Semi-automatic Generation of Learning Domain Modules for Technology Supported Learning Systems

Miguel Larrañaga Olagaray

Supervisors: Ana Arruarte Lasa and Jon A. Elorriaga Arandia

December 2012
Outline

1. Introduction
2. Acquisition of the Domain Module
4. Evaluation
5. Final Remarks: Conclusions and Future Lines
Outline

1 Introduction

2 Acquisition of the Domain Module

3 Dom-Sortze: a Framework for the Semi-automatic construction of the Domain Module

4 Evaluation

5 Final Remarks: Conclusions and Future Lines
Technology Supported Learning Systems

- More and more Technology Supported Learning Systems are being used
- Learning Management Systems such as Moodle or Blackboard
- Massive Open Online Courses
  - coursera.org ≈ 200 courses
- Intelligent Tutoring Systems - Algebra, SQL …
Domain Module

- Core of the Technology Supported Learning Systems
- Domain topics/learning units, pedagogical relationships, didactic resources . . .
- It also allows Technology Supported Learning Systems to plan the learning sessions and to evaluate the performance of the learners
Development of the Domain Module (I)

Entails . . .

- Identifying the topics to be mastered

- Organising the structure of the course based on pedagogical principles
  - Structure of the topics
  - Sequence of the topics
  - . . .

- Providing resources (definitions, examples . . .) that allow learning about each topic
However . . .

- Time and effort consuming
- The development of the Domain Module takes over 50% of the cost (Anderson, 1988)
- Authoring tools are usually not intended for “average” teachers

(Murray, 1999)

“The average teacher should not be expected to design ITSs any more than the average teacher should be expected to author a textbook in their field”
Idea

Electronic textbooks can be used to build the Domain Module.
Electronic textbooks can be used to build the Domain Module.
Idea

Electronic textbooks can be used to build the Domain Module.
Electronic textbooks can be used to build the Domain Module.
Planets are non-self luminous celestial bodies orbiting around a star.

Def.

Prob.

Ex.

Electronic textbooks can be used to build the Domain Module
Goals

Aim of this work

Automatise the generation of the Domain Module

- Collaborative supervision
- Knowledge reuse
- Domain-independence
- Language-independence
Approach

Knowledge Representation

- Ontologies
  - Interoperability (Uschold & Grunninger, 1996)
- Standards and Specifications
  - Learning Objects: “Reusable pieces of educational material intended to be strung together to form larger educational units such as activities, lessons or whole courses” (Brooks, 2003)

Knowledge Acquisition

- Ontology-guided processes
- Heuristic Reasoning
- NLP Techniques
  - The Basque Language
Working Methodology

Methodology

- Manually analyse small sets of documents
- Identify patterns/heuristics that facilitate knowledge elicitation
- Test on a wider set of documents
- Gold Standard-based evaluation
Outline

1. Introduction

2. Acquisition of the Domain Module


4. Evaluation

5. Final Remarks: Conclusions and Future Lines
Domain Module

Representation

- Learning Domain Ontology
  - Topics
  - Pedagogical relationships
    - Structural: isA and partOf
    - Sequential: prerequisite and next

- Learning Objects

Assumptions

- Topics are referred in the document
- Structure of the document underlies pedagogical relationships
- Didactic resources provided throughout the body of the textbook

Reflected in the outline + body of the textbook
General Process for the Acquisition of the Domain Module

1. **Electronic Textbook**
2. **Preprocess**
   - Document Outline Internal Representation
   - Document Internal Representation
3. **LDO Gathering**
   - Learning Domain Ontology
4. **LOs Gathering**
   - Domain Module
General Process for the Acquisition of the Domain Module

1. Preprocess
   - Electronic Textbook
   - Document Outline Internal Representation
   - Learning Domain Ontology
   - Document Internal Representation

2. LDO Gathering
   - LOs Gathering
   - Domain Module
### Some aspects to be considered

- **Document formats**
  - pdf
  - doc/docx
  - odt
  - ...

- **Document outlines**
  - Numbering/indentation
  - Location
    - At the beginning
    - At the end
Preparing the Textbook for the Knowledge Acquisition Processes (II)
Preparing the Textbook for the Knowledge Acquisition Processes (II)

Preprocess

1. Document Outline Extraction
2. Internal Representation Building
3. Part-of-Speech Analysis

Electronic Textbook

Document Outline Internal Representation

Lemma, morphological category...

Document Internal Representation
General Process for the Acquisition of the Domain Module

1. Electronic Textbook
2. Preprocess
   - Document Outline Internal Representation
   - Learning Domain Ontology
3. LDO Gathering
4. LOs Gathering
5. Domain Module
Gathering the Learning Domain Ontology

Entails

- Extracting the topics
- Extracting structural relationships
- Extracting sequential relationships

Similar to the ontology learning approach described by (Maedche & Staab, 2000)
Gathering the Learning Domain Ontology
Gathering the Learning Domain Ontology

Document Outline Internal Representation

Outline Analysis

Learning Domain Ontology

Whole Document Analysis

Enriched Learning Domain Ontology
Analysis of the Document Outline

**Assumptions**
- Each outline item represents a unique domain topic
- The structure of the outline underlies pedagogical relationships

**Procedure**
- Basic analysis
  - Identify topics and candidate relationships
- Heuristic analysis
  - Identify and categorise pedagogical relationships
    - Structural relationships: isA, partOf
    - Sequential relationships: next, prerequisite
Identifying the Structural Relationships

**Structural Relationships**
- Between an outline item and its subitems
- By default, partOf

**Heuristics**
- Group heuristics
  - 2 heuristics (e.g., CHe+MWH)
- Individual heuristics
  - 5 heuristics for isA (e.g., MWH or AH) + 1 for partOf (PGH1)
## Group Heuristics: Example

### Common Head + Multi-Word Heuristic (CHe+MWH)

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2 Zenbakizko taxonomia</td>
<td>5.2 Numerical classification</td>
</tr>
<tr>
<td>5.2.1 Clustering banatzailea</td>
<td>5.2.1 Exclusive clustering</td>
</tr>
<tr>
<td>5.2.2 Clustering hierarkikoa</td>
<td>5.2.2 Hierarchical clustering</td>
</tr>
</tbody>
</table>
Group Heuristics: Example

Common Head + Multi-Word Heuristic (CHe+MWH)

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2 Zenbakizko taxonomia</td>
<td>5.2 Numerical classification</td>
</tr>
<tr>
<td>5.2.1 Clustering banatzailea</td>
<td>5.2.1 Exclusive clustering</td>
</tr>
<tr>
<td>5.2.2 Clustering hierarkikoa</td>
<td>5.2.2 Hierarchical clustering</td>
</tr>
</tbody>
</table>

Numerical classification

- Hierarchical clustering
- Exclusive clustering
# Individual Heuristics: Examples

## Multi-Word Heuristic (MWH)

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3.4.2 Agenteak</strong></td>
<td><strong>3.4.2 Agents</strong></td>
</tr>
<tr>
<td><strong>3.4.2.1 Agente mugikorrak</strong></td>
<td><strong>3.4.2.1 Mobile agents</strong></td>
</tr>
</tbody>
</table>

## Acronyms Heuristic (AH)

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>4.2.2 Interfaze grafikoak sortzeko lengoaiak</strong></td>
<td><strong>4.2.2 Languages for building graphical interfaces</strong></td>
</tr>
<tr>
<td><strong>4.2.2.1 XUL</strong></td>
<td><strong>4.2.2.1 XUL</strong></td>
</tr>
</tbody>
</table>
### Multi-Word Heuristic (MWH)

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2 Agenteak</td>
<td>3.4.2 Agents</td>
</tr>
<tr>
<td>3.4.2.1 Agenteak mugikorrak</td>
<td>3.4.2.1 Mobile agents</td>
</tr>
</tbody>
</table>

### Acronyms Heuristic (AH)

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2.2 Interfaze grafikoak sortzeko lengoaiak</td>
<td>4.2.2 Languages for building graphical interfaces</td>
</tr>
<tr>
<td>4.2.2.1 XUL</td>
<td>4.2.2.1 XUL</td>
</tr>
</tbody>
</table>
### Individual Heuristics: Examples

#### Multi-Word Heuristic (MWH)

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2 Agenteak</td>
<td>3.4.2 Agents</td>
</tr>
<tr>
<td>3.4.2.1 Agente mugikorrak</td>
<td>3.4.2.1 Mobile agents</td>
</tr>
</tbody>
</table>

#### Acronyms Heuristic (AH)

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2 Languages for building graphical interfaces</td>
<td>4.2 Languages for building graphical interfaces</td>
</tr>
<tr>
<td>isA.</td>
<td>isA. isA.</td>
</tr>
<tr>
<td>XUL</td>
<td>XUL</td>
</tr>
<tr>
<td>4.2.2.1 XUL</td>
<td>4.2.2.1 XUL</td>
</tr>
</tbody>
</table>
Identifying the Structural Relationships

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Palm OS</td>
<td>6.1 Palm OS</td>
</tr>
<tr>
<td>6.2 Windows CE</td>
<td>6.2 Windows CE</td>
</tr>
<tr>
<td>6.3 Linux Familiar 0.5 Distribuzioa</td>
<td>6.3 Linux Familiar 0.5 Distribution</td>
</tr>
</tbody>
</table>

**Palm OS**
- Category: Concept
- Difficulty: High
- Relevance: High

**Windows CE**
- Category: Concept
- Difficulty: High
- Relevance: High

**Linux Familiar 0.5 Distribution**
- Category: Concept
- Difficulty: High
- Relevance: High
Identifying the Structural Relationships

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laguntzaile Pertsonal Digitalentzako Sistema Eragileak</td>
<td>Operative Systems for Personal Digital Assistants</td>
</tr>
<tr>
<td>6.1 Palm OS</td>
<td>6.1 Palm OS</td>
</tr>
<tr>
<td>6.2 Windows CE</td>
<td>6.2 Windows CE</td>
</tr>
<tr>
<td>6.3 Linux Familiar 0.5 Distribuzioa</td>
<td>6.3 Linux Familiar 0.5 Distribution</td>
</tr>
</tbody>
</table>

isA
Confidence: 1
Inferred by: heuristic
Heuristic: EH

Operative Systems for Personal Digital Assistants
Category: Concept
Difficulty: High
Relevance: High

Palm OS
Category: Concept
Difficulty: High
Relevance: High

Windows CE
Category: Concept
Difficulty: High
Relevance: High

Linux Familiar 0.5 Distribution
Category: Concept
Difficulty: High
Relevance: High
Identifying the Structural Relationships

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1 Palm OS</td>
<td>6.1 Palm OS</td>
</tr>
<tr>
<td>6.2 Windows CE</td>
<td>6.2 Windows CE</td>
</tr>
<tr>
<td>6.3 Linux Familiar 0.5 Distribuzioa</td>
<td>6.3 Linux Familiar 0.5 Distribution</td>
</tr>
</tbody>
</table>

**Palm OS**
- Category: Concept
- Difficulty: High
- Relevance: High

**Windows CE**
- Category: Concept
- Difficulty: High
- Relevance: High

**Linux Familiar 0.5 Distribution**
- Category: Concept
- Difficulty: High
- Relevance: High

**isA**
- Confidence: 0.90
- Inferred by: combined
- Heuristic: str-default
## Identifying the Sequential Relationships

### Sequential Relationships

- **prerequisite**: states that a topic **must** be mastered before addressing the other one
  - Between an item and any previous item
- **next**: states that a topic is **recommended** to be learned after mastering the other one
  - Default between two consecutive outline items at the same nesting level
- 6 heuristics for **prerequisite** (e.g., RH)
Reference Heuristic (RH)

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Unibertsoa</td>
<td>2. Universe</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>7. Unibertsoa ezagutzeko teknologia</td>
<td>7. Technology to explore the Universe</td>
</tr>
</tbody>
</table>
# Heuristics for the Sequential Relationships: Example

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. <strong>Unibertsoa</strong>...</td>
<td>2. <strong>Universe</strong>...</td>
</tr>
<tr>
<td>7. <strong>Unibertsoa</strong> ezagutzeko teknologia</td>
<td>7. Technology to explore the <strong>Universe</strong></td>
</tr>
</tbody>
</table>
### Identifying the Sequential Relationships

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Unibertsoa</td>
<td>2. Universe</td>
</tr>
<tr>
<td><img src="image.png" alt="Diagram" /></td>
<td><img src="image.png" alt="Diagram" /></td>
</tr>
<tr>
<td><img src="image.png" alt="Technology to explore the universe" /></td>
<td><img src="image.png" alt="Galaxies" /></td>
</tr>
<tr>
<td>Category: Concept</td>
<td>Category: Concept</td>
</tr>
<tr>
<td>Difficulty: High</td>
<td>Difficulty: High</td>
</tr>
<tr>
<td>Relevance: High</td>
<td>Relevance: High</td>
</tr>
</tbody>
</table>

- **Universe**
  - Category: Concept
  - Difficulty: High
  - Relevance: High

- **Galaxies**
  - Category: Concept
  - Difficulty: High
  - Relevance: High
Identifying the Sequential Relationships

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Unibertsoa</td>
<td>2. Universe</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>7. Unibertsoa ezagutzeko teknologia</td>
<td>7. Technology to explore the universe</td>
</tr>
</tbody>
</table>

**Universe**
- Category: Concept
- Difficulty: High
- Relevance: High

**Galaxies**
- Category: Concept
- Difficulty: High
- Relevance: High

Inferred by: heuristic
Heuristic: seq-default

Confidence: 0.76
Identifying the Sequential Relationships

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Unibertsoa</td>
<td>2. Universe</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>7. Unibertsoa ezagutzeko teknologia</td>
<td>7. Technology to explore the universe</td>
</tr>
</tbody>
</table>

**Prerequisite**
- **Confidence**: 0.93
- **Inferred by**: heuristic
- **Heuristic**: RH

**Next**
- **Confidence**: 0.76
- **Inferred by**: heuristic
- **Heuristic**: seq-default

**Universe**
- Category: Concept
- Difficulty: High
- Relevance: High

**Galaxies**
- Category: Concept
- Difficulty: High
- Relevance: High

**Technology to explore the universe**
- Category: Concept
- Difficulty: High
- Relevance: High
Gathering the Learning Domain Ontology

- Document Outline Internal Representation
- Outline Analysis
- Learning Domain Ontology
- Whole Document Analysis
- Enriched Learning Domain Ontology
Enhance the Learning Domain Ontology with new topics and relationships

- **New topics**: Erauzterm (Gurrutxaga et al., 2005)
- **New relationships**: Patterns
  - MWH: Eklipse – Eguzki-eklipse (Eclipse – Sun-eclipse)
  - Grammar for Pedagogical Relationships

<table>
<thead>
<tr>
<th></th>
<th>isA</th>
<th>partOf</th>
<th>prerequisite</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of patterns</strong></td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
### Example of a pattern for `isA` relationships

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pattern</strong></td>
<td>@Topic referred to as @Topic</td>
</tr>
<tr>
<td><strong>Example</strong></td>
<td>Esne bidea izeneko galaxiak 100 mila milioi izar dituela uste dute zientzalariek.</td>
</tr>
<tr>
<td></td>
<td>Scientists believe that the galaxy referred to as the Milky Way has over 100 billion stars.</td>
</tr>
</tbody>
</table>
### Example of a pattern for isA relationships

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>@Topic ize(\text{n})ko  @Topic</td>
<td>@Topic referred to as @Topic</td>
<td></td>
</tr>
<tr>
<td>Example</td>
<td>Esne bidea ize(\text{n})ko galaxiak 100 mila milioi izar dituela uste dute zientzalariek.</td>
<td>Scientists believe that the galaxy referred to as the Milky Way has over 100 billion stars.</td>
</tr>
</tbody>
</table>
General Process for the Acquisition of the Domain Module

1. **Electronic Textbook**
2. **Preprocess**
   - Document Outline Internal Representation
   - Document Internal Representation
3. **LDO Gathering**
   - Learning Domain Ontology
4. **LOs Gathering**
   - Domain Module
Gathering the Learning Objects

**Procedure**

- Ontology-driven process
- Identify the fragments of the textbook with educational purpose related to the domain topics
- Build Learning Objects from the identified didactic resources
- Store the Learning Objects in the Learning Object Repository
Gathering the Learning Objects

- Discourse Markers
- DR Grammar
- Document Internal Representation
- Generation of DRs
- Learning Domain Ontology
- Didactic Ontology
- Didactic Resources
Generation of the Didactic Resources

3.1.1 Labeled Document Internal Representation

Learning Domain Ontology

Dom-Sortze
Generation of the Didactic Resources

- Acquisition of the Domain Module
- Dom-Sortze
- Evaluation
- Final Remarks

- Generation of the Didactic Resources
  - Document Internal Representation
  - Labeling Domain Topics
    - Labeled Document Internal Representation
    - Identification of DRs
      - Atomic Didactic Resources
      - 3.1.2 DR Grammar
DR Grammar (I)

DR Grammar

- Syntactic structures used for definition, problem-statements ...

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patterns</td>
<td>21</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>18</td>
</tr>
</tbody>
</table>

Definition

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pattern</td>
<td>@Topic definition (DAT)</td>
<td>definition to be called @Topic</td>
</tr>
<tr>
<td>@Topic deitu</td>
<td>Unibertsoa astro guztien</td>
<td>The whole set of celestial bodies and the</td>
</tr>
<tr>
<td></td>
<td>multzoari eta betetzen</td>
<td>space they fill is called Universe.</td>
</tr>
<tr>
<td></td>
<td>duten espazioari deitzen</td>
<td></td>
</tr>
<tr>
<td></td>
<td>zaio.</td>
<td></td>
</tr>
</tbody>
</table>
**Problem-statement**

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Idatzi @Topic</strong> (DAT) <strong>buruz jakin</strong></td>
<td>Write know about @Topic</td>
<td></td>
</tr>
<tr>
<td><strong>Example</strong></td>
<td><strong>Idatzi</strong> <strong>eklipseei buruz dakizun guztia.</strong></td>
<td>Write all you know about eclipses.</td>
</tr>
</tbody>
</table>
Building the Atomic Didactic Resources

**Procedure**

- Identify the sentence containing part of the didactic resource
- Add the following sentences, if they are content-related
- Unless the DR Grammar has identified another didactic resource in the sentence
Generation of the Didactic Resources

- Acquisition of the Domain Module
- Dom-Sortze
- Evaluation
- Final Remarks

1. Learning Domain Ontology
2. Didactic Ontology
3. Labeled Document Internal Representation
4. Document Internal Representation
5. Labeling Domain Topics
6. Identification of DRs
7. Didactic Resources
8. Atomic Didactic Resources
9. Enhancements of DRs
10. Discourse Markers

3.1.3 Didactic Resources
Enhancing the Didactic Resources

- Consecutive didactic resources are combined if they are similar
**Enhancing the Didactic Resources: Similarity**

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planetak berezko argirik ez duten gorputzak dira, eta izar baten inguruan biraka mugitzen dira. Uste denez, Eguzki-Sistemako planetak Eguzkiarekin batera eratu ziren, eta pentsa daiteke antzeko planeta ugari izango direla beste izar batzuen inguruan.</td>
<td>Planets are non-self-luminous celestial bodies orbiting around a star. Solar system planets are thought to be formed together with the Sun, and it is also conceived that many similar planets might be found orbiting other stars.</td>
</tr>
<tr>
<td>Lurra planeta bat da.</td>
<td>The Earth is a planet.</td>
</tr>
</tbody>
</table>
## Enhancing the Didactic Resources: Similarity

<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Planetak</strong> berezeko argirik ez duten gorputzak dira, eta izar baten inguruan biraka mugitzen dira. Uste denez, Eguzki-Sistemako planetak Eguzkiarekin batera eratu ziren, eta pentsa daiteke antzeko planeta ugari izango direla beste izar batzuen inguruan.</td>
<td><strong>Planets</strong> are non-self-luminous celestial bodies orbiting around a star. Solar system planets are thought to be formed together with the Sun, and it is also conceived that many similar planets might be found orbiting other stars.</td>
</tr>
<tr>
<td><strong>Lurra planeta bat da.</strong></td>
<td>The <strong>Earth</strong> is a <strong>planet</strong>.</td>
</tr>
</tbody>
</table>
Resemblance of the Didactic Resources

Similarity Measuring Approaches

- Different methods were tested
- Ontology-based methods performed best (Larrañaga, Elorriaga and Arruarte, 2008)

DRs are combined considering

- Content similarity
  - Learning Domain Ontology
- Compatibility of the kind of DRs
  - Didactic Ontology (Leidig, 2001)
# Resemblance of the Didactic Resources

## Similarity Measuring Approaches

- Different methods were tested
- Ontology-based methods performed best (Larrañaga, Elorriaga and Arruarte, 2008)

## DRs are combined considering

- Content similarity
  - Learning Domain Ontology
- Compatibility of the kind of DRs
  - Didactic Ontology (Leidig, 2001)

## UKB Algorithm

(Agirre, López de Lacalle and Soroa, 2009)
Enhancing the Didactic Resources

- Consecutive didactic resources are combined if they are similar.
- Didactic resources might be enhanced with preceding text to keep the cohesion.
  - Discourse markers: words or expressions used to connect text fragments.
### Basque

<table>
<thead>
<tr>
<th>Online</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>The planets of the Solar System are classified in two groups: 1) Inner planets, those that are closest to the sun, Mercury, Venus and Earth. 2) Outer planets, the other planets. Sometimes, Mars is also included in the first group.</td>
</tr>
<tr>
<td>...</td>
<td>In fact, besides being further from the Sun than the Earth, like Mercury, Venus and Earth, Mars has a relatively small size and high density, i.e., it is heavy for its size.</td>
</tr>
</tbody>
</table>


Izan ere, Eguzkitik Lurra baina urrunago egon arren, Merkurio, Artizarra eta Lurra bezalaxe, dimentsio txiki samarrekoa da Marte, eta dentsitate handia du, hau da, duen tamainarako astuna da.
<table>
<thead>
<tr>
<th>Basque</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eguzki-sistema planetak bi taldetan banatu ohi dira: 1) Barneko planetak, eguzkitik hurbilen daudenak, Merkurio, Artizarra eta Lurra. 2) Kanpoko planetak, gainerako guztiak. Batzuetan, Marte ere lehenengo taldearen barruan sartzen da.</td>
<td>The planets of the Solar System are classified in two groups: 1) Inner planets, those that are closest to the sun, Mercury, Venus and Earth. 2) Outer planets, the other planets. Sometimes, Mars is also included in the first group.</td>
</tr>
<tr>
<td><strong>Izan ere</strong>, Eguzkitik Lurra baina urrunago egon arren, Merkurio, Artizarra eta Lurra bezalaxe, dimentsio txiki samarrekoa da Marte, eta dentsitate handia du, hau da, duen tamainarako astuna da.</td>
<td><strong>In fact</strong>, besides being further from the Sun than the Earth, like Mercury, Venus and Earth, Mars has a relatively small size and high density, i.e., it is heavy for its size.</td>
</tr>
</tbody>
</table>
Gathering the Learning Objects

- Discourse Markers
- DR Grammar
- Document Internal Representation
- Generation of DRs
- Learning Domain Ontology
- Didactic Resources
- Didactic Ontology
- ALOCOM Ontology
- Learning Objects

Building the LOs

3.2
Building the Learning Objects

Representation of the Learning Objects
- XML-based formalism, based on ALOCOM (Verbert, 2008)

Automatic Metadata Annotation
- Initial metadata
  - SAmgl (Meire, Ochoa & Duval, 2007)
- Enriching the metadata
  - **Keywords**
    - *Earth, Mars, Mercury, Venus → Planet*
  - **Kind**: Patterns used to identify the didactic resource
    - ALOCOM ontology used as vocabulary
Gathering the Learning Objects

1. Discourse Markers
2. DR Grammar
3. Document Internal Representation
4. Generation of DRs
5. Didactic Ontology
6. Learning Domain Ontology
7. Didactic Resources
8. Building the LOs
9. ALOCOM Ontology
10. Storing the LOs

- LOR
- Learning Objects
Storing the Learning Objects

LO Storage

- Learning Objects are stored in the LOR
- In the case of aggregated Learning Objects, their components are also stored
- The Learning Object Repository is based on the ARIADNE Knowledge Pool System (Ternier et al., 2009)
Outline

1 Introduction

2 Acquisition of the Domain Module

3 Dom-Sortze: a Framework for the Semi-automatic construction of the Domain Module

4 Evaluation

5 Final Remarks: Conclusions and Future Lines
DOM-Sortze

- Framework for the semi-automatic construction of the Domain Module
- Web-service based architecture
- Applications for the Domain Module generating tasks
  - Preprocessor – preparing the textbook for the knowledge acquisition tasks
  - LDO Builder – gathering the Learning Domain Ontology
  - ErauzOnt – gathering the Learning Objects
- Application for the collaborative supervision of the Domain Module
  - ElkarDOM
DOM-Sortze: Architecture

- Elkar-DOM
- ErauzOnt
- Preprocessor
- LDO Builder
DOM-Sortze: Architecture

Diagram of the architecture:

- **Elkar-DOM**
- **ErauzOnt**
- **LD Service**
- **Preprocessor**
- **LDO Builder**
- **Pdf2XML Service**
- **DIR Builder**
- **NLP Analysis Service**
DOM-Sortze: Architecture
DOM-Sortze: Architecture

- LO Preview Repository
- LO Repository
- Lucene Index
- COI Service
- RD Service
- Elkar-DOM
- ErauzOnt
- LD Service
- HC Service
- CG Service
- Preprocessor
- LDO Builder
- Pdf2XML Service
- DIR Builder
- NLP Analysis Service
- UKB Service
- SAmgl Service
DOM-Sortze: Architecture
DOM-Sortze: Architecture

- LO Preview Repository
- LO Repository
- Lucene Index
- Domain Module Repository
- SQI Service
- COI Service
- RD Service
- Elkar-DOM
- ErauzOnt
- LD Service
- HC Service
- CG Service
- Preprocessor
- LDO Builder
- Pdf2XML Service
- DIR Builder
- NLP Analysis Service
- UKB Service
- SAmgl Service
Elkar-DOM: a Tool for the Collaborative Supervision of the Domain Module

**Elkar-DOM**
- Concept map-based tool
- Collaborative supervision of the Domain Module

**Learning Domain Ontology**
- Topics and Relationships are gathered automatically
- Domain Module authors can supervise the results
  - Correct errors
  - Adapt it to their preferences

**Supervise the LOs for each topic**
- Select the most appropriate LOs
Supervision of the Learning Domain Ontology

Supervise the relationship. It adjusts the confidence of the heuristic.
Supervision of the Learning Objects (I)
Supervision of the Learning Objects (I)
Supervision of the Learning Objects (I)
Supervision of the Learning Objects (I)

List of selected Learning Objects

Manage the Learning Objects: remove, look for Learning Objects
Supervision of the Learning Objects (II)

Browse the Learning Objects that match the query and select the most appropriate Learning Objects for the topic.
Supervision of the Learning Objects (II)

Browse the Learning Objects that match the query and select the most appropriate Learning Objects for the topic.
Supervision of the Learning Objects (II)

Browse the Learning Objects that match the query and select the most appropriate Learning Objects for the topic.
Outline

1. Introduction
2. Acquisition of the Domain Module
4. Evaluation
5. Final Remarks: Conclusions and Future Lines
Evaluation

Methodology

- **Gold standard**
  - Instructional designers in cooperation with teachers or textbook authors defined the reference output

- **Measures**
  - **Recall**: proportion of relevant elements that have been extracted and correctly classified
  - **Precision**: proportion of extracted elements that are relevant and correctly classified
  - **F-measure**: harmonic mean of precision and recall

- **Incremental approach**
  - Evaluation of the acquisition of the Learning Domain Ontology from Outlines
  - Evaluation of the acquisition of the Learning Objects
  - Evaluation of the whole process
Evaluation of the Acquisition of the Learning Domain Ontology from Outlines (I)

Test Set
- Tested on 150 outlines from different subjects taught at the University of the Basque Country
- Social Sciences, Economy, and Engineering

Gold Standard
Only the topics referred in the outlines and the pedagogical relationships among those topics were defined in the reference Learning Domain Ontologies
Evaluation of the Acquisition of the Learning Domain Ontology from Outlines (II)

Statistics on the Acquisition of the LDO

<table>
<thead>
<tr>
<th></th>
<th>PartOf</th>
<th>IsA</th>
<th>Next</th>
<th>Prerequisite</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall (%)</td>
<td>99.15</td>
<td>79.38</td>
<td>99.86</td>
<td>92.95</td>
<td>98.15</td>
</tr>
<tr>
<td>Precision (%)</td>
<td>98.19</td>
<td>89.27</td>
<td>99.61</td>
<td>97.75</td>
<td>98.36</td>
</tr>
<tr>
<td>F-measure (%)</td>
<td>98.67</td>
<td>84.04</td>
<td>99.74</td>
<td>95.29</td>
<td>98.26</td>
</tr>
</tbody>
</table>
Evaluation of the Acquisition of the Learning Domain Ontology from Outlines (II)

Statistics on the Acquisition of the LDO

<table>
<thead>
<tr>
<th></th>
<th>PartOf</th>
<th>IsA</th>
<th>Next</th>
<th>Prerequisite</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall (%)</td>
<td>99.15</td>
<td>79.38</td>
<td>99.86</td>
<td>92.95</td>
<td>98.15</td>
</tr>
<tr>
<td>Precision (%)</td>
<td>98.19</td>
<td>89.27</td>
<td>99.61</td>
<td>97.75</td>
<td>98.36</td>
</tr>
<tr>
<td>F-measure (%)</td>
<td>98.67</td>
<td>84.04</td>
<td>99.74</td>
<td>95.29</td>
<td>98.26</td>
</tr>
</tbody>
</table>
Evaluation of the Acquisition of the Learning Domain Ontology from Outlines (III)
### Evaluation of the Acquisition of the Learning Domain Ontology from Outlines (IV)

#### Per Domain

<table>
<thead>
<tr>
<th></th>
<th>Structural</th>
<th></th>
<th>Sequential</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>partOf</td>
<td>isA</td>
<td>next</td>
<td>prereq.</td>
<td></td>
</tr>
<tr>
<td>Eng.</td>
<td>Recall (%)</td>
<td>93.63</td>
<td>59.09</td>
<td>99.66</td>
<td>95.05</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>97.80</td>
<td>89.66</td>
<td>99.74</td>
<td>96.35</td>
</tr>
<tr>
<td></td>
<td>F-measure (%)</td>
<td>98.71</td>
<td>71.23</td>
<td>99.70</td>
<td>95.69</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>98.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eco.</td>
<td>Recall (%)</td>
<td>98.66</td>
<td>84.76</td>
<td>100.00</td>
<td>90.52</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>98.55</td>
<td>85.80</td>
<td>99.50</td>
<td>98.96</td>
</tr>
<tr>
<td></td>
<td>F-measure (%)</td>
<td>98.60</td>
<td>85.28</td>
<td>99.75</td>
<td>94.55</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>98.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soc.</td>
<td>Recall (%)</td>
<td>99.42</td>
<td>91.78</td>
<td>100.00</td>
<td>94.29</td>
</tr>
<tr>
<td></td>
<td>Precision (%)</td>
<td>98.28</td>
<td>97.10</td>
<td>99.64</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>F-measure (%)</td>
<td>98.84</td>
<td>94.37</td>
<td>99.82</td>
<td>97.06</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>98.62</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Evaluation of the Acquisition of the Learning Domain Ontology from Outlines (IV)

#### Per Domain

<table>
<thead>
<tr>
<th>Domain</th>
<th>Structural</th>
<th>Sequential</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>partOf</td>
<td>isA</td>
<td>next</td>
</tr>
<tr>
<td>Eng.</td>
<td>Recall (%)</td>
<td>Precision (%)</td>
<td>F-measure (%)</td>
</tr>
<tr>
<td></td>
<td>93.63</td>
<td>97.80</td>
<td>98.71</td>
</tr>
<tr>
<td>Eco.</td>
<td>Recall (%)</td>
<td>Precision (%)</td>
<td>F-measure (%)</td>
</tr>
<tr>
<td></td>
<td>98.66</td>
<td>98.55</td>
<td>98.60</td>
</tr>
<tr>
<td>Soc.</td>
<td>Recall (%)</td>
<td>Precision (%)</td>
<td>F-measure (%)</td>
</tr>
<tr>
<td></td>
<td>99.42</td>
<td>98.28</td>
<td>98.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Evaluation of the Acquisition of the Domain Module: Dom-Sortze

### Final Remarks
Evaluation of the Acquisition of LOs (I)

Test set
- Tested on 4 textbooks provided by Euskal Herriko Ikastola
  - 5,100 – 7,900 words
- Nature Sciences: Geology and Biology
- 1st course, secondary education

Gold Standard

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>67</td>
<td>12</td>
<td>8</td>
<td>105</td>
<td>37</td>
<td>229</td>
</tr>
</tbody>
</table>

Evaluation Target
- Performance of the DR Grammar
- Acquisition of the Learning Objects
Evaluation of the Acquisition of LOs (II)

### Accuracy of the DR Grammar

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rule Firing</strong></td>
<td>103</td>
<td>22</td>
<td>35</td>
<td>1</td>
<td>342</td>
<td>504</td>
</tr>
<tr>
<td><strong>Correct</strong></td>
<td>43</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>308</td>
<td>371</td>
</tr>
<tr>
<td><strong>Accuracy (%)</strong></td>
<td>41.75</td>
<td>45.45</td>
<td>28.57</td>
<td>0.00</td>
<td>90.06</td>
<td>73.61</td>
</tr>
</tbody>
</table>
### Accuracy of the DR Grammar

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule Firing</td>
<td>103</td>
<td>22</td>
<td>35</td>
<td>1</td>
<td>342</td>
<td>504</td>
</tr>
<tr>
<td>Correct</td>
<td>43</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>308</td>
<td>371</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>41.75</td>
<td>45.45</td>
<td>28.57</td>
<td>0.00</td>
<td><strong>90.06</strong></td>
<td>73.61</td>
</tr>
</tbody>
</table>
## Evaluation of the Acquisition of LOs (III)

### Acquisition of the LOs

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recall (%)</strong></td>
<td>59.70</td>
<td>50.00</td>
<td>87.50</td>
<td>81.90</td>
<td>59.46</td>
<td>70.31</td>
</tr>
<tr>
<td><strong>Precision (%)</strong></td>
<td>91.14</td>
<td>96.30</td>
<td>100.00</td>
<td>88.55</td>
<td>97.84</td>
<td>91.88</td>
</tr>
<tr>
<td><strong>F-measure (%)</strong></td>
<td>72.14</td>
<td>65.82</td>
<td>93.33</td>
<td>85.10</td>
<td>73.97</td>
<td>79.66</td>
</tr>
</tbody>
</table>
Evaluation of the Whole Process (I)

Test Set
- Tested on a 8,495 words textbook provided by the Euskal Herriko Ikastola
- Nature Sciences: Astronomy
- 1st course, secondary education

Evaluation Target
- Acquisition of the Learning Domain Ontology
- Acquisition of the Learning Objects
Gold Standard

Learning Domain Ontology

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topics</td>
<td>38</td>
<td>45</td>
<td>83</td>
</tr>
<tr>
<td>Relations</td>
<td>108</td>
<td>27</td>
<td>135</td>
</tr>
<tr>
<td>partOf</td>
<td>36</td>
<td>7</td>
<td>43</td>
</tr>
<tr>
<td>isA</td>
<td>51</td>
<td>15</td>
<td>66</td>
</tr>
<tr>
<td>prerequisite</td>
<td>13</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>next</td>
<td>8</td>
<td>5</td>
<td>13</td>
</tr>
</tbody>
</table>

Learning Objects

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>27</td>
<td>6</td>
<td>21</td>
<td>6</td>
<td>6</td>
<td>66</td>
</tr>
</tbody>
</table>
## Evaluation of the Whole Process (III)

### Acquisition of the Topics

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recall (%)</strong></td>
<td>Outline</td>
<td>28.95</td>
<td>13.33</td>
</tr>
<tr>
<td></td>
<td>Whole Document</td>
<td>63.16</td>
<td>71.11</td>
</tr>
<tr>
<td><strong>Precision (%)</strong></td>
<td>Outline</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>Whole Document</td>
<td>100.00</td>
<td>14.18</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>92.11</td>
<td>84.44</td>
</tr>
</tbody>
</table>

### Acquisition of the Pedagogical Relationships

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recall (%)</strong></td>
<td>20.00</td>
<td>20.74</td>
<td>40.74</td>
</tr>
<tr>
<td><strong>Precision (%)</strong></td>
<td>80.10</td>
<td>63.27</td>
<td>72.50</td>
</tr>
<tr>
<td><strong>F-measure (%)</strong></td>
<td>32.53</td>
<td>31.24</td>
<td>52.17</td>
</tr>
</tbody>
</table>
### Evaluation of the Whole Process (III)

#### Acquisition of the Topics

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recall (%)</strong></td>
<td>Outline</td>
<td>28.95</td>
<td>13.33</td>
</tr>
<tr>
<td></td>
<td>Whole Document</td>
<td>63.16</td>
<td>71.11</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>92.11</td>
<td>84.44</td>
</tr>
<tr>
<td><strong>Precision (%)</strong></td>
<td>Outline</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td></td>
<td>Whole Document</td>
<td>100.00</td>
<td>14.18</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td>100.00</td>
<td>14.60</td>
</tr>
</tbody>
</table>

#### Acquisition of the Pedagogical Relationships

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recall (%)</strong></td>
<td>20.00</td>
<td>20.74</td>
<td>40.74</td>
</tr>
<tr>
<td><strong>Precision (%)</strong></td>
<td>80.10</td>
<td>63.27</td>
<td>72.50</td>
</tr>
<tr>
<td><strong>F-measure (%)</strong></td>
<td>32.53</td>
<td>31.24</td>
<td>52.17</td>
</tr>
</tbody>
</table>
### Evaluation of the Whole Process (III)

#### Acquisition of the Topics

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recall (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outline</td>
<td>28.95</td>
<td>13.33</td>
<td>20.48</td>
</tr>
<tr>
<td>Whole Document</td>
<td>63.16</td>
<td>71.11</td>
<td>67.47</td>
</tr>
<tr>
<td>Total</td>
<td>92.11</td>
<td>84.44</td>
<td>87.95</td>
</tr>
<tr>
<td><strong>Precision (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outline</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td>Whole Document</td>
<td>100.00</td>
<td>14.18</td>
<td>16.36</td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>14.60</td>
<td>17.48</td>
</tr>
</tbody>
</table>

#### Acquisition of the Pedagogical Relationships

<table>
<thead>
<tr>
<th></th>
<th>Level 1</th>
<th>Level 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recall (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20.00</td>
<td>20.74</td>
<td>40.74</td>
<td></td>
</tr>
<tr>
<td><strong>Precision (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80.10</td>
<td>63.27</td>
<td>72.50</td>
<td></td>
</tr>
<tr>
<td><strong>F-measure (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32.53</td>
<td>31.24</td>
<td>52.17</td>
<td></td>
</tr>
</tbody>
</table>
## Evaluation of the Whole Process (IV)

### Acquisition of the Learning Objects

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Recall (%)</td>
<td>74.07</td>
<td>33.00</td>
<td>95.24</td>
<td>50.00</td>
<td>100.00</td>
<td>77.27</td>
</tr>
<tr>
<td>Precision (%)</td>
<td>78.79</td>
<td>100.00</td>
<td>93.33</td>
<td>71.43</td>
<td>86.21</td>
<td>84.50</td>
</tr>
<tr>
<td>F-measure (%)</td>
<td>76.36</td>
<td>50.00</td>
<td>94.28</td>
<td>58.82</td>
<td>92.59</td>
<td>80.73</td>
</tr>
</tbody>
</table>
Evaluation of the Whole Process (V)

Some Reflections

- Acquisition of the Learning Domain Ontology
  - 87.95% recall on the identification of topics
  - 40.74% recall on the identification of pedagogical relationships
    - Outline was too simple
    - Considering the precision of the heuristics for the outlines, better results could be obtained if the outline was richer

- Acquisition of the Learning Objects
  - Similar results to those obtained in the previous evaluation
Goal Achievement (I)

Semi-automatic Generation of the Domain Module
- Automatic knowledge elicitation + collaborative supervision
- Preprocess
- Acquisition of the Learning Domain Ontology
- Acquisition of the Learning Objects

Collaborative Supervision
- Elkar-DOM

Knowledge Reuse
- Ontologies + Learning Objects
## Goal Achievement (II)

### Domain-independence
- No implicit domain-specific knowledge used
- Tested on diverse domains – similar results

### Language-independence
- The process is not language-dependent
- Language-dependent resources can be easily adapted
  - Heuristics
  - Grammars
  - Discourse markers
- First experiment for English (Conde et al., 2012)
  - *Principles of Object-Oriented Programming*, 29,300 words
Multilingual generation of Domain Modules
- Ontologies support multilingualism
- Machine translation ⇒ *draft* versions the Learning Objects

Machine Learning

Versioning and maintenance

Integration of DOM-Sortze with other applications
- Translators for the Domain Module
- Arikiturri (Aldabe, 2011)

Enhance collaboration
Semi-automatic Generation of Learning Domain Modules for Technology Supported Learning Systems

Miguel Larrañaga Olagaray

Supervisors: Ana Arruarte Lasa and Jon A. Elorriaga Arandia

December 2012