

Basi di dati sanitari multimediali

Carlo Combi

Dipartimento di Matematica e Informatica
Universita' degli Studi di Udine

Sommario

- **Introduzione**

- **La multimedialita' dei dati clinici**
- **Basi di dati multimediali**
- **L'approccio orientato agli oggetti**

- **Sistemi di gestione di dati sanitari multimediali**

- **Gestione integrata di dati alfanumerici ed immagini**

- **Modellazione, interrogazione e viste per basi di dati multimediali orientate agli oggetti**

- **Modellazione**
- **Interrogazione**
- **Viste**

La multimedialita' dei dati clinici

Le informazioni cliniche possono essere espresse in differenti forme:

- **stringhe alfanumeriche**
risultati di esami, osservazioni a testo libero, storia clinica del paziente
- **suoni**
auscultazioni, registrazioni contenenti pareri di colleghi, commenti legati all'osservazione di parametri clinici
- **immagini statiche**
TAC, RM, PET, esami di medicina nucleare, di radiologia

La multimedialita' dei dati clinici

- **immagini dinamiche**
film angiografici, ecografie
- **segnali**
ECG, EMG, EEG
- **grafici**
esame audiometrico, esame spirometrico, rilevazione automatica o manuale di contorni sulle immagini
- **figure**
rappresentazione stilizzata di alberi coronarici, ricostruzioni bi- o tri-dimensionali di particolari anatomici

La multimedialita' dei dati clinici

- ◆ I diversi fini per cui tali informazioni sono raccolte,
 - ◆ la specializzazione presente nell'esaminare ogni tipo di tali dati,
 - ◆ le differenti provenienze sia anatomiche sia strumentali di tali informazioni,
 - ◆ l'elevata ed eterogenea correlazione fra i dati clinici,
- ⇒ hanno spinto a non considerare mai pienamente la gestione davvero completa ed omogenea di tali dati.

• Sul piano tecnologico, inoltre, la memorizzazione, la elaborazione e la trasmissione di tali informazioni solo in questi ultimi anni possono fare affidamento su un certo numero di soluzioni e di strumenti che si propongono come sensati a soddisfare le esigenze presenti in campo clinico.

Basi di dati multimediali

- aspetti generali -

L'impiego dei tradizionali servizi forniti da un sistema di gestione di basi di dati (DBMS: Data Base Management System) offre numerosi vantaggi nella gestione di grandi moli di dati multimediali:

- l'indipendenza dei dati dai programmi (astrazione sui dati);**
- l'accesso ai dati ad alto livello attraverso linguaggi di interrogazione;**
- l'indipendenza rispetto alle applicazioni;**
- il controllo dell'accesso ai dati da parte di piu' utenti nello stesso tempo (gestione della concorrenza);**
- la tolleranza ai guasti hardware e software (transazioni e ripristino);**
- la riservatezza nell'accesso ai dati.**

Basi di dati multimediali

- temi specifici -

- **modellazione dei dati e/o definizione di opportuni linguaggi di interrogazione;**
- **definizione di interfacce evolute per sistemi di basi di dati multimediali;**
- **memorizzazione fisica e ricerca sui dispositivi di memorizzazione dei dati stessi**
- **prestazioni globali di basi di dati distribuite nella memorizzazione e trasmissione di dati**
- **progettazione concettuale, logica e fisica di sistemi di basi di dati multimediali**

Modellazione e interrogazione di basi di dati multimediali clinici

Progettazione concettuale e logica di basi di dati attraverso il modello dei dati orientato agli oggetti

L'approccio orientato agli oggetti nelle basi di dati

- concetti di base e notazione -

- Un *oggetto* puo' modellare qualsiasi entita' del mondo reale, ad esempio un paziente, una terapia, un'immagine angiografica, e cosi' via. Oggetti complessi possono essere definiti sulla base di oggetti piu' semplici.
- La caratteristica fondamentale di un oggetto e' la sua identita', che e' immutabile, persiste per tutta l'esistenza dell'oggetto, e' indipendente dalle proprieta' dell'oggetto o dal suo comportamento.

L'approccio orientato agli oggetti nelle basi di dati

- concetti di base e notazione -

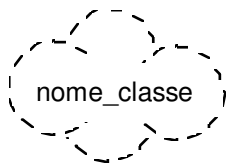
- Un oggetto e' caratterizzato da uno *stato*, espresso attraverso attributi, non accessibile dall'esterno, e da un'*interfaccia*, definita attraverso dei metodi, che consiste nella descrizione delle modalita' attraverso le quali si puo' agire sull'oggetto stesso. La caratteristica che consiste nel rendere non accessibile dall'esterno lo stato di un oggetto e' detta *encapsulation*.
- Gli oggetti vengono creati come istanze di una *classe*. Una classe caratterizza il comportamento delle sue istanze descrivendo i *metodi* che possono essere applicati agli oggetti. Da un punto di vista notazionale, il riferimento ad un metodo $m(\dots)$ applicato ad un oggetto o avverra' attraverso la notazione $o.m(\dots)$.

L'approccio orientato agli oggetti nelle basi di dati

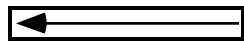
- concetti di base e notazione -

- Possono essere definite gerarchie di classi, attraverso relazioni di ereditarietà (IS_A) fra le classi. Ad esempio, dopo aver definito la classe *person*, contenente i dati che permettono di identificare una persona (cognome, nome, indirizzo), e' possibile definire la classe *patient*, che eredita dalla classe *person*, aggiungendo solo la descrizione dei dati tipici di un paziente (codice SSN, ricoveri, medico curante, ...).
- Ulteriori caratteristiche sono mutuare dai linguaggi orientati agli oggetti (overriding - overloading - late binding, completezza computazionale, estensibilita') e dalle basi di dati (persistenza, gestione della memoria secondaria, concorrenza, recovery, presenza di opportuni linguaggi di interrogazione).

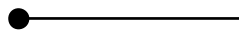
La notazione grafica di Booch



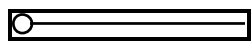
Classe



Relazione di ereditarieta' (IS_A)



Relazione di appartenenza (HAS_A)



Relazione di uso (USES)

Sistemi di gestione di dati sanitari multimediali

- **Disponibilita' di DBMS (Data Base Management System) in grado di considerare dati non alfanumerici: BLOB (Binary Large Object).**

- **La disponibilita' di DBMS orientati agli oggetti (OODBMS) ha reso possibile l'impiego di strumenti che permettono in fase di progettazione di definire in modo generale strutture di dati complesse, quali quelle legate alla memorizzazione di immagini, segnali, grafici, insieme alle operazioni che su tali dati sono ammesse.**

Gestione integrata di dati alfanumerici ed immagini

- **OSIRIS**
- **ARTEMIS-2**
- **PICQUERY+**
- **ARCADIA**

OSIRIS

- **Provenienza: Universita' di Ginevra.**
- **Scopo:** visualizzazione, elaborazione e gestione di immagini mediche.
- **Piattaforma HW/SW: inizialmente per macchine Unix con librerie X11 e OSF/Motif e per Macintosh con librerie MacApp: l'intero progetto e' stato sviluppato in C++. Attualmente OSIRIS e' presente su altre numerose piattaforme, fra le quali Windows NT e Windows 95.**
- **Caratteristiche:**
 - ◆ OSIRIS consente di gestire bioimmagini di qualsiasi tipo, purché esse siano scritte secondo il formato Papyrus, conforme allo standard ACR-NEMA.
 - ◆ L'interfaccia realizzata, basata su finestre, bottoni e menu', rende OSIRIS facilmente usabile anche da utenti senza una particolare preparazione di tipo informatico.
 - ◆ Il sistema permette anche la visualizzazione in una opportuna finestra di annotazioni sotto forma di testo libero, che facciano riferimento a particolari dell'immagine.

ARTEMIS-2

- **Provenienza: Broussais University Hospital di Parigi**
- **Scopo: gestione di informazioni provenienti da pazienti ipertesi; esso e' la seconda versione, orientata alla gestione di dati multimediali, di un sistema sviluppato ed impiegato in alcune centri per la cura dell'ipertensione nell'area di Parigi.**
- **Piattaforma HW/SW: workstation Unix (Digital, Sun, HP e IBM).**
- **Caratteristiche:**
 - ◆ gestione ed elaborazione di immagini angiografiche e delle relative informazioni grafiche ed alfanumeriche.
 - ◆ Controllo nei pazienti ipertesi delle anomalie presenti nelle arterie renali.

ARTEMIS-2

- integrazione dati e immagini -

- **sull'immagine angiografica sono applicati algoritmi di elaborazione, per una valutazione quantitativa delle eventuali lesioni presenti sulle arterie renali;**
- **i risultati dell'applicazione di tali algoritmi sono elaborati al fine di avere una descrizione del profilo del vaso arterioso in termini di regioni normali e anormali;**
- **tali descrizioni, opportunamente trattate, vanno a formare il risultato della fase di interpretazione dell'immagine stessa;**
- **il medico ha poi l'opportunita' di completare o aggiornare quanto ottenuto dal sistema, esprimendo in linguaggio naturale una descrizione dell'immagine;**
- **tale descrizione viene strutturata dal sistema in modo da permettere la definizione di interrogazioni anche sull'informazione inizialmente espressa in linguaggio naturale.**

PICQUERY+

- **Provenienza:** Università di Los Angeles (UCLA).
- **Scopo:** poter formulare interrogazioni complesse sia sui dati che sulle immagini. Come dominio di applicazione e di verifica esso è stato applicato alla medicina (KMED: Knowledge-based Medical Multimedia Distributed database system).
- **Piattaforma HW/SW:** base di dati distribuita in architettura client/server, ove sono presenti più sistemi: un sistema informativo ospedaliero, un sistema PACS, un sistema informativo di radiologia.
- **Caratteristiche:**
 - ◆ PICQUERY+ è un linguaggio di interrogazione per dati ed immagini che può essere impiegato in qualsiasi settore.
 - ◆ Impiego di un modello dei dati temporali, TEDM (Temporal Evolutionary Data Model), utile, ad esempio, per controllare l'andamento della crescita ossea di un paziente.

PICQUERY+

- interrogazioni su dati ed immagini -

- *"Ricerca un'immagine con la falange prossimale del quinto dito di un certo paziente e valuta la lunghezza dell'asse maggiore di tale osso".*

Questo genere di interrogazioni e' reso possibile grazie alla presenza di ulteriori elaboratori che hanno lo scopo di applicare alcuni modelli clinici di simulazione e di ricostruzione di immagini e di viste per estrarre caratteristiche da una immagine e verificare se essa soddisfi alle condizioni di selezione.

- **Sulla base di TEDM, PICQUERY+ permette di esprimere interrogazioni come: *"mostra le immagini relative alla fusione dell'epifisi metacarpale e dell'osso metacarpale tubolare del pollice e la loro evoluzione durante la crescita tra i periodi G ed H cosi' come definiti da Tanner e Whitehouse"*.**
- **Presenza di una interfaccia a finestre che consente agevolmente la formulazione di interrogazioni senza richiedere una dettagliata conoscenza del sistema e della sua architettura.**

ARCADIA

- A Realization of a Coronary-Artery Data and Image Administrator -

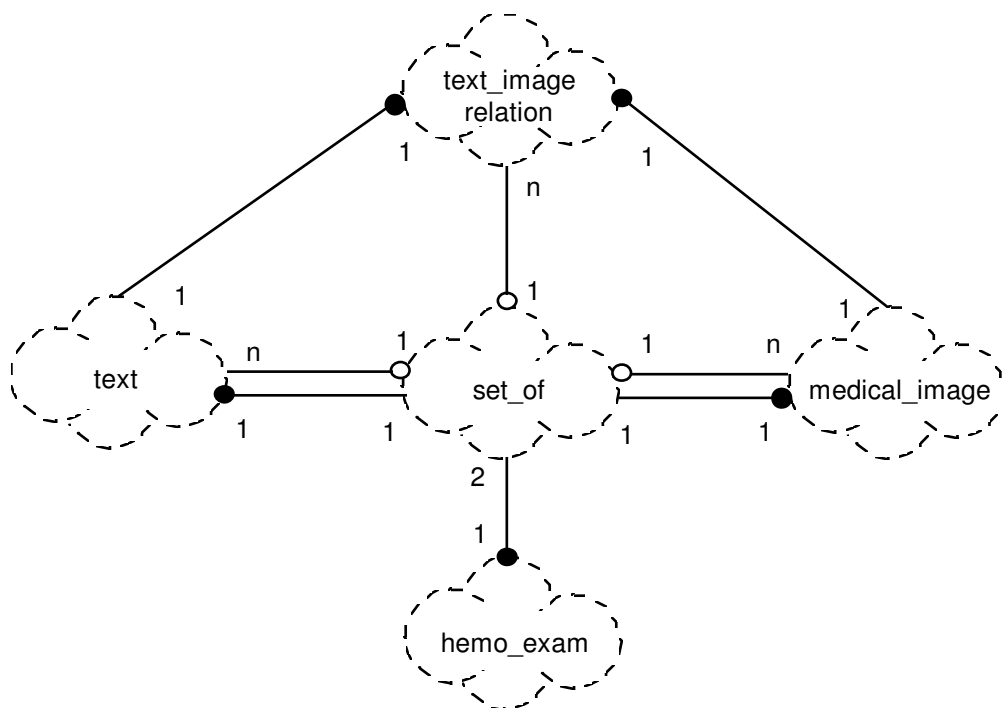
- **Provenienza: Politecnico di Milano, Laboratorio di Informatica Medica del Dipartimento di Bioingegneria.**
- **Scopo: gestione integrata di dati multimediali provenienti da pazienti sottoposti ad angioplastica coronarica.**
- **Piattaforma HW/SW: Sun workstation, Solaris.**
- **Caratteristiche:**
 - ◆ **La gestione dei dati multimediali avviene nel contesto di una cartella clinica orientata temporalmente, dove sono opportunamente considerati anche i dati anamnestici ed i parametri rilevati nel corso delle visite periodiche, caratterizzati dall'averne una dimensione temporale.**

ARCADIA

- integrazione di dati ed immagini -

- **L'attenzione e' stata posta inizialmente nella gestione delle relazioni fra le immagini angiocardiografiche e le diagnosi di stenosi su di esse effettuate.**
- **L'approccio e' completamente orientato agli oggetti sia per quanto riguarda la fase di progettazione e modellazione sia per quanto riguarda la fase di implementazione, fondata sull'utilizzo di un OODBMS.**
- **L'interfaccia utente e' fondata sull'utilizzo di finestre e di sistemi di puntamento, nell'ambiente grafico di Sun, fondato sulle librerie X11.**

*Il modello dei dati
adottato in ARCADIA per l'integrazione di
dati ed immagini statiche*



ARCADIA

- gestione dei dati multimediali -

- **immagini coronarografiche, dati relativi a stenosi diagnosticate, definizione della stenosi attraverso oggetti grafici;**
- **immagini ventricolografiche, dati relativi al ventricolo, sia quantitativi sia descrittivi, eventualmente ottenuti dall'applicazione di metodi quantitativi su contorni;**
- **film ventricolografico ed informazioni cliniche da esso desumibili, non legate a singole immagini.**

Modellazione, interrogazione e viste per basi di dati multimediali orientate agli oggetti - un esempio clinico -

Dati provenienti da follow-up di pazienti sottoposti ad intervento di angioplastica coronarica.

- **definizione grafica delle diagnosi effettuate sulle stenosi individuate;**
- **definizione della funzionalita' del ventricolo sinistro:**
 - ♦ **e' il medico che, visivamente, osservando il ventricologramma, stabilisce la funzionalita' ventricolare;**
 - ♦ **vengono usati metodi numerici applicati al contorno del ventricolo tracciato su due immagini opportunamente selezionate dal film.**
- **commenti generali legati a parti di film, ma non riconducibili a fotogrammi particolari, relativi, ad esempio, alla cinetica del muscolo cardiaco.**

Modellazione di dati multimediali

```
class visual_object {
public:
void display();
set_of_<observation> observations();
};

class observation {
public:
set_of_<visual_object> related_vis_obj();
};

class video_sequence:
public visual_object {
public:
set_of_<video_sequence>
                video_sub_sequences();
set_of_<image> images_of_video();
};

class image: public visual_object {
public:
set_of_<ROI> ROI_of_image();
};

class ROI: public image {
public:
polygon contours();
image rel_image();
boolean left(ROI);
boolean right(ROI);
boolean under(ROI);
boolean over(ROI);
};
```


};

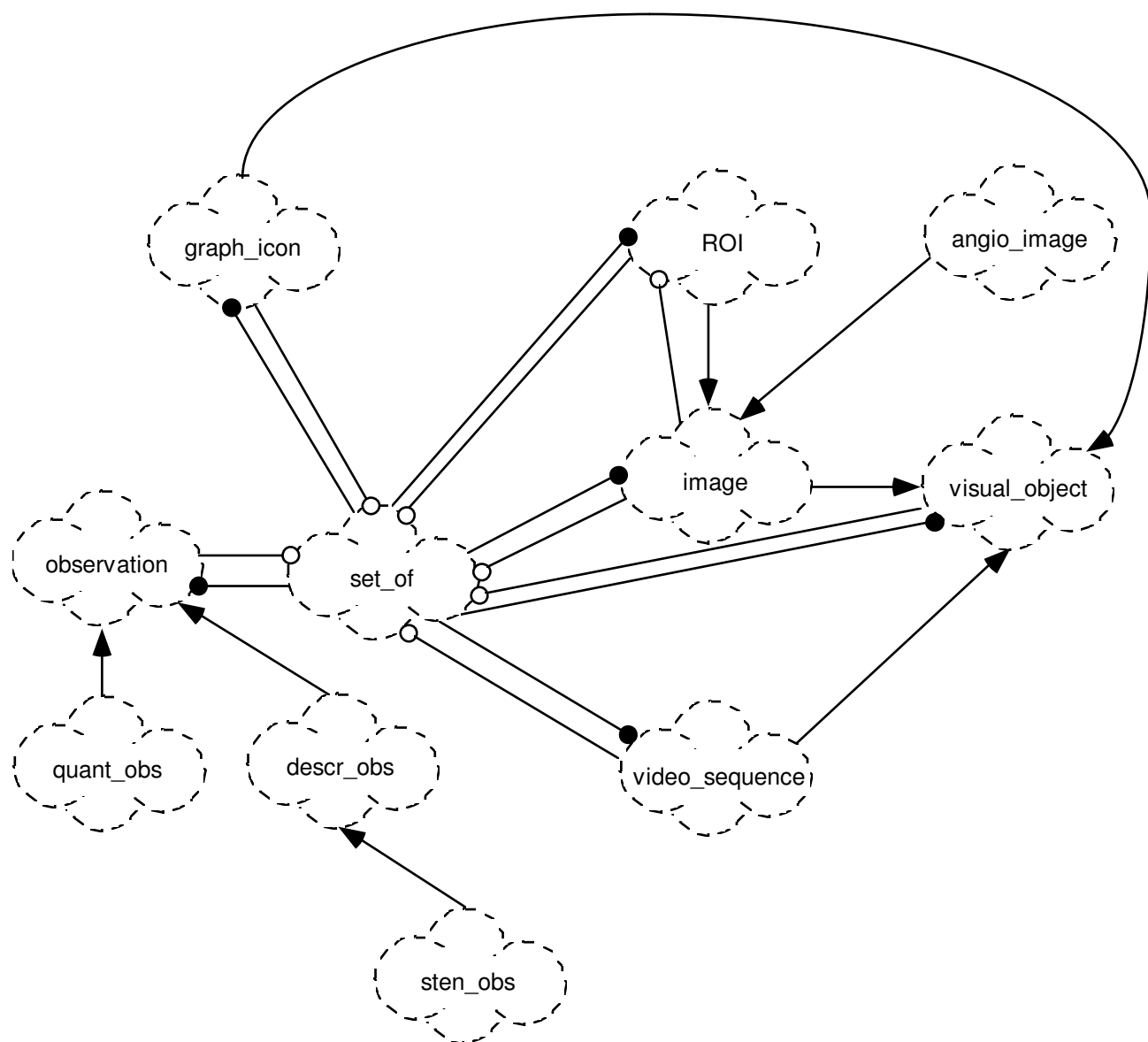
Modellazione di dati multimediali

```
class descr_obs: public observation {  
public:  
char* description();  
};
```

```
class sten_obs: public descr_obs {  
public:  
float intensity();  
boolean calcium_presence();  
char* lesion();  
char* location();  
.....  
.....  
};
```

```
class angio_image: public image {  
public:  
catheter cath();  
set_of_<sten_obs> observations();  
};
```

Il modello dei dati multimediali definito



Interrogazione di basi di dati multimediali

- La fase di presentazione dei dati, risultato dell'interrogazione, e' assai piu' importante e delicata di quanto avviene per i dati alfanumerici, usualmente riportati in forma tabellare.

*- composizione dei media,
operazioni interattive, sincronizzazione -*

- Nella definizione delle condizioni di interrogazione su immagini o film, non e' sempre elementare impiegare espressioni alfanumeriche ne' e' sempre possibile stabilire condizioni certe sui dati.

- interrogazioni ben-definite/interrogazioni vaghe -

Approcci:

- *Linguaggi di interrogazione per parole-chiave*
- *Linguaggi di interrogazione visuali*
- *Linguaggi di interrogazione semantici*

Un esempio di interrogazione

"Mostra il tratto di film relativo alla ventricolografia di tutti i pazienti con una stenosi di gravita' superiore al 75%".

```
select a.angiocardiography() .  
        get_ventricolography().display()  
from patient a, sten_obs b  
where a.angiocardiography() .  
        get_coronarography().rel_obs().  
        contains(b)  
        AND b.intensity() > .75
```

Viste multimediali

- **Visione sui dati a livelli di astrazione differenti**
 - ◆ **Ruolo dell'utente**
 - ◆ **Scopo nell'interazione con la base di dati**

- **Viste e basi di dati orientate agli oggetti**
 - ◆ **Viste materializzate o valutate run-time**
 - ◆ **Viste integrate nella gerarchia delle classi o viste definite con gerarchie separate**

Viste multimediali

Esempio: differenti viste associate alle immagini ed ai dati legati alla diagnosi di stenosi.

- Vista orientata allo *specialista*
 - ◆ **esigenza dell'utente: rappresentare tutto l'iter diagnostico a cui e' stato sottoposto il paziente**
 - ◆ **caratteristiche della vista: sono rappresentati i dati relativi a tutti gli esami emodinamici sostenuti, compreso l'intero film, alle diagnosi effettuate.**

- Vista orientata al *responsabile* della Divisione
 - ◆ **esigenza dell'utente: avere una sintesi che evidenzi il grado di stenosi residua dopo l'eventuale intervento di angioplastica coronarica**
 - ◆ **caratteristica della vista: vengono considerate e rappresentate solo poche immagini significative e le relative diagnosi.**