Two Level Approach to SRL

Milen Kouylekov (kouylekov@celi.it)

Roma, 24 January 2012

Outline

Frame Labeling
Boundary Detection
Semantic Role Labeling

Frame Labeling

1) Extract the **Context** of the Frame Lemma
2) Create a **set** of words for each frame
   - **CauseHarm** "*sfruttare, torturare, gravemente, minacciare incident*"
3) Compare the **context** of a candidate with the **set** of each Frame using **cosine similarity**

Boundary Detection

- LFC parser
- Example:

"A favore delle popolazioni di regioni colpite da catastrofi"

  - PP: di regioni colpite
  - PP: a favore delle popolazioni
  - …

Semantic Role Labeling

- Our approach to SRL we aimed at:
  - Maximize precision over recall.
- Roles can be assigned mostly on the basis of the subcategorization list
  - sembrare VERB^OBJ Inference
  - sembrare NOUN^SUBJ Phenomenon
- Not Working: Parsing error; Verbal alternation; Ambiguity of verbs; Ambiguity of arguments

Markov Logic Networks

- Reidel 2008 – theBeast
- Relational Learning language
  - First Order Logic
  - Markov Networks
- Template rules
  - Expanded in all Logical Possibilities
  - Assigned a Weight

Example

for Int a, Int p, Lemma l, Role r, FrameLabel fr

if plemma(a,l) &
isPredicate(p) &
possibleArgument(a)&
evoke(p,fr)

add[role(p,a,r)] * w_lemma_sframe_a(l,r,fr);

Type of Rules

1)Linear rules considering word features (POS) in a certain window

2)Rules considering distance between the predicate and its possible arguments.

3)Rules taking into account the compatibility of certain features of the predicate word and its possible argument word (lemma, surface form, part of speech, frame…)

4)Rules considering dependency relations between the predicate and its possible arguments.

5)Rules taking into account the computed subcategorization list and the features of the possible argument.

Analysis

- Low Recall
  - Role assignment based on the sub-categorization list were missing
  - Some role wrongly assigned on the basis of word combination rather than dependency
- Small dimensions of the corpus

Tuffy

- "**Correct**" the output of the trained theBeast system
- Rules are manually coded
- Manually assigned weights
- All dependency based
- Example:
  - !evoke(v0, v1, Statement_proclamare) v !dep(v0, v1, v2, SUBJ) v assignT(v0, v1, v2, Speaker)
  - !evoke(v0, v1, Statement_dire) v !dep(v0, v1, v2, OBJ) v assignT(v0, v1, v2, Occasion)

Results

	AC (p)	AC (r)	AC (f)	Token (p)	Token(r)	Token (f)
With Tuffy	**75.0**	40.18	52.33	**83.24**	51.49	63.62
No Tuffy	73.23	32.32	44.86	76.58	36.34	49.29

- The mixed approach with learnt weights and manually coded rules seems promising.
- Poor recall due to parser errors.

Linguagrid

http://www.linguagrid.org

Share and Control Web Services

europeana
pensare la cultura

organic
lingua

CACAO
PROJECT
CROSS-LANGUAGE ACCESS TO CATALOGUES AND ON-LINE LIBRARIES

CALATEAS

Scoop

LAW