CWL

CWL Conversion Framework

(Version 0.1 beta)

31 Oct. 2007

Vahan Avetisyan
Institute for Informatics and Automation Problems of NAS RA

1. Introduction

In this document the draft version of CWL Conversion Framework is illustrated. This Framework serves as a tool for conversions among CWL.unl, CWL.cdl and CWL.rdf. It is implemented as a web application written using Java/JSP, DHTML/Ajax and VML.

2. Application Flow

The basic flow of the conversion considers:
1. Parsing of input document
2. Building of object graph, where nodes correspond to entities and arches correspond to relations/predicates.
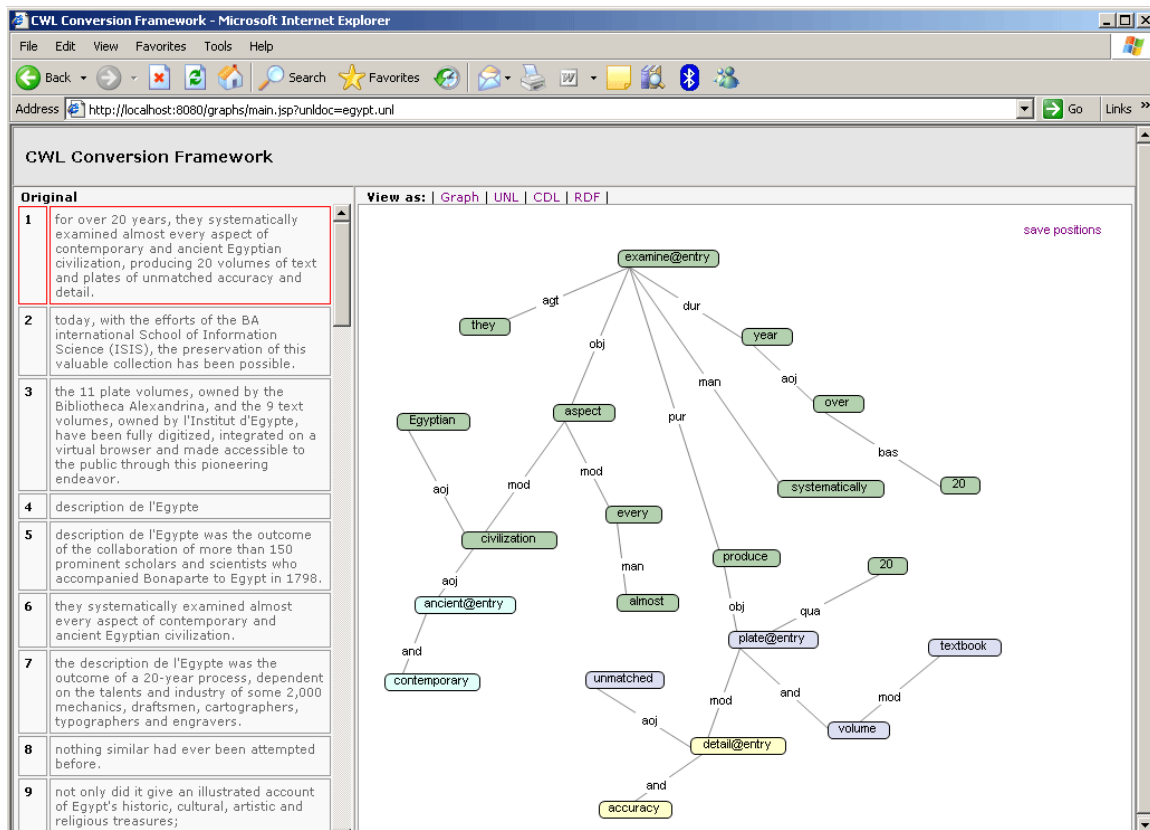3. Generation of various views (UNL, CDL, RDF, Graphical)

All basic functionality is implemented as a Java library which can be reused for various applications.

3. View Rendering

Object graphs can be rendered in different ways but in general they reflect the same relational network from the input document.

Graph View

The default is a graphical view where the sentence from natural language is presented as a graph (*see Picture 1*). Nodes of the graph contain Universal Words from UNL KB corresponding to the entities building the natural language sentence.

Picture 1: Graph View

Arches of the graph correspond to the relations among the entities. The direction of the relation can be seen by clicking on any node. In that case you will see incoming relations colored as red lines and outgoing relations will remain black. Full versions of Universal Words can be seen by double-clicking on the nodes, otherwise only headwords are shown by default. The rendering is done using VML (Vector Markup Language). Nodes of the graph are dragable, thus each user can customize the look for his convenience. After dragging the nodes, the overall positions of the nodes can be saved by clicking corresponding link in upper right corner of rendering area.

Sub sentences of the sentence are rendered using different colors for nodes. All sub sentences have root entry marked with "@entry" string after headword. Same refers to the root entry of the sentence itself.

UNL View

The UNL View renders the graph as a sequence of UNL relations (see Picture 2). The code is highlighted accordingly. The sequence of the relations considers main sentence relations first and sub sentence relations afterwards.
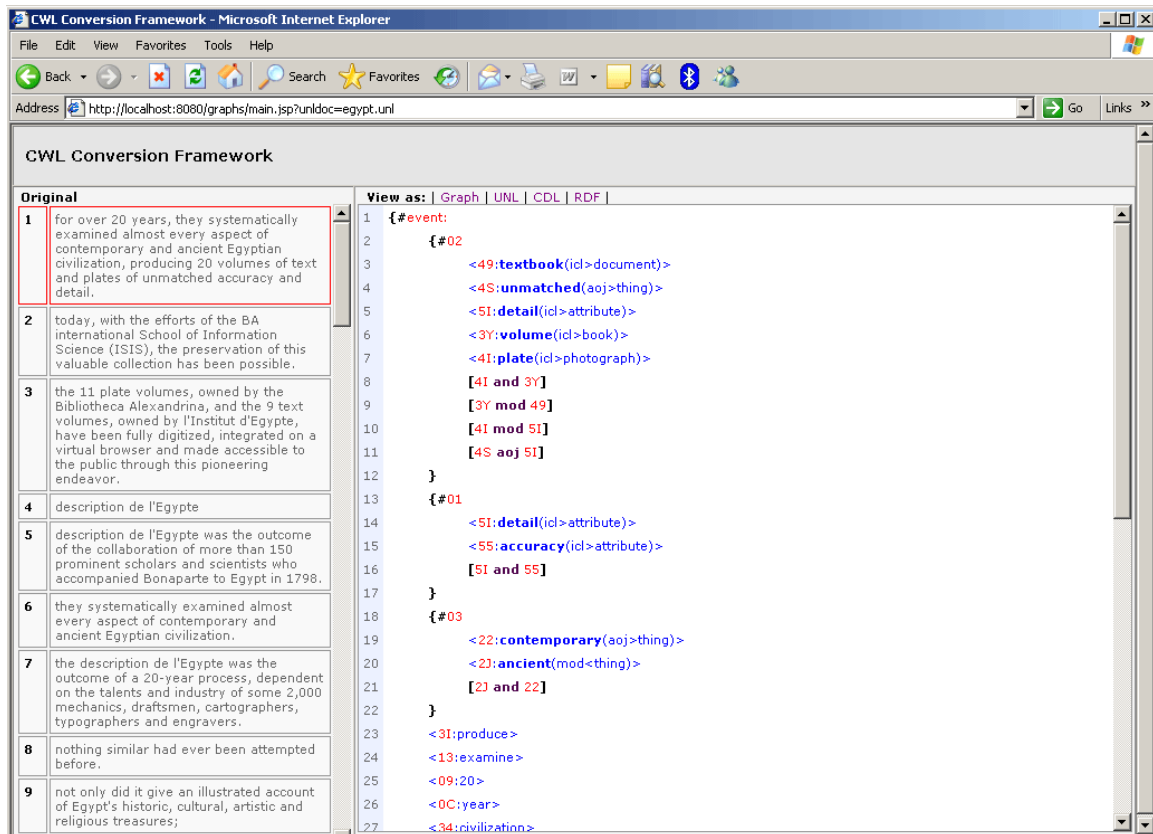

Picture 2: UNL View

CDL View

The CDL View renders the graph as sequences of entries and relations (see Picture 3).
At first the entity sequences appear where UWs from UNL KB are being used. After which there relations among listed entities are rendered. Similar structure is being used for sub sentences. The code is highlighted accordingly.

CWL Conversion Framework - Microsoft Internet Explorer

File   Edit   View   Favorites   Tools   Help

Back | Search   Favorites

Address http://localhost:8080/graphs/main.jsp?unldoc=egypt.unl   Go   Links »

CWL Conversion Framework

Original

1	for over 20 years, they systematically examined almost every aspect of contemporary and ancient Egyptian civilization, producing 20 volumes of text and plates of unmatched accuracy and detail.
2	today, with the efforts of the BA international School of Information Science (ISIS), the preservation of this valuable collection has been possible.
3	the 11 plate volumes, owned by the Bibliotheca Alexandrina, and the 9 text volumes, owned by l'Institut d'Egypte, have been fully digitized, integrated on a virtual browser and made accessible to the public through this pioneering endeavor.
4	description de l'Egypte
5	description de l'Egypte was the outcome of the collaboration of more than 150 prominent scholars and scientists who accompanied Bonaparte to Egypt in 1798.
6	they systematically examined almost every aspect of contemporary and ancient Egyptian civilization.
7	the description de l'Egypte was the outcome of a 20-year process, dependent on the talents and industry of some 2,000 mechanics, draftsmen, cartographers, typographers and engravers.
8	nothing similar had ever been attempted before.
9	not only did it give an illustrated account of Egypt's historic, cultural, artistic and religious treasures;

View as: | Graph | UNL | CDL | RDF |

```
1    {#event:
2        {#02
3            <49:textbook(icl>document)>
4            <4S:unmatched(aoj>thing)>
5            <5I:detail(icl>attribute)>
6            <3Y:volume(icl>book)>
7            <4I:plate(icl>photograph)>
8            [4I and 3Y]
9            [3Y mod 49]
10           [4I mod 5I]
11           [4S aoj 5I]
12       }
13       {#01
14           <5I:detail(icl>attribute)>
15           <55:accuracy(icl>attribute)>
16           [5I and 55]
17       }
18       {#03
19           <22:contemporary(aoj>thing)>
20           <2J:ancient(mod<thing)>
21           [2J and 22]
22       }
23       <3I:produce>
24       <13:examine>
25       <09:20>
26       <0C:year>
27       <34:civilization>
```

Picture 3: UNL View

RDF View

The RDF View renders the graph as an RDF document where Sentences(sub sentences), Relations and Entities (Universal Words) are defined as objects using "<rdf:description>" tags (see Picture 4). Corresponding namespaces and unique URIs are being generated and referenced. The code is highlighted accordingly.

CWL Conversion Framework - Microsoft Internet Explorer

File   Edit   View   Favorites   Tools   Help

Back | Search | Favorites | Links »

Address http://localhost:8080/graphs/main.jsp?unldoc=egypt.unl | Go

CWL Conversion Framework

Original

1 — for over 20 years, they systematically examined almost every aspect of contemporary and ancient Egyptian civilization, producing 20 volumes of text and plates of unmatched accuracy and detail.

2 — today, with the efforts of the BA international School of Information Science (ISIS), the preservation of this valuable collection has been possible.

3 — the 11 plate volumes, owned by the Bibliotheca Alexandrina, and the 9 text volumes, owned by l'Institut d'Egypte, have been fully digitized, integrated on a virtual browser and made accessible to the public through this pioneering endeavor.

4 — description de l'Egypte

5 — description de l'Egypte was the outcome of the collaboration of more than 150 prominent scholars and scientists who accompanied Bonaparte to Egypt in 1798.

6 — they systematically examined almost every aspect of contemporary and ancient Egyptian civilization.

7 — the description de l'Egypte was the outcome of a 20-year process, dependent on the talents and industry of some 2,000 mechanics, draftsmen, cartographers, typographers and engravers.

8 — nothing similar had ever been attempted before.

9 — not only did it give an illustrated account of Egypt's historic, cultural, artistic and religious treasures;

View as: | Graph | UNL | CDL | RDF |

```
1   <rdf:RDF xmlns:rdf='http://www.w3.org/1999/02/22-rdf-syntax-ns#'
        xmlns:unl='http://www.undl.org/unl#' xmlns:rdfs='http://www.w3.org/2000/01/rdf-schema#'>
2       <rdf:Description rdf:about='http://www.undl.org/sentence#s1'>
3           <rdf:type rdf:resource='http://www.undl.org/sentence'>
4           <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r0-dur'/>
5           <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r1-agt'/>
6           <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r2-man'/>
7           <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r3-obj'/>
8           <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r4-aoj'/>
9           <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r5-bas'/>
10          <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r6-man'/>
11          <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r7-mod'/>
12          <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r8-mod'/>
13          <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r9-aoj'/>
14          <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r10-aoj'/>
15          <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r11-pur'/>
16          <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r12-obj'/>
17          <unl:hasRelation rdf:resource='http://www.undl.org/unl/relation#s1-ss-r13-qua'/>
18          <unl:hasSubSentence rdf:resource='http://www.undl.org/unl/sentence#s1-ss02'/>
19          <unl:hasSubSentence rdf:resource='http://www.undl.org/unl/sentence#s1-ss01'/>
20          <unl:hasSubSentence rdf:resource='http://www.undl.org/unl/sentence#s1-ss03'/>
21      </rdf:Description>
22      <rdf:Description rdf:about='http://www.undl.org/unl/relation#s1-ss-r0-dur'>
23          <rdf:type rdf:resource='http://www.undl.org/unl/relation'>
24          <unl:relationType rdf:resource='http://www.undl.org/unl/relation-type#dur'/>
25          <unl:sourceEntity rdf:resource='http://www.undl.org/unl/uw#s1-13'/>
26          <unl:targetEntity rdf:resource='http://www.undl.org/unl/uw#s1-0C'/>
```

Picture 4: RDF View