



PROGETTAZIONE DI DATABASE

Introduzione alle Basi di Dati

Lezione 01



Questa dispensa è rilasciata sotto la licenza Creative Common CC BY-NC-SA. Chiunque può copiare, distribuire, modificare, creare opere derivate dall'originale, ma non a scopi commerciali, a condizione che venga riconosciuta la paternità dell'opera all'autore e che alla nuova opera vengano attribuite le stesse licenze dell'originale.

Versione del: **26/09/2017**

Revisione numero: **3**

Prof. Andrea Zoccheddu
Dipartimento di Informatica

Immagine di copertina da: <http://www.iamsterdam.com/en-GB/living/education/Dutch-Education-System>





INTRODUZIONE

SISTEMI INFORMATIVI



Fonte (informazione tratta dal seguente sito) :

http://it.wikipedia.org/wiki/Sistema_informativo

Il sistema informativo è costituito dall'insieme delle informazioni utilizzate, prodotte e trasformate da un'azienda durante l'esecuzione dei processi aziendali, dalle modalità in cui esse sono gestite e dalle risorse sia umane sia tecnologiche coinvolte. Tutto questo partendo da dati che descrivono fenomeni di carattere aziendale o ambientale. Esso non va confuso con il sistema informatico, che indica la porzione di sistema informativo che fa uso di tecnologie informatiche e automazione.

Elementi del sistema informativo

Come dice la parola stessa, un sistema informativo è un "insieme" di elementi (nel nostro caso informazioni) unite in un unico agglomerato. Possiamo dare una prima classificazione delle caratteristiche essenziali che connotano questo tipo di sistema:

- 1) **Dati:** sono la componente essenziale del sistema, ma dal momento che non sono ancora stati elaborati, si presentano in uno stato primitivo.
- 2) **Informazioni:** insieme di dati già elaborati, strettamente collegati tra di loro con un fine preciso.
- 3) **Persone:** coloro che si occupano di raccogliere e catalogare i dati di interesse (opportunosamente registrati), affinché possano essere poi elaborati dalle strutture competenti. Sono anche destinatari delle informazioni già manipolate.
- 4) **Strumenti:** è l'insieme delle attrezzature che sono in grado di far viaggiare le informazioni tra fornitore e acquirenti, tra diverse aziende, e in genere tra punti diversi di un'azienda. Possiamo anche inserire in questa categoria tutte quelle infrastrutture in grado di trasformare i dati in informazioni. Ovviamente al giorno d'oggi si tratta di mezzi altamente tecnologici.
- 5) **Procedimenti:** è l'insieme delle procedure che permettono di capire in che maniera vengono raccolti ed elaborati i dati. Per ogni singola finalità, le persone di competenza devono scegliere la modalità per elaborare i dati, dal momento che ogni azienda ha una propria esigenza da soddisfare.

Definizione di Base di Dati (DATABASE)

Un dato è un valore (numerico, alfabetico, logico o di altro tipo) che si può conservare (in memoria) ed elaborare. Una informazione è un dato o un insieme di dati a cui può essere associato un significato.

Una Base di Dati (Data Base o DB) può essere definita come una massa di dati strettamente connessi e relativi ad uno specifico contesto che si intende gestire. I dati possono essere memorizzati su un supporto di memoria di massa e mantenuti attraverso un insieme di regole di gestione che stabiliscono il profilo informativo.

Il DB, pertanto, è costituito da dati e da regole di gestione.

Il DB è un modo per rappresentare la realtà con un modello.



Esempio

Si pensi ad una compagnia di voli aerei. L'insieme dei dati relativi ai mezzi, alle prenotazioni e vendite di biglietti, ai dati dei passeggeri e del personale, ai voli e orari costituisce la base dati.

La realtà che si intende rappresentare e gestire viene interpretata come un insieme di dati, tutti connessi tra loro perché appartenenti allo stesso ambito.



Definizione di DBMS (data base management system):

Un DBMS è un sistema per la gestione di basi dati.

Un DBMS è visto come un software che provvede alla memorizzazione (archiviazione) dei dati, al loro trattamento (inserimento, aggiornamento, modifica e cancellazione, elaborazione di ricerca e di informazione aggiunta) ed alla relativa informazione (verifica della coerenza dei dati rispetto alla realtà rappresentata, della loro consistenza rispetto a dei vincoli di validità, della loro protezione e del recupero di dati corrotti).

Il DBMS deve necessariamente occuparsi della manutenzione dell'insieme di registrazioni che devono essere memorizzate in supporti di memoria di massa (dischi, nastri, unità ottiche o magnetiche, ecc ...).

Uno stesso DB potrebbe essere gestito da diversi DBMS.

Il DBMS, pertanto, è un insieme di programmi, che sono rivolti alla gestione di dati memorizzati in archivi. Ovviamente, tra Data Base e DBMS esiste una forte iterazione per cui spesso si tende a confonderli, ma sono comunque due cose ben diverse e distinte.



Esempio

Se un ragazzo fa collezione di figurine e le raccoglie, conserva, dispone e ricerca secondo certe regole, allora il ragazzo può essere visto come un DBMS mentre il DB gestito sono le figurine.

Progettazione di basi di dati

Fasi della progettazione

Di solito si progetta un DB concernente uno specifico argomento di interesse.

L'oggetto di interesse di solito è commissionato da qualcuno (il cliente, il capo, il docente) o scelto arbitrariamente per motivi personali del progettista. In ogni caso la prima attività da affrontare è quella della Analisi della Realtà.

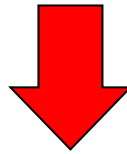
La realtà viene osservata ed indagata e l'analista deve cogliere quali siano gli aspetti importanti e quelli trascurabili. La distinzione tra le due categorie di elementi non è banale e in ogni caso fa riferimento sia al committente sia alla sensibilità dell'analista.

Questo implica che l'analista deve sforzarsi di considerare e rispettare le richieste del committente ma che talvolta il committente stesso non si accorge di elementi rilevanti ma talmente ovvi da averli trascurati nelle richieste.

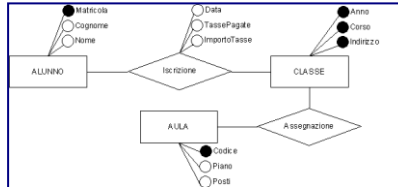
Nel corso di questo anno vedremo delle tecniche di modellazione e realizzazione della base di dati. Il procedimento che seguiremo, abbastanza standard, si svolge attraverso le seguenti fasi:



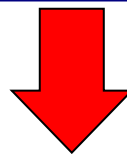
Fase 1
Progettazione Concettuale



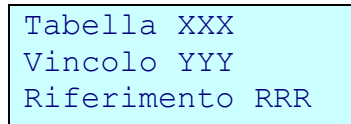
In questa fase si analizza la realtà e la si rappresenta con un modello astratto lontano dalla implementazione sulla macchina. Il risultato è un Progetto Concettuale



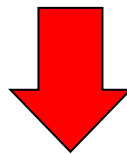
Fase 2
Progettazione Logica



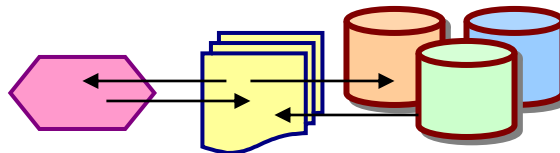
In questa fase si traduce il progetto concettuale in un modello per DB, un po' più vicino all'implementazione sulla macchina. Il risultato è un Progetto Logico.



Fase 3
Implementazione



In questa fase si realizza un DataBase su uno specifico DBMS usando un ambiente e dei linguaggi. Il risultato è un Progetto Fisico.





Progettazione Concettuale

Come abbiamo visto la realtà deve essere osservata ed indagata ed è necessario distinguere quali siano gli aspetti importanti e quali quelli trascurabili.

In alcuni casi questo lavoro non è facile e talvolta ci si accorge successivamente di errori, trascuratezze e imprecisioni.

Tra gli elementi che occorre valutare ricordiamo che si dovrebbe:

- classificare gli elementi ritenuti importanti ed elencarli con appositi schemi
- individuare tra gli elementi se ci sono omonimi (ovvero due elementi diversi ma con stesso nome) e cercare di evitarli (rinominare un elemento)
- individuare tra gli elementi se ci sono sinonimi (ovvero due elementi concettualmente equivalenti ma annotati con nomi diversi) e cercare di evitarli (unire gli elementi)
- individuare tra gli elementi se alcuni ne comprendono altri (ovvero un elemento ne contiene un altro) e valutare le diverse soluzioni possibili
- evitare la ridondanza di informazioni (due informazioni possono essere ricavati con due procedimenti diversi o in due luoghi diversi)

La fase di progettazione concettuale solitamente termina con un Progetto Concettuale costituito da:

- Schema Entity-Relationship (ma anche altri modelli di rappresentazione)
- Dizionario dei Dati (una breve descrizione dei dati in italiano)

Progettazione Logica

La progettazione logica è un modo per trasformare il progetto concettuale, troppo astratto, in un modello più vicino alla rappresentazione di macchina.

La progettazione logica può essere svolta usando diversi modelli, ciascuno dei quali si rivolge ad una specifica realizzazione con un DBMS. Questi modelli sono:

- Modello gerarchico
- Modello reticolare
- Modello relazionale
- Modello ad oggetti

Nel **modello gerarchico** i dati sono organizzati in record connessi tra loro secondo strutture ad albero. Ogni record del database, che non sia la radice dell' albero, deve avere uno e un solo padre. Possono quindi esserci due record, su alberi diversi, che rappresentano la stessa informazione. Questo comporta problemi di ridondanza nel database e richiede controlli di consistenza durante il suo uso. Inoltre non è possibile memorizzare informazioni che non hanno padre.

La struttura gerarchica impone delle regole rigide sul modo di eseguire gli aggiornamenti e le interrogazioni: il livello più alto è il punto di accesso a tutte le informazioni. La cancellazione di un record del database comporta l' eliminazione di tutti i record dipendenti gerarchicamente da esso. L' aggiornamento di un dato richiede l' accesso e la modifica di diversi record per assicurare la consistenza del database.

Il modello gerarchico comporta la dipendenza dei programmi dalle strutture, quindi non possiamo modificare le strutture senza modificare i programmi.

Nel modello reticolare i record sono legati tra loro con strutture ad anello (puntatori) che permettono all' utente di accedere ai dati più facilmente, senza i vincoli rigidi della struttura gerarchica. Ogni nodo può essere il punto di partenza per raggiungere un determinato campo. Un record può avere uno o più record padre e ciò permette di evitare i problemi di ridondanza. Rimangono il problema della dipendenza dei programmi dalle strutture e il



Ciclo vita dei sistemi informativi



Fonte (informazione tratta dal seguente sito) :

http://it.wikipedia.org/wiki/Sistema_informativo

Le fasi di realizzazione di un sistema informativo sono generalmente le seguenti:

- **nascita dell'esigenza:** tende a migliorare la situazione esistente tramite il perseguimento di determinati obiettivi.
- **studio di fattibilità:** serve a definire in maniera per quanto possibile precisa i costi delle varie alternative possibili, ad effettuare la conseguente analisi costi/benefici ed a stabilire le priorità della realizzazione delle varie componenti del sistema.
- **raccolta e analisi dei requisiti:** consiste nell'individuazione e nello studio delle proprietà e delle funzionalità che il sistema informativo dovrà avere. Questa fase richiede un'interazione con gli utenti del sistema e produce una descrizione completa ma generalmente informale dei dati coinvolti e delle operazioni su di essi. Vengono inoltre stabiliti i requisiti software e hardware del sistema informativo.
- **progettazione:** si divide generalmente in progettazione dei dati, progettazione delle applicazioni e progettazione dell'architettura tecnica di sistema. Nella prima si individua la struttura e l'organizzazione che i dati dovranno avere, nella seconda si definiscono le caratteristiche dei progetti applicativi. Queste due attività sono complementari e possono procedere in parallelo o in cascata. Le descrizioni dei dati e delle applicazioni prodotte in questa fase sono formali e fanno riferimento a specifici modelli. La progettazione dell'architettura tecnica di sistema, infine, rappresenterà l'infrastruttura individuandone le caratteristiche in termini di sistemi (server), connettività, sicurezza fisica e logica.
- **sviluppo:** consiste nella realizzazione del sistema informativo secondo la struttura e le caratteristiche definite nella fase di progettazione. Viene costruita e popolata la base di dati e viene prodotto il codice dei programmi.
- **validazione e collaudo:** serve a verificare il corretto funzionamento e la qualità del sistema informativo. La sperimentazione deve prevedere, per quanto possibile, tutte le condizioni operative.
- **avviamento:** è la fase di messa in funzione del sistema. Si erogano i corsi di formazione, si travasano i dati da eventuali realizzazioni preesistenti che vengono sostituite, si attivano i collegamenti con le altre applicazioni e si parte con l'operatività reale.
- **funzionamento:** in questa fase il sistema informativo diventa operativo a regime ed esegue i compiti per i quali era stato originariamente progettato. Se non si verificano malfunzionamenti o revisioni delle funzionalità del sistema, questa attività richiede solo operazioni di gestione e manutenzione.
- **manutenzione:** con la manutenzione *correttiva* si consolida il sistema, mentre con la manutenzione *evolutiva* lo si completa ed arricchisce di funzionalità inizialmente non individuate.

Va precisato che il processo non è quasi mai strettamente sequenziale, in quanto spesso durante l'esecuzione di una attività citata bisogna rivedere decisioni prese nell'attività precedente. Quello che si ottiene è un ciclo di operazioni. Inoltre alle attività citate si aggiunge quella di [prototipizzazione](#), che consiste nell'uso di specifici strumenti software per la realizzazione rapida di una versione semplificata del sistema informativo, con la quale sperimentare le sue funzionalità. La verifica del prototipo può portare a una modifica dei requisiti e una eventuale revisione del progetto.

Il suddetto ciclo di vita ha costituito la base sulla quale è stato poi formalizzato il ciclo di vita delle architetture telematiche.



MODELLI, SCHEMI E ISTANZE

Modello



Fonte:

[http://it.wikipedia.org/wiki/Modello_\(scienza\)](http://it.wikipedia.org/wiki/Modello_(scienza))

Nell'uso scientifico e tecnico-progettuale, **un modello è una rappresentazione** di un oggetto o di un fenomeno, che corrisponde alla cosa modellata per il fatto di riprodurne (evidentemente alla luce di precisi riscontri fattuali ottenuti a partire da un metodo che sia garanzia di controllabilità) alcune caratteristiche o comportamenti fondamentali; in modo tale che questi aspetti possano essere mostrati, studiati, conosciuti laddove l'oggetto modellato non sia direttamente accessibile.

La **realizzazione di un modello** scientifico o tecnico, per quanto possa essere genericamente orientata o guidata in partenza da una teoria metafisica, deve sempre essere il risultato di un contesto della prova rigoroso, predisposto in modo tale da non essere minimamente influenzato dalle aspettative e dall'interpretazione soggettiva dell'osservatore (si dice che l'osservazione e l'esperienza scientifiche, su cui si fonda la formulazione di modelli teorici validi, sono invarianti rispetto all'osservatore).

Nella teoria dei Database, **un modello è un insieme di regole, procedure e tecniche per rappresentare la realtà**, selezionando solo quegli aspetti interessanti dal punto di vista del progettista. Un modello quindi è un modo per costruire una rappresentazione della realtà, necessariamente in modo parziale e sommario, mirato a descrivere le funzionalità operative (dati, vincoli, procedimenti, comportamenti). I modelli che analizzeremo possono essere di livello concettuale (il più astratto, che riassume una descrizione generale della realtà) o di livello logico (meno astratto, più vicino a una rappresentazione di macchina).

Ogni modello deve specificare quali costrutti (elementi di rappresentazione) siano ammessi. Ogni costrutto è un modo per rappresentare qualcosa, un dato, un vincolo, un procedimento, un comportamento.

Schema

Uno schema è una tecnica di rappresentazione grafica (visiva) per illustrare la realtà secondo un modello. Ogni modello ammette alcuni schemi possibili. Uno schema è una raffigurazione della realtà, dove ogni elemento (costrutto) ha una sua rappresentazione grafica. Ogni costrutto dello schema può essere riprodotto sia a livello intensionale (schema che rappresenta in astratto quella realtà generica, es. una generica scuola) sia a livello estensionale (schema che rappresenta con fattispecie una istanza di quella realtà, es. questa scuola specifica).

Preso un modello di dati, per ogni costrutto dello schema occorre specificare:

- la pragmatica, ovvero il suo significato intuitivo, che determina il significato naturale dell'informazione
- la sintassi, le regole formali per descriverlo a livello intensionale
- la semantica, le regole di trascrizione per istanze di fattispecie a livello estensionale

Istanza

Un'istanza è una situazione specifica (concreta o simulata) che raffigura elementi dei costrutti in casi reali (o verosimili). Per istanze quindi si intendono degli esempi concreti di oggetti o eventi illustrati in uno schema. A seconda delle situazioni possiamo parlare di istanze di database (esempi di basi di dati con strutture concretamente realizzate, un database realizzato davvero), di istanze di entità (elementi concreti di categorie generiche, come quattro cani descritti opportunamente), di istanze di schemi logici (registrazioni di elementi).