Resources and tools for natural language design: Intents and entities

Deborah Dahl
Conversational Technologies
Conversational Interaction Conference
San Jose
March 11, 2019
Virtual assistant development tools

- Amazon Alexa Skills Kit
- Microsoft LUIS
- Google DialogFlow
- Nuance Mix
- IBM Watson
Meaning is in the form of Intents and Entities

- Concepts for representing meaning in a natural language application
- Provided by application development tools
- Developers annotate utterances with entities and intents
- Machine learning recognizes new utterances
Virtual Assistant Development steps

Collect data
- Agent logs
- Agent scripts and training materials
- Activity tracker from existing website

Diagram:
- Annotate data
- Train machine learning
- Test
- Deploy
Virtual Assistant Development steps

Collect data
- Agent logs
- Agent scripts and training materials
- Existing website

Identify intents
Identify entities and values
Structure concepts
Develop annotation guidelines

Annotate data
Train machine learning
Test
Deploy
### Annotation in the Alexa Skills Kit

#### Intents / GetNewFactIntent

#### Sample Utterances (14)

What might a user say to invoke this intent?

- tell me something about the {pieceFact} or the {pieceFact}
- what can you tell me about the {pieceFact}
- tell me an interesting fact about the {pieceFact}
Some entities (slots) in the Alexa Skills Kit

<table>
<thead>
<tr>
<th>Slot Type</th>
<th>ID (Optional)</th>
<th>Synonyms (Optional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>attack</td>
<td>Enter ID</td>
<td>Add synonym +</td>
</tr>
<tr>
<td>capture</td>
<td>Enter ID</td>
<td>Add synonym +</td>
</tr>
<tr>
<td>file</td>
<td>Enter ID</td>
<td>Add synonym +</td>
</tr>
<tr>
<td>rank</td>
<td>Enter ID</td>
<td>Add synonym +</td>
</tr>
</tbody>
</table>
Entity design

- What entities are needed?
- What are their possible values?
- What intents use them?
- What kinds of complex entities are needed?
  - Hierarchies
  - Composite entities
Entity design considerations

- Machine learning
- Capabilities of platform
- Ease of annotation, ability to be consistent
- Types of entities
  - Hierarchical
  - Composite
Where do we get ideas for entities?

Toolkits assume developers are starting with a design -- they give developers a way to execute the design.

Without a design, changes will involve substantial rework:

- Reannotation
- Re-do back-end integration

Can’t eliminate rework, even with the best design, but we can minimize it.
Finding entity candidates in user inputs: Resources

Generic NLP technology
- Stanford CoreNLP
- OpenNLP
- GATE
- NLTK

Meaning resources
- Wordnet
- dbPedia
- Knowledge graphs
- Wolfram Alpha
- Named Entity recognition tools
Natural Language Resources: Stanford Core NLP

01 Verbs → intents
02 Nouns → entities
03 Noun modifiers → entity values
04 Prepositional phrases → entities for time and location
Semantic resources: Wordnet
Is there a good Japanese or Indian restaurant near here?

possible intent lookFor
possible entity restaurant
possible compound intent lookForRestaurant
possible entity value good
possible entity value Japanese
possible entity value Indian
kinds of restaurant are [bistro, brasserie, brewpub, cafe]
restaurant is a kind of [building, edifice, structure, construction, artifact, artefact, whole, unit, object, physical_object, physical_entity, entity]
synonyms of restaurant are [restaurant, eating_house, eating_place]
Look at suggested concepts
Use suggestions to get ideas about entities and values

1. Synonyms: what might someone say instead of “restaurant”?
2. Are there specific categories of restaurant like “bistro”, “pub” that should be included?
3. What other values might there be like “Mexican”, “Japanese”, etc.?
Conclusions: Design First!

<table>
<thead>
<tr>
<th>Find Concepts</th>
<th>Organize</th>
<th>Develop</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Find intent and entity candidates</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• The application API</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Natural language processing tools</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Semantic resources</td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Concept Suggester</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Compound intents</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Complex entities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>• Create guidelines</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>THEN use vendor annotation and training tools</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>