Generating New Course Material
From Existing Courses

Vincent Oria
Department of Computer and Information Science
New Jersey Institute of Technology
Newark, NJ 07102-1982

Joint work with:
Silvia Hollfelder and Peter Fankhauser, GMD-IPSI
Motivations

- Specialists in some domains (e.g. IT) are not easy to find
- Worker in evolving domains (e.g. IT) need to regularly update their knowledge
  - distance learning or part-time study
    - For worker seeking new degrees
  - Training courses
    - Where to find the appropriate content?
Courseware-on-Demand and Teachware-on-Demand

(modifiable) teaching materials

Choosing and structuring of right fragments

Course designer

Author

Integration of teaching materials

new teaching material

Current situation

Courseware on Demand

repository
Teachware-on-Demand: a cooperation of Fraunhofer ISST, GMD IPSI and Fraunhofer IESE supported by the Deutsche Telekom AG, Control Data Institute and HTTC. Teachware on Demand is funded by the German Ministry of Education and Research (bmb+f).

Courseware-on-Demand: a research project at NJIT in cooperation with Fraunhofer ISST (Germany), GMD-IPSI (Germany) and University of Waterloo (Canada).
Outline

- Background: Standards and Tools
  - Multimedia Databases and MPEG-7
  - Metadata Standards and Learning Objects
- Courseware-on-Demand
  - System Architecture and Metadata
  - Indexing and Querying
  - Distribution and Interoperability
What are multimedia data?

- MM data
  - Text Data
  - Image
  - Video
  - Audio
  - Graphics
  - Generated media (Animation, midi)

- Common characteristics
  - Size of the data (in term of bytes)
  - Real-time nature of the information content
  - Raw or uninterpreted
Multimedia Database Architecture

Additional Information

MM Data Preprocessor

Recognized components

MM Data Instance

MultiMedia Data Preprocessing System

MultiMedia DBMS

Query Interface

Users

MM Data

Meta-Data

Database Processing
Multimedia Database Architecture: Documents

DTD files
- DTD Parser
- DTD Manager
- Type Generator

XML or SGML Document Instance
- SGML/XML Parser
- Parse Tree

Instance Generator
- C++ Types
- C++ Objects

Multimedia DBMS
- Query Interface
- Multimedia DBMS

Users

Document Processing System

Database Processing
Multimedia Database Architecture: Video

- Video Analysis and Pattern Recognition
- Key Frames
- Meta-Data
- Video Annotation
- Image
- Video Processing System
- Multimedia DBMS
- Query Interface
- Users

Video Content Description
Video
Multimedia DBMS
Database Processing
Multimedia Data and Metadata

- **Technical Metadata**
  - camera movements, number of scenes, ...

- **Editorial Metadata**
  - e.g., author, date, ...

- **Semantic Metadata**
  - content, persons, objects, relationships, ...
MPEG-7 Objectives

MPEG-7, formally called "Multimedia Content Description Interface", will standardise:

- A language to specify description schemes, i.e. a Description Definition Language (DDL).
- A set of Description Schemes and Descriptors A scheme for coding the description

http://www.csel.it/mpeg/standards/mpeg-7/mpeg-7.htm
MPEG 7 Context and Objectives

- Content Description
  - format independent
  - may be applied to analogue media
  - different description granularities

- Supplementary Data

- Application Types
MPEG-7 Framework

- **Description Generation**
- **MPEG7 Description**
- **Encoder**
- **MPEG7 Coded Description**
- **Decoder**
- **MPEG7 Description Definition Language (DDL)**
- **MPEG7 Description Schemes (DS) & Descriptors (D)**
- **Search / Query Engine**
- **Filter Agents**
- **User**
Where are We?

- Background: Standards and Tools
  - Multimedia Databases and MPEG-7
  - Metadata Standards and Learning Objects
- Courseware-on-Demand
  - System Architecture and Metadata
  - Indexing and Querying
  - Distribution and Interoperability
IEEE Learning Object Metadata (LOM)

- Defined by the IEEE Learning Technology Standards Committee (LTSC)
- Builds on the metadata work done by the Dublin Core group [http://purl.org/dc](http://purl.org/dc)
- Objective: Propose a structured metadata model for learning objects
  - syntax, semantics
- LOM supports security, privacy, commerce, and evaluation
A learning object is an entity, digital or non-digital, that can be used, re-used or referenced during technology-supported learning.

- Learning objectives, persons, organizations or events

A learning object can have more than one description.
The Base Scheme is composed of 9 categories:

- General
  - context-independent features and semantic descriptors
- Lifecycle
  - features linked to the lifecycle of the resource
- Meta-metadata
  - features of the description itself
- Technical
  - technical features of the resource
LOM Metadata Structure (cont)

- Educational
  - educational and pedagogic features of the resource

- Rights
  - features dealing with condition of use

- Relation
  - link to other resources

- Annotation
  - comments on the educational use

- Classification
  - characteristics of the resource described by entries in classifications
IMS Metadata

Record
  general
    title
      langstringtype
        language: “en-US”
        string: Becoming a Metadata Expert
    catalogentry
    lifecycle
      (version, status, etc)

Catalog: “ISBN”
entry: “0-13-659707-6”
Background: Standards and Tools (Conclusion)

- Multimedia databases
  - access and extract part of learning objects
- MPEG-7
  - use to describe the audio-visual content of learning objects
- IEEE LOM and IMS Metadata
  - provide metadata for learning object
  - needs to be extended to learning fragments
Where are We?

- Background: Standards and Tools
  - Multimedia Databases and MPEG-7
  - Standards and Learning Objects
- Courseware-on-Demand
  - System Architecture and Metadata
  - Indexing and Querying
  - Distribution and Interoperability
Courseware-on-Demand: Architecture

- Course Designer
- Layout and Presentation
- Goal Selector
- Sub-Graph Selector
- C. M. Extractor and Composer
- Query Processor
- Metadata Manager
- Type System
- Indexer
- Course Material Manager
- Object Repository
- Meta-Data
- Slides, Video, Audio, Text, Meta-Data
- Course Materials
A learning fragment is a learning unit of a course

- sequence of learning pieces (notions)
- has different versions (e.g. overview, short, long)

Level of granularity

- depends on
  - the course
  - the author

Intuitively the finer, the better
- Need of experimental results
Fragmenting Course Content

contains

Course or
course module

Elementary
fragment
Fragmenting Course Content: Correctness Rules

- Completeness:
  - If a learning object $L$ is decomposed into $L_1$, $L_2$, ..., $L_n$, every learning notion in $L$ should also be found in one or more of $L_i$’s

- Reconstruction:
  - If a learning object $L$ is decomposed into $L_1$, $L_2$, ..., $L_n$, it should be possible to define an order $r$ such that $r(L_i)=L$

- Disjointness:
  - If a learning object $L$ is decomposed into $L_1$, $L_2$, ..., $L_n$, and a learning notion $l_i$ is in $L_j$, it is not in any other fragment $L_k$ (k≠j)
Prerequisite and Precedence can be modeled by the temporal relationship “Before”

Pushed to the fragments

A node A is before a node B if there exist a fragment a (ANC(a)=A), a fragment b (ANC(b)=B) and a “Before” b
Sharing The Same Fragments

A contains B

Course or course module

Elementary fragment
A Course with Alternatives

C = D \land (E \lor F)
Logical and Physical Learning Fragments

- Logical fragment
- Course or course module

Diagram:
- Physical Fragments
- Contains
- Logical fragment
- Course or course module

Legend:
- Physical Fragments
- Logical fragment
- Course or course module
- Contains
Fragmenting Learning Objects

- Decompose a learning object based on their learning content
  - a new type of fragment (physical learning fragment)
  - the fragment mentioned before (logical learning fragment)

- Correctness
  - Similar to the correctness rules defined for the logical learning objects

- “Referencial Integrity”
  - Every physical learning fragment should refer to an existing logical learning fragment
Where are We?

- Background: Standards and Tools
  - Multimedia Databases and MPEG-7
  - Standards and Learning Objects
- Courseware-on-Demand
  - System Architecture and Metadata
  - Indexing and Querying
  - Distribution and Interoperability
Indexing and Querying

- Representation: XML
  - Hierarchical structure of courses
  - Interoperability
- Indexing
  - Indexing XML document
  - Indexing learning objects on logical fragments
Chapter

Section 1
Para 1
Hypertext
Internet
Multimedia

Section 2
Para 2
Hypertext
Internet
Java

Hypertext
Browser

BUS Document Tree with Index Terms
BUS Accumulation Method

- Elements accumulated into parents from bottom up until user query level reached
- Parent Unique Element Identifier (UID) calculated with following formula

\[
\text{Parent}(\text{UID}) = \frac{\text{UID} - 2}{k} + 1
\]

where \( k = \text{maximum number of children} \)
BUS Limitations

- Storage overhead - 240% of original document size
- Indexing time is long - over 4 hours for 250 MB
- Query time is long - up to 6.5 seconds
- Inefficient update method - sometimes have to modify entire indexing system
- No ability to do similarity queries
Our Solutions

- Reduce storage overhead and indexing time - Index fewer keywords
- Reduce query time - Traverse smaller tree
- Create efficient update method - Remove the need for a fixed number of maximum children
- Allow for similarity queries - Create hierarchical dictionary, create concept hierarchy, use hierarchical queries
Some document updates cause $k$ to be increased.

Requires re-indexing entire document to assign new UIDs.

Solution: Eliminate need for calculation and therefore $k$.

How? Collect elements by storing Parent ID with the element in the database.
Concept Hierarchy and Hierarchical Dictionary

- Dentistry
  - Orthodontics
  - Periodontics
- Surgery
- Neurology
- Pediatrics
- Journals
- Transplants
- Brain Surgery
- Growth Disorders
  - Joints
  - Muscles
  - Bones
- Behavioral Disorders
Hierarchical Query Example

```
SELECT subject, parent_subject, level
FROM subject_relationships
CONNECT BY PRIOR parent_subject = subject
START WITH subject = 'Pediatrics'
```

Result:

<table>
<thead>
<tr>
<th>Subject</th>
<th>Parent</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pediatrics</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Growth Disorders</td>
<td>Pediatrics</td>
<td>2</td>
</tr>
<tr>
<td>Joints</td>
<td>Growth Disorders</td>
<td>3</td>
</tr>
<tr>
<td>Muscles</td>
<td>Growth Disorders</td>
<td>3</td>
</tr>
<tr>
<td>Bones</td>
<td>Growth Disorders</td>
<td>3</td>
</tr>
<tr>
<td>Behavioral Disorders</td>
<td>Pediatrics</td>
<td>2</td>
</tr>
</tbody>
</table>
Querying

- Selecting the goals: Querying XML documents
- Finding a subgraph in the fragment network that contains all the concepts in the goal
  - exponential complexity
- Bottom-up approach [Caumanns 98]
  - Selection of the most appropriate fragments
  - Sequencing fragments
- Extraction and composition of new learning objects
Where are We?

- Background: Standards and Tools
  - Multimedia Databases and MPEG-7
  - Standards and Learning Objects
- Courseware-on-Demand
  - System Architecture and Metadata
  - Indexing and Querying
  - Distribution and Interoperability
Distribution and Interoperability

Course Designer

Global Request Manager

Global vocab. catalog

Integrated request result (XML)

Local Request Manager

Local vocab. catalog

Request (XML)

Request result (XML)

Query processor

Teaching Material Repository

Local Request Manager

Local vocab. catalog

Request (XML)

Request result (XML)

Query processor

Teaching Material Repository
The research work includes:

- the definition of the metadata model and implementation of the metadata type system;
- the development of new indexing tools;
- development of a new query processor that combines traditional query techniques and path theory;
- development of a distributed and interoperable middleware to integrate several distributed teaching material repositories.