

Tesi di laurea

Reingegnerizzazione ed estensione di uno strumento per la generazione di siti Web

Candidato
Gianluca Di Tomassi
matr. 60018/17

Relatore
Prof. P. Atzeni
Università di Roma Tre

Correlatore
Ing. P. Merialdo
Università di Roma Tre

Contesto

Progetto di ricerca **Araneus**: definizione e sviluppo di metodologie e strumenti per la gestione dei dati sul WWW.

Oggetto della tesi: realizzazione di un nuovo prototipo per la generazione di siti Web

Sommario

- L'approccio Araneus per la generazione di siti Web
 - Il modello ADM
 - Il linguaggio Penelope
- Limiti del prototipo iniziale
- Reingegnerizzazione ed estensioni
- Conclusioni e sviluppi futuri

Il modello ADM

- Il modello ADM è orientato alle pagine
- Un page-scheme descrive l'organizzazione di un insieme omogeneo di pagine Web
- Un'istanza di un page-scheme è una pagina Web

Due classi di attributi:

- Attributi semplici
Text, Image, Link-To
- Attributi composti
List-OF, MAP, FORM

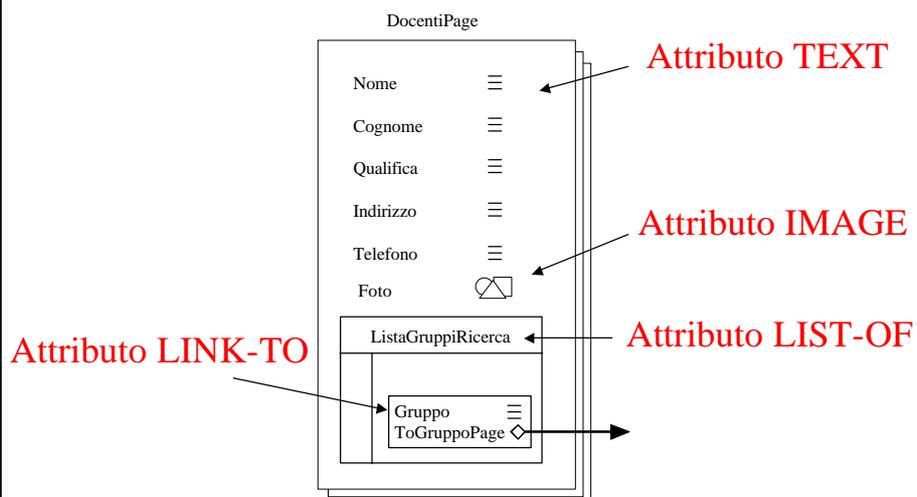
Una pagina Web



Gianluca Di Tomassi

5

Esempio di page-scheme



Gianluca Di Tomassi

6

Il linguaggio Penelope

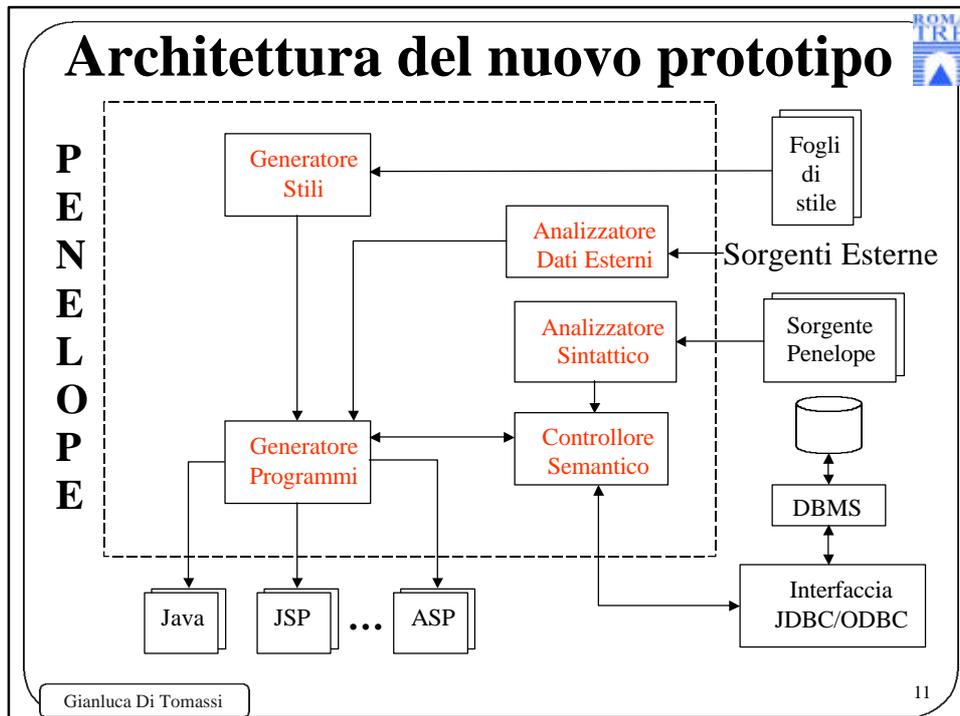


- È un linguaggio dichiarativo
- È basato sul modello ADM
- Permette di generare siti Web (in formato HTML o XML) a partire da dati memorizzati su database relazionali.

ESEMPIO



```
DEFINE PAGE DocentiPage : <Professore.Nome>
AS URL(<Professore.Nome>);
Nome : TEXT = <Professore.Nome>;
Cognome : TEXT = <Professore.Cognome>;
.....
Foto : IMAGE = <Professore.Foto>;
ListaGruppiRicerca : LIST-OF (
    LINK-TO GruppoPage (
        URL(<NomeGruppo>)
        Gruppo : TEXT = <NomeGruppo>;
    );
);
USING Professore,
GruppiRicerca : (SELECT NomeGruppo, Professore.Nome
FROM Professore, Gruppi
WHERE Professore.Ricerca = Gruppi.Settore )
END
```

Descrizione moduli

L'Analizzatore Sintattico: verifica la correttezza sintattica delle istruzioni del sorgente Penelope

Il Controllore Semantico verifica la correttezza delle relazioni tra lo schema della base di dati e quello del sito

Il Generatore di stili trova le associazioni tra la presentazione grafica e gli attributi delle pagine

Analizzatore dati esterni gestisce dati che possono essere generati dinamicamente da sorgenti esterne

Il Generatore di programmi crea fisicamente i programmi a partire dal sorgente Penelope

12

Il codice generato

- **Generazione programmi Java** (1 per ogni page-scheme)
 - Portabilità
 - Per generare il sito non occorre disporre del prototipo
 - **Generatore pagine dinamiche**
 - JSP (Java Server Page)
 - ASP (Active Server Page)
- OSSERVAZIONE** : L'algoritmo proposto, nella sua generalità, è applicabile per la generazione di qualsiasi tipo di implementazione.
- Occorre gestire connessioni efficienti con il database

Gestione delle connessioni al database

Generazione di pagine dinamiche \Rightarrow *Accessi concorrenti in tempo reale al database*

Studio di una soluzione efficiente

- Definizione ed implementazione di un pattern:
DBConnectionManager

Gradi di libertà :

- (1) Numero max di connessioni
- (2) Numero di connessioni da utilizzare per gestire un elevato numero di richieste concorrenti

Fonti dati eterogenee per generare pagine Web



- Acquisizione dei dati da una qualsiasi sorgente esterna disponibile sul web.

```
DEFINE PAGE DocentiPage : <Professore.Nome>
AS URL(<Professore.Nome>);
Nome : TEXT = <Professore.Nome>;
.....
USING Professore: (SELECT NomeGruppo, Professore.Nome
                   FROM http://vesuvio.dia.it/professore
                   WHERE ...
                   )
....
```

- Formato standard di scambio: XML

Generazione di pagine dinamiche con dati estratti dinamicamente da sorgenti esterne e dal database

Algoritmo di generazione



Algoritmo di generazione *PRECEDENTE* :

- Materializzazione delle viste definite dall'utente
- Definizione dei join tra le tabelle del database
- Generazione dei cursori contenenti i dati di interesse
- Generazione delle pagine web

Algoritmo di generazione *ATTUALE* :

- Definizione dei join tra le viste definite dall'utente e le tabelle del database
- Definizione dei cursori che conterranno i dati di interesse
- Generazione dei programmi (java, jsp, asp, ...)

Comparazione tra i due algoritmi

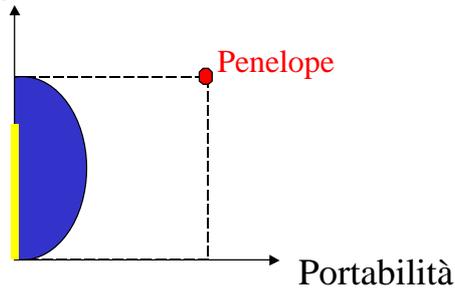
Page-scheme	Cardinalità	Livelli di nidificazione	Dimensione media pagina (kb)	Algoritmo precedente (ms)	Algoritmo attuale (ms)	Variazione %
IndexTerms	922	1	4	164.116	108.035	52%
OrdinaryIssue	51	3	15	99.984	25.566	291%
Proceedings	18	3	65	112.061	40.388	177%

S.O. utilizzato : Windows NT

DBMS utilizzato : MS Access 97

Scenario

Flessibilità
(struttura - grafica)



- Prodotti RAD
- Prodotti messi a disposizione dai DBMS

Conclusioni



- Analisi e studio del prototipo Penelope esistente
- Studio delle tecnologie e dei formati di nuova generazione per la generazione di pagine Web
- Definizione e test di nuovi algoritmi per la generazione di pagine
- Definizione delle specifiche, dell'architettura e della implementazione del nuovo prototipo Penelope (Realizzato in Java circa 22000 righe di codice)
- Estensione del prototipo per la gestione di dati proveniente da sorgente dati dinamiche
- Applicazione Web per la pubblicazione, in diversi formati, dei dati estratti da ACM SIGMOD

<http://vesuvio.dia.uniroma3.it/applicazioni>

Gianluca Di Tomassi

19

Applicazione Web - ACM Sigmod



Year	Month	Volume	Number	Notes
1999	March	Vol. 26	No. 1	
	June	Vol. 26	No. 2	
	September	Vol. 26	No. 3	
	December	Vol. 26	No. 4	
1997	March	Vol. 25	No. 1	
	June	Vol. 25	No. 2	ACM SIGMOD International Conference on Management of Data 1997
	September	Vol. 25	No. 3	
	December	Vol. 25	No. 4	

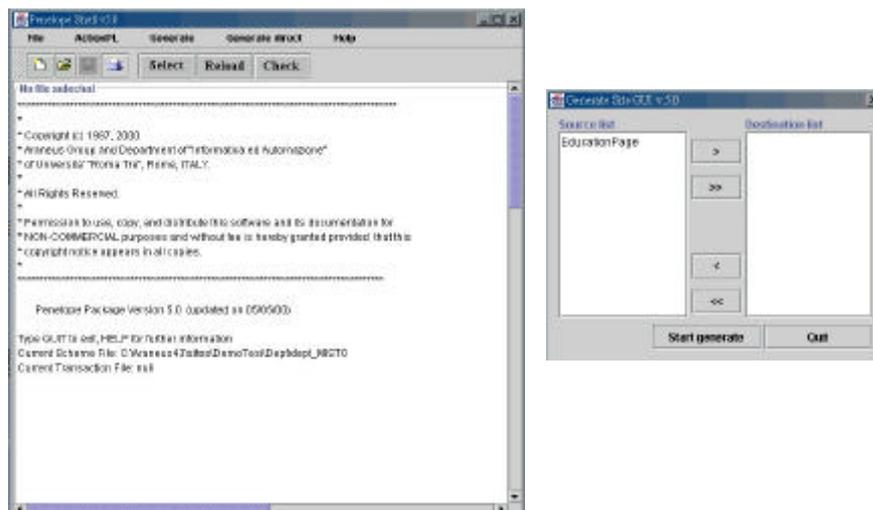
Gianluca Di Tomassi

20

Sviluppi futuri

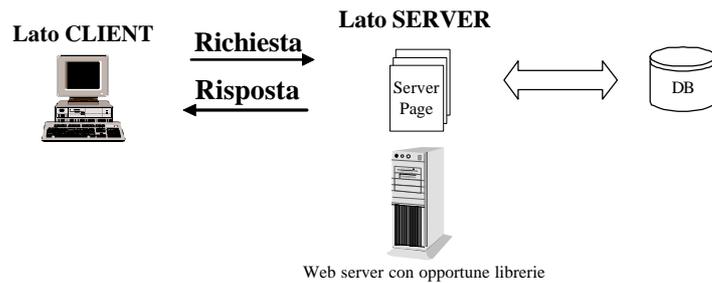
- Generazione di sorgente WML (Wireless Markup Language)
- Riscrittura del parser Penelope
- Generazione di un sito parallelo che permetta gli aggiornamenti della base di dati on-line

Interfacce grafiche implementate



Le Server Pages

- Pagine all'interno delle quali è possibile inserire codice
- Il server compila il codice e lo esegue



⇒ Pagine sempre aggiornate rispetto alla base di dati

