

AgileLab

Marco Savino

1 Introduzione

Step per entrare in AgileLab:

1. Colloquio conoscitivo HR
2. Test di conoscenza tecnica
3. Intervista tecnica

2 Preparazione al test

Il test è diviso in tre macro aree:

- Concetti di software engineering
- Database
- Sistemi distribuiti (engine)
- Machine learning (in minor parte)

3 Software Engineering

3.1 Programmazione orientata agli oggetti

La programmazione orientata agli oggetti è un paradigma di programmazione basato sul concetto di oggetto, ovvero un'entità che può contenere dati e operazioni. I dati sono contenuti in campi conosciuti come attributi e le operazioni sono codice scritto in forma di procedure chiamate metodi.

Classi Per poter applicare il concetto di oggetto, è necessario prima definirlo: un oggetto viene definito attraverso una classe. La classe serve a stabilire il formato dei dati e le operazioni disponibili dell'oggetto creato tramite essa. Gli oggetti sono quindi delle istanze, ovvero la manifestazione fisica dei dati e delle operazioni descritte all'interno nella classe.

Concetti della Programmazione Orientata agli Oggetti (A. P.I.E.) I concetti alla base di questo paradigma hanno lo scopo di creare metodi e variabili funzionanti, che permettano poi il riutilizzo del codice senza comprometterne la sicurezza.

A - Abstraction (Astrazione) L'Astrazione è un concetto per il quale si nascondono all'utente i dettagli superflui dell'implementazione, permettendo a questo di concentrarsi sullo sviluppo della logica al di sopra dell'entità astratta fornita, senza doversi preoccupare di come è fatta al suo interno.

Esempio - Una macchina del caffè rappresenta un'astrazione per l'utente: questo non sa come viene prodotto fisicamente il caffè, ma conosce quanto basta per comandare alla macchinetta di produrne uno. In questo modo, è persino possibile cambiare il meccanismo interno della macchina senza che l'utente ne debba essere a conoscenza.

P - Polymorphism (Polimorfismo) Il polimorfismo è un fenomeno in cui una stessa entità assume proprietà differenti a seconda del contesto. In informatica, questo significa poter accedere, in base alla situazione, a diverse classi tramite la stessa interfaccia o a diversi metodi tramite lo stesso nome. In un linguaggio orientato agli oggetti esistono due tipi di polimorfismo:

- *Statico* - All'interno di una stessa classe è possibile definire metodi con lo stesso nome ma con parametri diversi in numero, tipo o ordine: questa proprietà è definita come *overload* del metodo. Al momento dell'invocazione del metodo, verrà eseguito quello selezionato non solo in base al nome, ma anche attraverso le caratteristiche dei parametri della chiamata: si fa così riferimento alla *firma (signature)* del metodo. Si definisce polimorfismo *statico* perché l'associazione fra chiamata e metodo chiamato viene stabilita durante la compilazione.
- *Dinamico* - Data una gerarchia di classi, per una sottoclasse è possibile implementare un metodo con la stessa firma di uno della sua superclasse. Alla creazione dell'istanza della sottoclasse, quel metodo avrà la prevalenza su quello corrispondente della superclasse: questa proprietà prende il nome di *overriding*. Inoltre, è possibile che l'istanza di una sottoclasse venga assegnata alla reference della sua superclasse: l'oggetto così appare come una istanza della superclasse ed è possibile invocare solo i metodi di questa. Tuttavia, al momento della creazione dell'istanza della sottoclasse a run-time, alla chiamata del metodo della superclasse viene associato il metodo della sottoclasse. Questa associazione a run-time fra chiamata del metodo e metodo chiamato, selezionando prevalentemente quelli delle sottoclassi, è il fenomeno definito come *late binding* ed è alla base dell'applicazione di costrutti come le *interfacce*.

I - Inheritance (Ereditarietà) L'ereditarietà è il meccanismo attraverso cui la definizione di una classe deriva da quella di un'altra classe. La prima

prende così il nome di sottoclasse e la seconda di super-classe. La sottoclasse condivide lo stesso formato dei dati e le stesse operazioni della sua super-classe, in base alla loro visibilità, e in più ha la possibilità di definirne dei propri a sua volta. Una classe può ereditare una sola altra classe e un numero illimitato di interfacce.

E - Encapsulation (Incapsulamento) Per definizione, l'incapsulamento consiste nel combinare metodi e attributi in un'unica unità, come ad esempio una classe. Grazie a questo concetto, è possibile nascondere la rappresentazione interna dei dati o il loro stato, rendendoli accessibili solo tramite metodi appositi. Questa tecnica è chiamata *Information Hiding*. Nella pratica, un linguaggio orientato agli oggetti può prevedere principalmente quattro modificatori di visibilità per metodi e attributi:

- *Private* - L'elemento è accessibile solo dalla propria classe e nient'altro.
- *Package-private* - Se non si specifica il modificatore di visibilità, l'elemento è accessibile dalla propria classe e da tutte le classi ed eventuali sottoclassi appartenenti allo stesso pacchetto.
- *Protected* - L'elemento è accessibile dalla propria classe, dalle classi presenti nello stesso pacchetto e anche dalle sue eventuali sottoclassi presenti in altri pacchetti.
- *Public* - L'elemento è accessibile da qualunque classe.

3.2 S.O.L.I.D.

Questo acronimo indica un gruppo di cinque principi da seguire, affinché il codice sia più leggibile e più facile da cambiare in caso di necessità.

S - Single Responsibility Principle *Ogni classe è responsabile di un singolo compito; se cambia il compito, allora cambia anche la classe. Questa deve essere l'unica ragione per cui la classe cambia.*

Esempio: Una classe che si occupa di tenere conto del punteggio di un gioco, non si deve occupare di altro. Questa cambia solo se cambiano le regole del gioco

O - Open-Closed Principle *Una classe deve essere aperta all'utilizzo da parte di altre classi, come sua estensione, e chiusa alla possibilità di esser cambiata da altre classi.*

Questo principio si applica attraverso l'Ereditarietà e l'override: creando una sottoclasse ad hoc e sovrascrivendo il metodo di interesse in partenza, si evita di alterare cambiare la classe madre, proprio grazie all'estensione della stessa.

L - Liskov Substitution Principle *Una proprietà dell'oggetto X di una classe T deve essere valida anche per un oggetto Y di classe S , dove S è sottoclasse di T .*

Questo implica che, in un programma, le istanze di una classe dovrebbero essere sempre sostituibili dalle istanze di una sua sottoclasse. A livello pratico, ogni suo metodo deve poter accettare gli argomenti del corrispettivo metodo della classe madre e ritornare lo stesso tipo di valore. In alcuni casi però, questo ha dei limiti: una classe Rettangolo crea un'istanza grazie agli argomenti base e altezza, ma una sua sottoclasse Quadrato utilizzerebbe come argomento solo il lato; in questo caso, per rispettare il principio la classe Quadrato non deve essere sottoclasse di Rettangolo.

I - Interface Segregation Principle *"Un cliente non dovrebbe essere costretto a dipendere da una interfaccia che non usa" - Robert C. Martin*

Non bisogna aggiungere funzionalità che non vengono utilizzate a una interfaccia. Questo perché, aggiornando l'interfaccia, bisogna aggiornare anche tutto ciò che ne dipende. L'ideale è quindi creare un'interfaccia a parte che non abbia influenze su quella di partenza.

D - Dependency Inversion Principle *Moduli di altro livello non devono dipendere da moduli di basso livello, ed entrambi devono basarsi su astrazioni; le astrazioni non devono dipendere dai dettagli, ma i dettagli dalle astrazioni.*

Si dà importanza al disaccoppiamento fra moduli: l'implementazione di questi potrà cambiare senza influire sulla loro comunicazione. A livello pratico, si applica la cosiddetta *Dependency Injection*.

3.3 Sviluppo software

Ciclo di vita del software il *Software Development Life Cycle* viene definito con 6 fasi:

1. *Raccolta dei requisiti*: "Qual è il problema attuale?"
2. *Pianificazione*: "Che cosa vogliamo?"
3. *Progettazione*: "Come otteniamo quello che vogliamo?"
4. *Implementazione*: "Creiamo quello che vogliamo"
5. *Test e integrazione*: "Abbiamo ottenuto quello che volevamo? Funziona?"
6. *Manutenzione*: "Miglioriamo quello che abbiamo ottenuto"

Queste fasi possono essere gestite in numerosi modi diversi, definiti come modelli:

- *Modello a cascata*: le fasi vengono svolte linearmente una dopo l'altra.

- *Modello a V*: estensione del modello a cascata, testando però dopo ogni fase.
- *Modello iterativo*: si crea una versione base che viene migliorata tramite frequenti aggiornamenti uno dopo l'altro.
- *Modello agile*: il prodotto è gestito a cicli e con frequenti aggiornamenti, per migliorare in base al feedback dell'utente.
- *Modello a spirale*: Estensione del modello iterativo, vengono ripetute tutte le fasi del SDLC
- *Big Bang Model*: Viene lasciato tutto in mano a sviluppatori esperti, che si concentrano da subito sull'implementazione

Testing Ci sono diversi tipi di test, che servono a garantire il corretto funzionamento del codice in tutte le occasioni:

- *Analisi statica*: esamina tutti i possibili comportamenti del codice che potrebbero manifestarsi a run-time. Include sempre la revisione del codice, l'ispezione del codice, l'analisi algoritmica e la prova di correttezza.
- *Analisi dinamica*: coinvolge l'esecuzione del codice per esporre errori e malfunzionamenti.
- *Black box*: si testa il codice attraverso input e output, ignorando il modo in cui è implementato il codice.
- *White box*: presupponendo la conoscenza dell'implementazione, si effettuano i test concentrando sulla struttura del codice e sulla logica di business (vedi *statement coverage*, *branch coverage*, *path coverage*).
- *Scripted box*: specifico test basato su una serie di step prefissati pensato per validare un preciso requisito.
- *Exploratory*: il tester simula il comportamento dell'utente finale e si basa sull'intuito per rilevare problemi nascosti o non previsti in precedenza.
- *Manual*: viene eseguito da persone.
- *Automated*: viene eseguito da software automatizzati.

3.4 Design Pattern

3.4.1 Pattern Creazionali

Factory

Abstract Factory

Builder

Singleton

Prototype

3.4.2 Pattern Strutturali

Adapter

Bridge

Composite

Decorator

Facade

Flyweight

Proxy

3.4.3 Pattern Comportamentali

Chain of Responsibility

Command

Iterator

Mediator

Memento

Observer

State

Strategy

Template Method

3.5 Strutture dati

Binary Tree

3.6 Altri concetti

Dependency Injection A partire da un'interfaccia, che ha un livello di astrazione più alto di una classe, si applica la dipendenza, ovvero la classe che la implementa. La dipendenza può cambiare senza che l'interfaccia ne risenta, persino a run-time. In questo caso, si parla di **Binding Dinamico**.

Inversion of Control Principio di programmazione secondo cui il controllo del flusso dello sviluppo di una applicazione viene lasciato alla piattaforma che si utilizza, la quale prende il nome di **framework**. I vantaggi principali consistono nel riutilizzo del codice e nel permettere all'utente di concentrarsi sull'implementazione dei moduli, piuttosto che sulla loro comunicazione. La Dependency Injection è un modo di implementare il principio di Inversion of Control.

4 Database

4.1 A.C.I.D.

Proprietà che devono avere delle transazioni in un database.

A - Atomic Una transazione deve essere un'unità "atomica", ovvero non può essere suddivisa in più parti e, una volta iniziata, deve essere portata a termine.

C - Consistency La transazione deve rispettare i vincoli di integrità del database.

I - Isolation Una transazione deve essere eseguita in modo isolato e indipendente dalle altre.

D - Durability Definibile anche come *persistenza*, implica che, una volta che la transazione abbia effettuato una *commit*, i cambiamenti apportati devono essere conservati.

4.2 Relazioni

Uno a uno A un record di una tabella, corrisponde uno e un solo record di un'altra tabella. Si realizza ponendo il vincolo UNIQUE sulla foreign key.

Uno a molti A un record di una tabella corrispondono uno o più record di un'altra tabella. Si realizza ponendo il vincolo della foreign key.

Molti a molti A più record di una tabella possono corrispondere più record di un'altra tabella. Per realizzarlo fra due tabelle è necessario introdurre una terza tabella intermedia che leghi la foreign key della prima con la primary key della seconda.

4.3 Normalizzazioni

Forma zero La tabella dispone di chiave primaria.

Prima forma normale La tabella presenta campi atomici, senza tabelle al loro interno. Facendo riferimento al modello E-R, significa che eventuali attributi composti sono stati trasformati in entità, cioè tabelle, a loro volta.

Seconda forma normale In caso la chiave primaria sia composta da più campi, ogni altro campo del record deve essere *funzionalmente dipendente* da essa nella sua interezza, e non solo da alcune sue parti.

Terza forma normale Ogni campo di un record deve essere funzionalmente dipendente solo e soltanto dalla chiave primaria, in modo diretto.

Forma normale di Boyce e Codd Formulazione leggermente più forte della terza forma normale: ogni campo deve essere funzionalmente dipendente da una super-chiave.

5 Constraints

I vincoli sono delle regole che vengono assegnate per garantire l'accuratezza e l'affidabilità di una tabella. Possono essere *column level* oppure *table level*. I vincoli più comuni in SQL sono:

- **NOT NULL**: il campo della colonna non può avere valore NULL.
- **UNIQUE**: nella colonna non può esserci lo stesso valore più volte.
- **PRIMARY KEY**: unione dei vincoli NOT NULL e UNIQUE, ve ne è solo uno per tabella.
- **FOREIGN KEY**: associa i valori della colonna a quelli della chiave primaria di un'altra tabella; impedisce la cancellazione della tabella grazie a questo collegamento, se non specificato diversamente. Conosciuto anche come *vincolo di integrità referenziale*.
- **CHECK**: si assicura che i valori in una colonna soddisfino specifiche condizioni.
- **DEFAULT**: assegna un valore di default se il campo non viene valorizzato.

- *CREATE INDEX*: puntatore alla cella che permette il rapido accesso ai dati.

5.1 Big Data - Le cinque V

5.2 Altri concetti

Dipendenza funzionale In un record di una tabella, un campo è funzionalmente dipendente da un altro campo quando, in presenza di un determinato valore del primo, si ottiene un determinato valore del secondo.

OLD - Data Abstraction Il Data Abstraction è un design pattern per il quale i dati sono visibili solo alle funzioni con cui sono semanticamente collegati. Di conseguenza, si integra il Data Hiding come principio di design nei linguaggi di programmazione ad oggetti. Ad esempio, in Java si esplicita il grado di visibilità di un attributo o di un metodo attraverso le parole chiave private, protected e public.

OLD - P - Polymorphism (Polimorfismo) Il polimorfismo consiste nel fornire una stessa interfaccia a entità di tipi differenti, in modo da dare l'impressione che uno stesso oggetto sia in grado di avere funzioni differenti. Ne esistono di tre tipi: il polimorfismo ad-hoc, il polimorfismo parametrico e il polimorfismo di sottotipo (o di inclusione). Il primo si riferisce a funzioni che possono essere applicate a parametri di tipi differenti, ma che si comportano diversamente a seconda di questi: questo concetto è identificato con il termine di *overload*. Quello parametrico permette a un tipo di dato o a una funzione di essere scritto in modo generale, in modo che possa trattare i valori indipendentemente dal loro tipo. L'ultimo, il polimorfismo di sottotipo, enuncia che una sottoclasse può utilizzare un metodo della sua super-classe e questo potrebbe avere un comportamento diverso: ciò viene indicato come *override* del metodo e rimane in accordo con il Principio di Sostituzione di Liskov.

NoSQL

6 Sistemi distribuiti

Un sistema distribuito è un insieme di componenti che comunicano e coordinano le loro azioni attraverso lo scambio di messaggi.

6.1 Caratteristiche dei sistemi distribuiti

Vantaggi

- *Scalabilità* - Il sistema è scalabile orizzontalmente attraverso l'impiego di un numero potenzialmente illimitato di macchine economiche.

- *Affidabilità* - Se un componente cessa di funzionare, il resto del sistema può continuare a operare, anche se a regime ridotto.
- *Prestazioni* - Le applicazioni su sistemi distribuiti tendono a suddividere problemi complessi in numerosi parti più semplici, la cui elaborazione viene parallelizzata su più macchine.

Svantaggi

- *Coerenza o disponibilità* - Poiché i sistemi distribuiti offrono per definizione piena *partition tolerance* (*tolleranza delle partizioni*), bisogna scegliere

6.2 Hadoop

7 Fonti

Wikipedia

Stackify

- <https://stackify.com/oops-concepts-in-java/>
- <https://www.baeldung.com/cs/distributed-systems-guide>

EnjoyAlgorithms

Baeldung