The 13th Workshop on Domain-Specific Modeling

model[NL]generation: Natural Language Model Extraction

Lars Ackermann (M. Sc.)
E-Mail: lars.ackermann@uni-bayreuth.de

27.10.2013
George: Condi! Nice to see you. What's happening?
Condi: Sir, I have the report here about the new leader of China. [...] Hu is the new leader of China.
George: That's what I want to know.
Condi: That's what I'm telling you.
George: That's what I'm asking you. Who is the new leader of China?
Condi: Yes. [...] George: Will you or will you not tell me the name of the new leader of China?
Condi: Yes, sir.
George: Yassir? Yassir Arafat is in China? I thought he was in the Middle East. [...]
Larger Scope

- Information about source- and target language and speaker
- Ideally, this is achieved the same way as reconstructing a model
- There is no jump involved, levels are reconstructed from bottom to top

Informal, textual description

Information Extraction 
Natural Language Processing

Formal model

Model reconstruction

Information about modeling patterns

Model reconstruction

Formal description of a model instance

Information about source- and target language and speaker

Ideally, this is achieved the same way as reconstructing a model

There is no jump involved, levels are reconstructed from bottom to top

M0

M1

M2
Not only in IT (but notably there) ...

Up to now

Customer → Docu → Docu → Docu → Subjectively created target model

model[NL]generation

Customer → model[NL]generation → model[NL]generation → Systematically generated target model

Getting **flexibly, intuitively and interactively** towards the target model using **natural language**.
model[NL]generation: Natural Language Model Extraction

1. Flexible
Choosing the (domain specific) meta model

2. Intuitive
Training a system for processing natural language using natural language

Sample Sentences
“The manager notifies the customer.”
“After that, employees advise their colleagues.”

3. Interactive
Extracting target model

Model Type Library

ER
Processes
UML
Many users, even more formulations – one Activity: We need flexible detection rules

“The manager notifies the customer.”

“The customer is contacted by the manager.”

“The manager is responsible for the contact to the customer.”

Activity

name : contacts
First attempt of detection rule definition: Plain Syntax

- **Example:**
  
  "The manager notifies the customer."

- **Noun Design Strategy**
  
  [Rumbaugh1991]

- **Great Basis:**
  
  Such grammatical parsers already exist!
Everything is fine so far ...

... but only so far. Syntax is ambiguous!

- **Example 1:** “The manager *notifies* the customer.”
- **Example 2:** “The manager *is* the executive.”

- Adjust detection rules to "notifies"  
  → Lexicalization of production rules

- But ...
  “The managers notify the customer.”
  “The manager notifies the customers.”
  “The managers notify many customers.”
  “The MN notifies the customer.”
  “Managers notify the customer.”
  “Max notifies the customer.”
  “She notifies many customers.”
  “The alarm notifies the customer.”

**Actor = Person?!!**

---

**POS Tags**

- **notifies/VBZ**
- **is/VBZ**

<table>
<thead>
<tr>
<th>S</th>
<th>→ NP VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP</td>
<td>→ notifies NP</td>
</tr>
<tr>
<td>NP</td>
<td>→ DT NN</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S</th>
<th>→ NP VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>VP</td>
<td>→ notifies NP</td>
</tr>
<tr>
<td>VP</td>
<td>→ notify NP</td>
</tr>
<tr>
<td>VP</td>
<td>→ ...</td>
</tr>
<tr>
<td>NP</td>
<td>→ DT NN</td>
</tr>
<tr>
<td>NP</td>
<td>→ DT NNS</td>
</tr>
<tr>
<td>NP</td>
<td>→ ...</td>
</tr>
<tr>
<td>...</td>
<td>→ ...</td>
</tr>
</tbody>
</table>
Creating detection rules – 1

• Requirement 1: **Detection** ability

  - The manager notifies the customer.
  - Users meet an administrator.

• Requirement 2: **Disambiguation** ability

  - The manager is the executive.
  - The manager interacts with the executive.
Creating detection rules – 2

**Syntactic basis**

- Syntax tree for given samples
- Reduce them to be more general

Remaining issues:
- Disambiguation
- Specialization (Domain, User, ...)

Automatic feature extraction and abstraction

The manager notifies the customer
Creating detection rules – 3

Object of interest

Select the sentence parts which are the most relevant [Collins1999]

“The manager”
“All managers”
“The account manager”
“All assembly managers”

“notifies the customer”
“meets an administrator”
“informs the CEO”
“teach all pupils”

S

(manager)
(notifies)
(customer)

NP
NP
VP

The manager
notifies customer
VP
Creating detection rules – 4

**Annotation**

Automatically add disambiguating and specializing information

“The manager notifies the customer.”
“Users meet an administrator.”

“The manager is the executive.”

---

**Diagram**

```
S
 NP
 NP
 WNCat <person>
 TypDep nsubj_dep
 WNCat <manager>
 TypDep nsubj_dep
 TypDep det_gov
 TypDep cop_gov

VP
 VP
 WNCat <interact>
 TypDep nsubj_gov
 TypDep dobj_gov

S
 NP
 NP
 WNCat <executive>
 TypDep det_gov
 WNCat <person>
 TypDep nsubj_gov
 TypDep cop_gov

```

---

04.12.2013 | 12
2013 (C) Databases and Information Systems Group | Lars Ackermann
Mapping: Feature key paths

1. \(\text{Concat}(\text{Lemma}(S_{\text{VP}}), S_{\text{NP}})\)
2. \(\text{Concat}(\text{Lemma}(\text{advises}), \text{employee})\)
3. “advise employee”

Process meta model element

- Activity
  - name = advise employee

- Entity
  - name = car
    - Attribute
      - name = color
    - Attribute
      - name = speed

- Entity
  - name = car
  - Attribute
    - name = color
  - Attribute
    - name = speed
Vision

Dynamic Domains

Flexible linguistic knowledge

Bi-directional transformation

Formalized Knowledge

Source

Using linguistic knowledge

ER

Processes

UML

Extraction

Generation

Linguistic fingerprint

Domain-specific phrases
Main Literature:


Further Literature:


Images.

#2: http://invincibles.in/tag/marvin/, 20.10.2013
Thank you for your attention

Questions?

DON'T PANIC