

LA QUESTIONE DEL TEMPO NEI DATI: **DATABASE TEMPORALI**

Andrea Gottardi

Introduzione

- ✦ La vita è scandita sempre più fortemente dal **tempo**
- ✦ I database convenzionali, nella versione generalmente utilizzata, non trattano il tempo
- ✦ I database temporali sono nati per colmare il (potenziale) divario tra situazione registrata e realtà

Evoluzione della tecnologia

- ✦ **Ben-Zvi** (1982) e **Clifford** (1983): prima formalizzazione
 - ✦ Tempo effettivo e tempo di registrazione
- ✦ **Snodgrass** (fine anni 90)
 - ✦ Miglioramento linguaggio QUEL
 - ✦ Estensione di SQL-92 (TSQL)
 - ✦ Fondamenti di SQL3

Concetti fondamentali

I database temporali sono caratterizzati da tre aspetti:

- ✦ Tipi di **dato** temporale
- ✦ Tipi di **tempo** utilizzabili
- ✦ Tipi di **interrogazioni**

I dati temporali

- **Istanti**

- Un momento preciso nel tempo

- **Intervalli**

- Una quantità di tempo definita

- **Periodi**

- La naturale unione tra istanti e intervalli
 - Hanno una gestione particolare (relazioni, rappresentazione)

Tipi di tempo

- ✦ “User-defined time”
 - ✦ Sono paragonabili a qualsiasi altro campo
- ✦ Tempo di transazione
 - ✦ Quando la modifica viene registrata
- ✦ Tempo di validità
 - ✦ Quando la modifica vale effettivamente

Tipi di interrogazioni

- **Interrogazioni non temporali**

- Utilizzate nei database convenzionali

- **Interrogazioni correnti**

- Prendono in considerazione l'istante attuale

- **Interrogazioni sequenziali**

- Considerano tutto il tempo conosciuto

- **Interrogazioni non sequenziali**

- Danno importanza solo ai valori diversi dal tempo

Problemi da risolvere

Rispetto alle tabelle tradizionali, quelle temporali presentano alcuni problemi da risolvere:

- ✦ Definizione della **chiave primaria**
 - ✦ Non è sufficiente collegarla ai campi di dati
- ✦ Definizione dei record attualmente **validi**
 - ✦ Tre modi: NULL, data precedente, data futura
 - ✦ Non sempre sono fedeli alla realtà

Problemi da risolvere

- ✦ Definizione dell'**unicità** dei record
 - ✦ Ci sono quattro diversi tipi di duplicati
 - ✦ Alcuni sono accettabili, o addirittura obbligatori
 - ✦ Altri sono da evitare, o addirittura inutili

Tabelle temporali

- ✦ Sono tre tipi, e presentano particolarità diverse:
 - ✦ Tabelle con tempo di **validità**
 - ✦ Tabelle con tempo di **transazione**
 - ✦ Tabelle **bitemporali**

Aspetti comuni delle tabelle

- Vengono ottenute aggiungendo una o due coppie di **campi** alle tabelle tradizionali
- I campi rappresentano un periodo, in vari modi:
 - Coppia di istanti
 - Istante e intervallo
- Possono ammettere **duplicati** con valore equivalente
- Le operazioni canoniche devono subire **modifiche** notevoli

Tabelle con tempo di validità

- ✦ Registrano il tempo in cui il record ha validità **reale**
- ✦ Ammettono tutti i tipi di interrogazione

RESIDENZA				
ID	Nome	Residenza	Start	End
1	Mario Rossi	Trento	03.05.1984	31.12.9999
2	Giorgio Bianchi	Lavis	06.03.1992	20.07.2014
2	Giorgio Bianchi	Trento	20.07.2014	31.12.9999

Tabelle con tempo di transazione

- ✦ Registrano il tempo in cui il record si “ritiene valido”
- ✦ Ammettono solo interrogazioni **correnti**, con lo scopo di tenere traccia di ogni modifica
- ✦ Per questo motivo non è possibile correggere gli errori

RICH_RESIDENZA				
ID	Nome	Residenza	Tran_Start	Tran_End
1	Mario Rossi	Trento	28.04.1984	31.12.9999
2	Giorgio Bianchi	Lavis	03.03.1992	17.07.2014
2	Giorgio Bianchi	Trento	17.07.2014	31.12.9999

Tabelle bitemporali

- Unione tra le tabelle con tempo di validità e tabelle con tempo di transazione
- Hanno una gestione più potente, ma anche più complessa:
 - Come con le tabelle con tempo di transazione, anche qui le operazioni subiscono notevoli cambiamenti per ottenere lo stesso risultato
 - Non sono molto utilizzate: la difficoltà di gestione rende preferibile usare due tabelle, più flessibili e facili da mantenere
 - Consentono di risparmiare spazio e utilizzare meno operazioni, anche se sono più complesse

Tabelle bitemporali

PERSONALE

ID	Nome	Grado	Val_Start	Val_End	Tran_Start	Tran_End
1	Luca Verdi	2	01.07.2005	31.12.9999	10.06.2005	01.11.2014
2	Paolo Rossi	2	01.02.2009	31.12.9999	11.01.2009	15.07.2014
2	Paolo Rossi	2	01.02.2009	01.08.2014	15.07.2014	31.12.9999
2	Paolo Rossi	1	01.08.2014	31.12.9999	15.07.2014	31.12.9999
1	Luca Verdi	2	01.07.2005	01.01.2015	01.11.2014	31.12.9999

Novità in SQL-2011

SQL-2011 ha introdotto molte innovazioni in ambito temporale:

- Supporto ai periodi (costrutto **PERIOD**)
- Supporto adattato alle **relazioni** tra periodi
 - Unite relazioni logicamente simili
- Supporto alla gestione **automatica** di tempi di validità e transazione

Fondazione Bruno Kessler

- ✦ FBK utilizza i database temporali, anche se con alcune varianti rispetto alla teoria originaria
 - ✦ Vengono usate solo tabelle con tempo di validità, senza backup (ottenuti in altro modo)
 - ✦ Per questo motivo non vengono utilizzate nemmeno le tabelle bitemporali

Conclusioni

- ✦ I database temporali sono una soluzione fondamentale se è necessario avere dati coerenti in ogni momento
- ✦ Le strategie di utilizzo, a seconda delle esigenze, possono essere molto diverse
- ✦ Molte sono gli aspetti ancora da perfezionare, ma l'ultima versione dello standard ha permesso notevoli miglioramenti

GRAZIE PER L'ATTENZIONE