WP11
Multilingual Semantic Wiki

Kaarel Kaljurand

Institute of Computational Linguistics,

University of Zurich

MOLTO Project Review, Barcelona

2012-03-20

Multilingual Semantic Wiki

- *multilingual*

  - multilingual interface for editing and querying

  - synchronised multilingual content

- *semantic*

  - notions of consistency, entailment, automatic Q&A

- *wiki*

  - collaborative

  - fast and accessible

Existing background technologies

- Grammatical Framework and its Resource Grammar Library

- Attempto Controlled English (ACE): first-order language with English syntax

- AceWiki

Main tasks of WP11

- port the ACE grammar to 15 MOLTO languages (assisted by UGOT)

- develop AceWiki into a *multilingual* controlled language wiki-system

    ◦ modularize and extend AceWiki

    ◦ integrate existing GF tools for translation and smart text input

- evaluate the new multilingual wiki-system (assisted by UHEL)

AceWiki

- goal: user-friendly expressive semantic wiki system

- collaborative editing (web-based tool)

- background reasoning language: OWL

    ◦ expressive fragment of first-order logic

    ◦ decidable reasoning tasks: consistency checking, question answering, …

    ◦ complex syntax

- front-end language: ACE

    ◦ subset of natural English

    ◦ well-defined translation into first-order logic

    ◦ end-user documentation: construction and interpretation rules

- see more: http://attempto.ifi.uzh.ch/acewiki/

Bidirectional ACE-to-OWL mapping

Every country that does not border a sea is a landlocked-country.

```
SubClassOf(
    ObjectIntersectionOf(
        :country
        ObjectComplementOf(
            ObjectSomeValuesFrom(
                :border
                :sea
            )
        )
    )
    :landlocked-country
)
```

Which country is a landlocked-country?

```
ObjectIntersectionOf(
    :country
    :landlocked-country
)
```

Article (screenshot)

<Back Forward> Refresh

AceWiki

Article | Noun | References | Individuals | Hierarchy |

planet

▷ *We use here the definition of "planet" according to the International Astronomical Union (see http://www.iau.org) without the restriction to solar planets.*

▷ Every planet is a celestial-body.

▷ No planet is a star.

▷ No planet is a dwarf-planet.

▷ *The distinction between planet and dwarf-planet has been introduced by the International Astronomical Union in 2006.*

▷ No planet is a moon.

▷ Every planet orbits a star.

▷ Every planet that does not orbit the Sun is an extrasolar planet.

▷ Which planets orbit the Sun?
  - Earth
  - Jupiter
  - Mars
  - Mercury
  - Neptune
  - Saturn
  - Uranus
  - Venus

Navigation:
- Main Page
- Index
- Random Article
- Search:

Actions:
- New Word...
- Export...

Look-ahead editor (screenshot)

ACE Text Editor ✕

Every country that has a capital controls ...

< Delete

text

```
[                                                                    ]
```

function word	*proper name*	*new variable*	*reference*
a	new...	X	itself
an	Austria	Y	it
at least	Berlin	Z	the capital
at most	Bill	X1	the country
every	Canada	Y1	
everybody	France	Z1	
everything	Germany	X2	
exactly	Italy	Y2	
less than	Jane	Z2	
more than	John	X3	
no	Lara	Y3	
nobody	Mary	Z3	
nothing	Mike	X4	
somebody	Paris	Y4	
something	Rome	Z4	
what	Russia	X5	

OK   Cancel

Reasoning (screenshot)

AceWiki in MOLTO

- preserve the features of the AceWiki

  - controlled natural language

  - Attempto Controlled English

  - notions of consistency, automatic question answering, ...

  - wiki aspects (collaborative editing, multiple interlinked articles)

  - look-ahead editing

- extend AceWiki to be multilingual

  - base the implementation on GF

Generalization of AceWiki

- multiple languages

  - natural: English, German, ACE, ...

  - formal: ACE, Sage, ...

  - languages for content, UI, meta information

- multiple grammars

  - ACE (or its subsets)

  - Math Grammar Library

  - Phrasebook

  - ...

- multiple reasoners

  - ACE-based (RACE, OWL-based, ...)

  - math reasoners, e.g. Sage, WolframAlpha

  - ...

Multiple languages

- multiple languages for
    - content
    - UI (labels etc.)
    - meta queries (authors, edits)
- the content
    - viewable/editable/queryable in multiple languages
    - automatically kept in sync
- some languages are formal, i.e. they are (mainly) meant for the reasoners
- implemented using Grammatical Framework (GF)
    - **single** abstract syntax corresponds to **multiple** concrete syntaxes
    - Resource Grammar Library

(Stable AceWiki: single language: OWL-compatible subset of ACE)

Multiple grammars

- different grammars in different sections of the wiki

- interaction between grammars

- support for changing the grammars

    - full grammar editor

    - UI for adding/editing words and their forms

- implemented by integrating multiple GF grammars and a GF editor

(Stable AceWiki: single grammar: OWL-compatible subset of ACE)

Multiple reasoners

Syntactic vs semantic reasoning

- syntactic

    - queries over GF trees to answer e.g. "for a given sentence, show me all the syntactically similar sentences"

- semantic (e.g. ACE-based, math reasoning)

    - consistency checking

    - Q&A

    - ...

    - explanation of reasoning results

- implemented by using external reasoning tools

(Stable AceWiki: single reasoner: OWL-reasoner with multiple implementations)
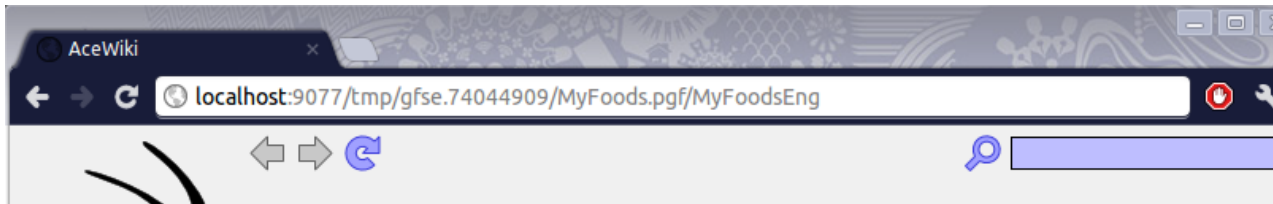
Current integration with GF

is based on the GF Webservice and the GF online editor.

- storage is now based on (the language-neutral) abstract trees. (In case of standard AceWiki, the storage is based on ACE.)

- access to multiple grammars (PGF)

- multilingual viewing and editing of wiki content

- look-ahead editing of sentences

- presentation of the GF-analysis of sentences (translations, word alignment diagrams, GF syntax trees, …)

- preliminary grammar editing using the online GF editor

    ◦ http://cloud.grammaticalframework.org/gfse/

- GF webservice accessed via a new Java API

    ◦ https://github.com/Kaljurand/GF-Java

AceWiki and the online GF editor

AceWiki

Wiki in
MyFoodsEng

Navigation:
- Main Page
- Index
- Search
- About
- Random Article

Actions:
- Edit grammar
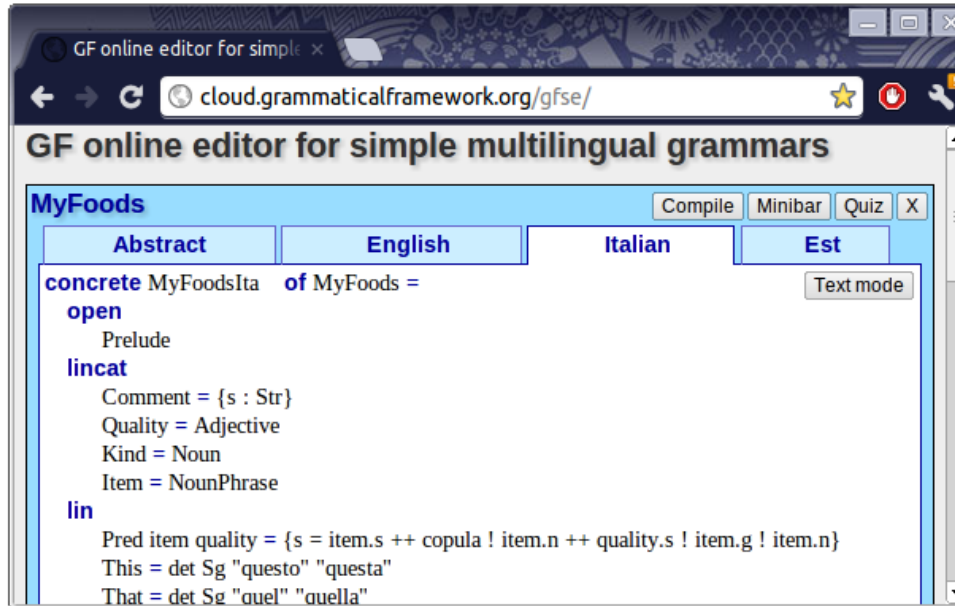- Export...

Languages:
- MyFoodsIta
- MyFoodsEst
- MyFoodsEng

localhost:9077/tmp/gfse.74044909/MyFoods.pgf/MyFoodsEng

Main Page

▷ Comment: This page contains sentences about food.
▷ these boring expensive Italian pizzas are delicious
▷ that cheese is fresh
▷ those very very very fresh wines are Italian
▷ this cheese is cold
▷

GF online editor for simple

cloud.grammaticalframework.org/gfse/

GF online editor for simple multilingual grammars

MyFoods   Compile | Minibar | Quiz | X

| Abstract | English | Italian | Est |

Text mode

```
concrete MyFoodsIta    of MyFoods =
  open
      Prelude
  lincat
      Comment = {s : Str}
      Quality = Adjective
      Kind = Noun
      Item = NounPhrase
  lin
      Pred item quality = {s = item.s ++ copula ! item.n ++ quality.s ! item.g ! item.n}
      This = det Sg "questo" "questa"
      That = det Sg "quel" "quella"
```

Immediate future work

- structuring the content

    ◦ multiple articles

    ◦ sentence order

- viewing/querying the content

    ◦ dynamic views based on queries over the GF abstract trees

ACE in GF

- starting point: *Angelov and Ranta (2009)*

  - based on ACE v6.0 (early 2009)

  - covers 7 languages (multilinguality via the GF RGL)

  - languages: *Eng*, *Ger*, *Ita*, *Swe*, *Fin*, *Fre*, *Urd*

- issues with *Angelov and Ranta (2009)*

  - coverage and precision not perfect with respect to ACE v6.6

  - no AceWiki-oriented subset

  - small number of languages

- new developments (https://github.com/Attempto/ACE-in-GF)

  - minor updates and fixes

  - new AceWiki-oriented subset (less overgeneration in the AceWiki context)

  - work on a precise AceWiki-oriented subset in GF continues as a collaboration between UZH and UGOT

Dissemination

- paper or poster at the CNL 2012

    ◦ August 2012 in Zurich, Switzerland

- AceWiki repository on GitHub

    ◦ https://github.com/AceWiki

    ◦ source code history, wiki, issues list

- Attempto mailing list

Thank You!