Chatbots for Language Learning

Anja Reusch

Technische Universität Dresden

Analyse eines Forschungsthemas (INF-D-960)

18.05.2018
Beginner talks to native speaker

Beginner has limited knowledge of vocabulary and grammar
What are you doing?

I am reading.

Interesting. What are you reading?

System

Beginner
Overview

1. Problem
2. Approaches for Dialogue Systems
   - Rule Based Systems
   - Statistical Systems
   - Neural Systems
3. Evaluation of Dialogue Systems
4. Approaches for Grammar Checkers
   - Rule Based Systems
   - Statistical Systems
   - Hybrid Systems
5. Evaluation of Grammar Checkers
6. Conclusion
Problem

- User...
  - has limited knowledge (vocabulary, grammar)
  - does not understand native speakers
  - needs to practice to get better

- Grammar checker...
  - checks the users input for grammar mistakes

- Dialogue system...
  - communicates with user
  - can handle a small set of topics (e.g., shopping, asking for directions)
  - uses vocabulary and grammar the user knows
Dialogue System

- System $C = (\Sigma, \Gamma, reply)$
- Input alphabet $\Sigma = \{I, you, go, dog, ...\}$, output alphabet $\Gamma$
- Dialog $D = (u_0, u_1), (u_2, u_3), ..., (u_{2i}, u_{2i+1})$ with $i \in \mathbb{N}$ and utterances $u_{2j} \in \Sigma^*, u_{2j+1} \in \Gamma^*, 0 \leq j \leq i$
- $\mathcal{D}$ set of all dialogues
- $reply : \Sigma^* \times \mathcal{D} \times \Theta \rightarrow \Gamma^*$
  - $reply(u, D, \theta) = u'$ defines the reply of the system
- User input $u$
- Dialog history $D$
- Context parameters $\theta$
Rule Based Systems [Dr.00]

- Matches input with rules

<table>
<thead>
<tr>
<th>Rule:</th>
<th>Hello *!</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hi, how are you?</td>
<td></td>
</tr>
</tbody>
</table>

Hello Bot!

Hi, how are you?
Analyzes input and generates output using a language model
- **word class**: set of words, e.g. \{pens, apples, computers\}
- **reply class**: type of the output, e.g. reply-items-positive

```
Input: Do you have apples?

Content Manager

Language Model → Scoring

Generate

Best

Substitution

Yes, we have <items>. we have yes. we
Yes, yes, yes ...

0.9 Yes, we have <items>. 0.5 we have yes. 0.2 we 0.2 Yes, yes, yes ...

Output: Yes, we have apples.
```
end-to-end approach using Long Short-Term Memory cells [HS97]

Words are encoded using word embeddings (high dimensional vectors)
Qualitative Evaluation

Rule Based Systems
- Rules written by hand are time consuming
- Language dependent
- Do not consider dialogue history

Statistical Systems
- More flexible than rule based systems
- Do not consider dialogue history
- Need annotated corpus

Neural Systems
- Consider dialogue history
- Needs large data set for training
Corpora and Metrics

**Corpora**
- Ubuntu Dialogue Corpus, 1 million multi-turn dialogues [LPSP15]
- Twitter dataset, [RCD10]

**Metrics**
- Hard to find suitable evaluation, [LLS^+16]
- Human judges, e.g., in [Jia09]
Grammar Checkers
Grammar Checkers

Correcting mistakes in sentences

\( f : \Sigma^* \rightarrow \Sigma^* \) ungrammatical sentence \( \mapsto \) correct sentence

Annotating mistakes

Example:
“This is you’re house.” \( \mapsto \{ (\{3\}, "Maybe you confused your and you’re") \} \)

\( g : \Sigma^* \rightarrow \mathcal{P}(\mathcal{P}([n]) \times \Sigma^*) \)

Input: potentially ungrammatical sentence

Output: One tuple for each error over positions and annotations with helpful information
Rule Based Systems [Nab03]

- Match input with rules
- If rule matches, message is displayed
- Example rule:

```xml
<rule>
  <pattern>
    <token postag="SENT_START"/>
    <marker>
      <token>Your</token>
    </marker>
    <token regexp="yes">not|an?|the</token>
  </pattern>
  <message>Did you mean <suggestion>You’re</suggestion>? </message>
</rule>
```

- Matches "Your a nice person."
Rule Based Systems

LanguageTool (https://languagetool.org/)
- Rule Based Grammar and Style Checker
- Community based
- Over 30 languages and dialects
- Over 2000 rules for European languages
Paste your own text here and click the 'Check Text' button. Click the colored phrases for details on potential errors. Or use this text to see an few of of the problems that LanguageTool can detect. What do you think of grammar checkers? Please note that they are not perfect. Style issues get a blue marker: It's 5 P.M. in the afternoon. LanguageTool 3.8 was released on Thursday, 27 June 2017. You're a nice person.

Did you mean "You're"?

- You're
- (another replacement)
- Ignore this type of error
- Report as false alarm...
- Examples...
Tries to correct ungrammatical sentence
Creates normal form:
  - Remove articles, prepositions and auxiliaries as "can" or "would"
  - Transform nouns into singular and verbs into infinitive
Builds a lattice by inserting all possible articles, prepositions,...
Uses a trigram language model to score sentences
Hybrid Systems [FYA⁺14]

- Combines rules with statistical machine translation
- Rules:
  - Rules are obtained from unigram, bigrams and trigrams of a learner corpus
  - n-gram contains error if 90 percent of its occurrences are wrong
- Translates ungrammatical sentence into grammatical sentence
- 10 candidates are created using a 4-gram language model
- Combine both corrections
Hybrid Systems \([FYA^{+}14]\)

- He go to school.
- I go to school.
- He goes to school.
- I goes to school.
- He went to school.
- I went to school.

**Rule Based**
- He go
- He goes

**Possible corrections**

**Machine Translation**
- 10 candidates

- He go to school.
- I go to school.
- He goes to school.
- I goes to school.
- He went to school.
- I went to school.
Qualitative Evaluation

Rule Based Systems
- Writing rules by hand is time consuming
- Language dependent
- Working open source system LanguageTool

Statistical Systems
- Good for Japanese English learners
- Not designed for other languages
- Does not consider context

Hybrid System
- Needs learner corpus
- Does not consider context
Metrics and Corpora

Metrics

- BLEU Score
- Precision, Recall, $F_{0.5}$-Score
- Human judges

Corpora

- Cambridge Learner Corpus (not open to the public) [Nic03]
- UD English-ESL / Treebank of Learner English [BKS+16]
- EAGLE: an Error-Annotated Corpus of Beginning Learner German [Boy10]
Conclusion

**Dialog Systems:**
- Statistic and neural systems are promising
- Not enough training data, especially in language learning topics
- Evaluation metrics not clear

**Grammar Checkers:**
- Good existing systems
- Largest corpus not open to the public

⇒ Due to these problems it is not possible to implement such a system quickly or easily.
The End


Daniel Naber, *A rule-based style and grammar checker*.

