuovo oggetto stringa, sovrascritto sulla variabile utilizzata per memorizzarla. Per ottenere la lunghezza di una stringa si utilizza il metodo **.length()**, per ottenere il carattere ad una determinata posizione si utilizza il carattere **.charAt(int i)**. Il metodo **.indexOf(char a)** fornisce la posizione di un carattere o il primo carattere di una sottostringa, per cui è un metodo sovraccarico, altrimenti restituisce **-1**. Per aggiornare una stringa si utilizza il metodo **.replace(Stringa s1, Stringa s2)**, che restituisce un nuovo oggetto dove è stata la sottostringa **s1** con la sottostringa **s2**.

Se viene eseguito il metodo **println()** su un riferimento stampa il valore del riferimento, non il riferimento stesso. Definendo il metodo **toString()** è possibile specificare cosa si vuole visualizzato quando viene chiamato il metodo di stampa a schermo. Un oggetto che possiede questo metodo può essere concatenato con una stringa, questo concatena la stringa con la string restituita dal metodo **toString()**. Se non viene definito questo metodo, il suo comportamento standard implicito stampa l'indirizzo di memoria dell'oggetto, come per il metodo **equals()**.

Un array è una struttura dati che memorizza dati omogenei, per dichiarare un array si inserisce accanto al tipo oppure al nome della variabile **[]**, si preferisce la prima per leggibilità. Sono oggetti per cui devono essere dichiarati con il metodo **new**. Le operazioni di accesso e modifica di un array sono le medesime del linguaggio C, per ottenere la lunghezza di un array si utilizza il metodo **.length()**. Da Java 5 è stato introdotto un modo per poter iterare su un array, senza gestire esplicitamente l'indice di iterazione **for(<tipo> <nome-variabile> : <nome-array>)**, in questo modo la variabile **<nome-variabile>** itera su ogni elemento dell'array **<nome-array>** dello stesso tipo **<tipo>**.

2 Qualità del Codice

2.1 Librerie

Come ogni linguaggio di programmazione, relativamente, recente, Java contiene numerose librerie specializzate per risolvere diversi problemi. Per cui è necessario utilizzare in maniera efficace ed efficiente queste librerie, composte da classi. Bisogna quindi conoscere per nome le classi più importanti di queste librerie, e sapere come cercare ed usare altre classi. Non è necessario conoscere l'implementazione di una classe, ma solo l'interfaccia o API fornito. Queste informazioni vengono fornite tramite la documentazione associata ad ogni libreria, in formato HTML `javadoc`, accessibili tramite un browser, oppure integrate all'interno dell'IDE. Per ogni classe viene fornito il suo nome, una descrizione generale della classe, dello scopo dei suoi costruttori e metodi, e valori di ritorno per i suoi costruttori e metodi. Non sono presenti invece tutti i campi e metodi privati, poiché non sono accessibili dall'esterno, né il suo codice, non necessario per il suo utilizzo.

In caso si voglia creare una nuova classe, bisogna documentarla allo stesso modo delle classi fornite dalle librerie, per permettere il loro utilizzo, senza conoscere in dettaglio la sua implementazione. La documentazione di una classe dovrebbe includere il nome, un commento che descrive la caratteristiche generali e lo scopo della classe, la versione, il nome dei metodi e dei costruttori, una descrizione del suo scopo, i parametri di ritorno e passati al costruttore o metodo. La documentazione può essere generata automaticamente tramite un'utilità chiamata `javadoc`, contenuta in Eclipse, i marcatori utilizzati per definire i comandi si trovano esclusivamente all'interno di un commento generato del genere `/** <comandi> */`:

```
/**
 * Nome-classe: breve descrizione della classe ed il suo scopo
 *
 * @author      autore della classe
 * @see         riferimento ad altre classi
 * @version     versione corrente della classe
 */
```

I commenti per un metodo o un costruttore vengono inseriti prima subito prima del metodo o costruttore stesso.

```
/**
 * Commento che descrive scopo e caratteristiche del metodo o costruttore
 *
 * @param nome-parametro  breve descrizione
 * @return                valore di ritorno
 */
```

La classi vengono raggruppate in **package**, per mantenere insieme classi concettualmente e logicamente correlate. Permette di creare spazi di nomi per evitare conflitti, in oltre permette di definire domini di protezione. Una classe può accedere a tutte le classi presenti nello stesso **package**, altrimenti per accedere a classi pubbliche di un altro pacchetto si può utilizzare il nome completo,

antepoendo il nome del pacchetto, oppure importando la classe ed utilizzando direttamente il nome della classe. Le classi devono dichiarare la propria appartenenza ad un pacchetto tramite la dichiarazione `package <nome-pacchetto>`, all'inizio del file. Una classe può appartenere a più di un pacchetto. Ogni pacchetto deve avere un nome univoco, di solito il nome comprende il nome del dominio Internet dell'organizzazione in ordine inverso. Una classe può essere usata al di fuori del pacchetto solamente se viene dichiarata pubblica, tramite un modificatore d'accesso `public`, se viene omesso questo modificatore, è accessibile solamente alle classi interne allo stesso pacchetto.

Il nome di un pacchetto possiede una struttura gerarchica, questa struttura deve avere una corrispondenza diretta nel file system.

2.2 Coesione e Accoppiamento

Il software evolve continuamente, viene modificato in continuazione, per estensioni, correzioni, mantenimento, etc., e se il costo di questa evoluzione è troppo elevato, il software viene gettato e viene implementato da zero. Questa evoluzione viene effettuata in tempi diversi da molte persone. Se il codice è di cattiva qualità, la sua manutenzione ha un costo relativamente alto. La qualità del codice dipende da due fattori importanti, la coesione e l'accoppiamento. L'accoppiamento tra due o più unità di un programma rappresenta l'impossibilità di modificare una sola unità senza l'eventuale modifica di altre unità accoppiate ad essa. Per cui si vuole evitare l'accoppiamento, per limitare il più possibile la modifica di ulteriori unità, poiché aumenterebbe esponenzialmente i costi di una singola modifica. Un sintomo del forte accoppiamento è la duplicazione del codice all'interno di un progetto. Un basso accoppiamento permette di capire il codice di una classe senza leggere i dettagli di tutte le altre a lei accoppiata, e quindi anche la modifica, senza modificare ulteriori unità.

La coesione si riferisce al numero e all'eterogeneità dei compiti di cui una singola classe è responsabile. Se ogni unità è responsabile di un singolo compito, allora possiede un'elevata coesione. La coesione si applica alle classi, ai metodi ed anche ai pacchetti. Un'alta coesione permette di definire moduli aventi uno scopo preciso ed un compito ben definito. Questo favorisce la comprensione dei compiti di una classe, l'utilizzo di nomi appropriati, ed il riuso di classi e di metodi, rendendo la manutenzione, meno costosa. Alta coesione e basso accoppiamento rappresentano due facce della stessa medaglia.

Per cui quando viene progettato il codice è utile pensare a quali cambiamenti futuri potranno essere richiesti, e come verrà usata la classe.

La manutenzione del codice spesso richiede l'aggiunta di nuovo codice, questo eventualmente degrada il codice nella sua totalità, aumentando l'accoppiamento e diminuendo la coesione, portando ad un'eventuale riorganizzazione del codice. Questa operazione viene chiamata "refactoring", senza modificare la funzionalità del codice. Per mantenere la correttezza del codice viene associato a dei test che stabiliscono i criteri di correttezza del codice, usati per valutare la correttezza di implementazioni interne diverse.

In generale un metodo è troppo lungo se è responsabile di più di un compito logico, mentre una classe è troppo lunga se comprende più di un concetto.

2.3 Tecniche di Testing

In generale i primi errori in un programma sono errori di sintassi, indicati dal compilatore in modo preciso e completo, anche se le specifiche di un messaggio di errore dipendono dal compilatore utilizzato, successivamente possono presentarsi errori logici o “bug”, su cui il compilatore non è in grado di fornire indicazioni. Alcuni errori logici non si presentano immediatamente, per la complessità del software, altri possono essere individuati a tempo di esecuzione, e la macchina virtuale fornisce informazioni precise sulla natura dell'errore, la maggior parte delle volte si tratta di un `NullPointerException`.

Un programma rappresenta una singola descrizione statica associata a molteplici possibili esecuzioni dinamiche. Il compilatore è in grado di individuare errori a livello statico, a tempo di scrittura o compilazione, e non è in grado di individuare possibili errori dovuti all'esecuzione e l'evoluzione del programma. I bug sono quindi errori che si presentano durante l'evoluzione dinamica del programma, che il compilatore non è in grado di prevedere. Il costo delle operazioni di debugging, rappresenta il costo principale di ogni moderno progetto di software, ed è interamente a carico del programmatore.

La correzione di un bug dipende da due grandezze, la dimensione del contesto, ovvero il numero di linee di codice in cui il bug si può annidare. Inoltre dipende dal tempo necessario al bug per manifestarsi, rappresenta una misura temporale tra la causa del bug ed il rilevamento dei suoi effetti.

Per rilevare i malfunzionamenti di un software vengono utilizzati delle tipologie di test, in tre fasi sequenziali:

- Mettere il sistema in uno stato iniziale noto;
- Iniziare a sollecitare il sistema;
- Controllare lo stato del sistema e confrontarlo con lo stato atteso.

Se i test sono progettati e mantenuti, allora permettono di individuare tempestivamente i bug all'interno del software, restringendo le cause del singolo bug. Se il test ha successo si ha una garanzia sul comportamento dinamico del codice, altrimenti fornisce informazioni sulla presenza di un bug. Ad ogni modifica è necessario ripetere un'operazione di testing, per identificare eventuali bug, inseriti durante la manutenzione del codice, in questo modo è possibile ricercare localmente l'errore, in maniera molto economica. Inoltre utilizzando questo processo è possibile prevenire la regressione, localizzando l'errore nel codice appena aggiunto al programma.

Per automatizzare questo processo si realizza, accanto al codice di produzione, il codice di test. Una possibile soluzione inizializza tutti i casi da testare sul codice di produzione, inserendo, ad ogni chiamata, il valore atteso confrontandolo con il valore ottenuto. I test devono essere automatici, efficienti ed isolati, per mantenere la località degli errori, inoltre devono essere separati dal codice applicativo ed eseguibili e verificabili separatamente. Esistono vari strumenti per assistere il programmatore nel testing, in particolare l'unit-testing, in Java il più noto ed utilizzato framework è “JUnit”, integrato all'interno di Eclipse. Per ogni classe creata si crea una classe parallela dedicata al testing, contenente una batteria di test, ognuno un metodo differente per testare diversi casi, preceduto dall'annotazione `@Test`, contenente un'asserzione sul risultato aspettato tramite il metodo `assertEquals()`.

Tutte le classi di test hanno la stessa struttura, si inseriscono nello stesso pacchetto della classe che devono testare, e vengono chiamate rispettando la convenzione `test<nome-classe>`, anche se non è più necessario utilizzare questa notazione, semplicemente tramite l'annotazione `@Test`, anche se viene favorito il suo utilizzo. I risultati attestati sono documentati tramite asserzioni esplicite, non mediante stampe, se l'asserzione è vera il test è andato a buon fine, altrimenti è presente un errore nella sezione di codice testato. Sono possibili altre asserzioni fornite da JUnit, sovraccariche e facilmente intercambiabili, per cui è sempre favorito utilizzare la versione più pertinente.

Una variante di `assertEquals()`, prevede una stringa, da stampare per fornire informazioni sul risultato del test.

Per facilitare la scrittura di tutti i metodi di test, è comodo utilizzare degli oggetti predisposti per l'utilizzo da parte di tutti i test-case operazione chiamata Fixture. JUnit permette di confinare in un unico metodo `setUp()` la creazione di tutti gli stati noti a priori tramite l'annotazione `@BeforeEach`. In questo modo i test diventano estremamente semplici. La lunghezza ottimale di un singolo test-case consiste di una singola riga di codice, rappresentata dall'asserzione. Per favorire la semplicità dei test si può fattorizzare il codice di creazione della fixture. L'uso di test minimali favorisce la leggibilità e la ricerca di errori, ed è sempre conveniente partire da test minimali per individuare i bug minori, per poi aumentare la complessità dei test ed individuare bug sempre più complessi.

In generale è buona norma scrivere i test prima ancora di aver scritto il programma, una tecnica chiamata “test-driven-development”.