Integrating a Dialog Component into a Framework for Spoken Language Understanding

Sebastian Weigelt, Tobias Hey, and Mathias Landhäußer
“The only way a person can truly concentrate on his problem and solve it [. . .] are if he is able to communicate directly with the computer without having to learn some specialized intermediate language.”

Jean E. Sammet, 1966
Programming should never be a one way street

Traditional Programming

Your condition misses a dependent statement!

What does „it“ refer to?

Programming in Natural Language
Dialog Systems (Related Work)

Rule-based Dialog Systems

- call-flow-based
- frame-based
- agenda-based
- information state-based

Statistical Dialog Systems

- neuronal networks-based
- POMDP-based

[Mct98] [Pie01] [Sen00] [Boh03] [Rud99] [Mor14] [Ser16] [Wen16] [Li16] [Roy00] [Tho08] [You10] [Gas14]
Dialog Systems: Integration Challenges

Integrating a Dialog Component into a Framework for Spoken Language Understanding | S. Weigelt, T. Hey, and M. Landhäußer
Dialog Systems: Integration Challenges

Modular System

Dialog

Integrating a Dialog Component into a Framework for Spoken Language Understanding | S. Weigelt, T. Hey, and M. Landhäußer
Dialog Systems: Goal

Existing System (PARSE)

Dialog Module
Application Example: PARSE

Architecture [Wei15]

Domain Ontology

SLU

Agent A
context acquisition [Wei17]

Agent B
conditional detection [Wei18]

Agent C
coreference

Pre
P1 ... Pn

Post
P1 ... Pn

ARMAR-III
Alice
(3D animations)
openHAB
(home automation)
# Challenges & Approach

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit in existing framework (PARSE)</td>
<td>Implement as PARSE agent</td>
</tr>
<tr>
<td>Reactivity</td>
<td>Use indicators for (unresolvable) language understanding problems</td>
</tr>
<tr>
<td>Extensibility</td>
<td>Dialog acts</td>
</tr>
<tr>
<td></td>
<td>Chain of responsibility</td>
</tr>
</tbody>
</table>
Dialog Systems: General Architecture

- Speech Recognition
- Language Understanding
- Dialog Management
- Speech Synthesis
- Language Generation
Approach: A Dialog Agent for PARSE

Pre-processing

P1  ...  Pn

SLU Framework

Agent A

Agent B

Agent C

Post-processing

Integrating a Dialog Component into a Framework for Spoken Language Understanding | S. Weigelt, T. Hey, and M. Landhäußer
Approach: A Dialog Agent for PARSE

Pre-processing
- Speech Recog.
- Agent A
- Dialog Agent
  - Dialog Management
  - Language Generation
  - Speech Synthesis
- Post-processing

Speech Recognition
...
Pn
SLU Framework
- Language Understanding
- Dialog Management

Agent B

Agent C
Approach: Dialog Acts & Chain of Responsibility

- **Dialog Acts**
  - One per problem class
  - It comprises
    - An indicator: graph to pattern to identify problems
    - A dialog model: agenda-based
    - A language generation: slot-filling

- **Chain of Responsibility**
  - Orders Dialog Acts
  - Extensible
Dialog Act: Coreference Ambiguity

```
"go to the dishwasher next to the fridge and open it"
```

**Coreference Ambiguity DA**

**Indicator:** coref\(f_1\)(X,A)\(\sim\)coref\(f_2\)(X,B)

**Dialog Model:**
- Match Response
- Update Graph

**Language Generation:**
- In the following what does REF.EX refer to CLAUSES?
- I can't understand you've mentioned E entities. Tell me what does REF.EX refer to LIST_ENT or LIST_EMIT?
## Problem classes: Overview

<table>
<thead>
<tr>
<th>Information</th>
<th>Problem</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>speech recognition</td>
<td>uncertainty</td>
<td>low word confidence</td>
</tr>
<tr>
<td>coreference</td>
<td>ambiguity</td>
<td>corefConf(x, a) \sim= \text{corefConf}(x, b)</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>pronoun without reference</td>
</tr>
<tr>
<td></td>
<td>uncertainty</td>
<td>low confidence of sole reference</td>
</tr>
<tr>
<td>conditional</td>
<td>incomplete</td>
<td>conditional clause without then-clause</td>
</tr>
<tr>
<td></td>
<td></td>
<td>…</td>
</tr>
</tbody>
</table>
Problem classes: Speech recognition uncertainty

<table>
<thead>
<tr>
<th>Information</th>
<th>Problem</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>speech recognition</td>
<td>uncertainty</td>
<td>low word confidence</td>
</tr>
<tr>
<td>coreference</td>
<td>ambiguity</td>
<td>corefConf(x, a) ~ corefConf(x, b)</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>pronoun without reference</td>
</tr>
<tr>
<td></td>
<td>uncertainty</td>
<td>low confidence of sole reference</td>
</tr>
<tr>
<td>conditional</td>
<td>incomplete</td>
<td>conditional clause without then-clause</td>
</tr>
</tbody>
</table>

…

“... take\textsubscript{0.91} the\textsubscript{0.82} right\textsubscript{0.72} fridge\textsubscript{0.89} ...”
Problem classes: Coreference ambiguity

<table>
<thead>
<tr>
<th>Information</th>
<th>Problem</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>speech recognition</td>
<td>uncertainty</td>
<td>low word confidence</td>
</tr>
<tr>
<td>coreference</td>
<td>ambiguity</td>
<td>corefConf(x, a) ~ corefConf(x, b)</td>
</tr>
<tr>
<td></td>
<td>missing</td>
<td>pronoun without reference</td>
</tr>
<tr>
<td></td>
<td>uncertainty</td>
<td>low confidence of sole reference</td>
</tr>
<tr>
<td>conditional</td>
<td>incomplete</td>
<td>conditional clause without then-clause</td>
</tr>
</tbody>
</table>

“... dishwasher and fridge ... open it”
### Problem classes: Incomplete Conditionals

<table>
<thead>
<tr>
<th>Information</th>
<th>Problem</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>speech recognition</td>
<td>uncertainty</td>
<td>low word confidence</td>
</tr>
<tr>
<td>coreference</td>
<td>ambiguity</td>
<td>corefConf(x, a) ( \sim ) corefConf(x, b)</td>
</tr>
<tr>
<td>coreference</td>
<td>missing</td>
<td>pronoun without reference</td>
</tr>
<tr>
<td>coreference</td>
<td>uncertainty</td>
<td>low confidence of sole reference</td>
</tr>
<tr>
<td>conditional</td>
<td>incomplete</td>
<td>conditional clause without then-clause</td>
</tr>
</tbody>
</table>

“if the laundry is done **put it into the dryer ...**”

**conditional clause**

**undetected then-clause**
Chain of responsibility

- Problems may be connected
- Solving a problem may make an invocation of another Act obsolete

Organize Dialog Acts in a chain of responsibility

DA1
Solve speech recognition uncertainties

DA2
Solve coreference ambiguities

DA3
Solve incomplete conditionals

...
Chain of responsibility

- Problems may be connected
- Solving a problem may make an invocation of another Act obsolete

Organize Dialog Acts in a chain of responsibility

- DA1: Solve speech recognition uncertainties
- DA2: Solve coreference ambiguities
- DA3: Solve incomplete conditionals

... you fill it into the cup then you ...
Evaluation: Setting

- 10 subjects
- 3 scenarios
  - Scenario 1: explain robot how to perform a task (laundry)
  - Scenarios 2 & 3: existing recordings (from previous studies)

Scenario 1: free dialog

“... robo take the laundry from the washing machine and put it into the dry hair and start it ...”
Evaluation: Setting

- 10 subjects
- 3 scenarios
  - Scenario 1: explain robot how to perform a task (laundry)
  - Scenarios 2 & 3: existing recordings (from previous studies)

Scenario 2: word errors and ambiguous coreference

“Armar, can you get the green cup … please Philip afterwards with water from the fridge … then you can bring the cup to me”
Evaluation: Setting

- 10 subjects
- 3 scenarios
  - Scenario 1: explain robot how to perform a task (laundry)
  - Scenarios 2 & 3: existing recordings (from previous studies)

Scenario 3: conditional (undetected then-statement)

“... if there are dirty dishes please put them into the dishwasher ...“
Evaluation: Results Speech Recognition Uncertainty

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø questions</td>
<td>1.81</td>
</tr>
<tr>
<td>error rate</td>
<td>0.02</td>
</tr>
<tr>
<td>success rate</td>
<td>0.25</td>
</tr>
<tr>
<td>resolution rate</td>
<td>0.23</td>
</tr>
</tbody>
</table>

\[
\text{error rate} = \frac{\# \text{ newly introduced errors}}{\# \text{ dialog acts}}
\]

\[
\text{resolution rate} = \frac{\# \text{ solved errors}}{\# \text{ dialog acts}}
\]

\[
\text{success rate} = \frac{\# \text{ successful dialog acts}}{\# \text{ dialog acts}}
\]
Evaluation: Results Speech Recognition Uncertainty

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø questions</td>
<td>1.81</td>
</tr>
<tr>
<td>error rate</td>
<td>0.02</td>
</tr>
<tr>
<td>success rate</td>
<td>0.25</td>
</tr>
<tr>
<td>resolution rate</td>
<td>0.23</td>
</tr>
</tbody>
</table>

\[
\text{error rate} = \frac{\# \text{ newly introduced errors}}{\# \text{ dialog acts}}
\]

\[
\text{resolution rate} = \frac{\# \text{ solved errors}}{\# \text{ dialog acts}}
\]

\[
\text{success rate} = \frac{\# \text{ successful dialog acts}}{\# \text{ dialog acts}}
\]
Evaluation: Results Coreference Ambiguity

<table>
<thead>
<tr>
<th>Scenario 2</th>
<th>you -&gt; you</th>
<th>it -&gt; green cup</th>
<th>∑</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø questions</td>
<td>2.00</td>
<td>1.50</td>
<td>1.75</td>
</tr>
<tr>
<td>resolution rate</td>
<td>0.40</td>
<td>0.60</td>
<td>0.50</td>
</tr>
</tbody>
</table>

\[
\text{error rate} = \frac{\text{# newly introduced errors}}{\text{# dialog acts}}
\]
\[
\text{resolution rate} = \frac{\text{# solved errors}}{\text{# dialog acts}}
\]
\[
\text{success rate} = \frac{\text{# successful dialog acts}}{\text{# dialog acts}}
\]
Evaluation: Results Incomplete Conditionals

<table>
<thead>
<tr>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ø questions</td>
</tr>
<tr>
<td>resolution rate</td>
</tr>
</tbody>
</table>

error rate = \[ \frac{\text{# newly introduced errors}}{\text{# dialog acts}} \]

resolution rate = \[ \frac{\text{# solved errors}}{\text{# dialog acts}} \]

success rate = \[ \frac{\text{# successful dialog acts}}{\text{# dialog acts}} \]
Conclusion

Objective: Integrate dialog into existing system
- Reactive
- Extensible
Conclusion

- Objective: Integrate dialog into existing system
  - Reactive
  - Extensible

- Approach: one dialog act per problem class
Conclusion

- **Objective**: Integrate dialog into existing system
  - Reactive
  - Extensible

- **Approach**: one dialog act per problem class
  - Indicators for language understanding problems
Conclusion

- Objective: Integrate dialog into existing system
  - Reactive
  - Extensible

- Approach: one dialog act per problem class
  - Indicators for language understanding problems
  - Dialog modelling: PARSE and dialog act
Conclusion

- **Objective:** Integrate dialog into existing system
  - Reactive
  - Extensible

- **Approach:** one dialog act per problem class
  - Indicators for language understanding problems
  - Dialog modelling: PARSE and dialog act
  - Chain of responsibility

- **Evaluation:** user study
  - Resolution rates up to 50% (negligible errors)

- **Future Work**
  - More dialog acts
  - Improved wording and dialog models
References (1)


[Gam94] Erich Gamma, Richard Helm, Ralph Johnson, and John Vlissides. 1994. Design Patterns: Elements of Reusable Object-Oriented Software. Pearson Education.


References (2)


