An Ontology-Based Conversation System for Knowledge Bases

Abdul Quamar¹, Chuan Lei¹, Dorian Miller², Fatma Özcan¹, Jeffrey Kreulen², Robert J Moore¹, Vasilis Efthymiou¹

¹IBM Research-Al, ²IBM WatsonHealth
Conversational system for exploring domain specific knowledge bases

- Enable interaction using a natural dialog
- Characterized by:
  - Ability to understand and respond in natural language
  - Persistent context across turns
  - Interactive experience for data exploration

Here are types of Psoriasis in children:
- Plaque Psoriasis
- Guttate Psoriasis
- Pustular Psoriasis
- Inverse Psoriasis
- Erythrodermic Psoriasis

Here are drugs that treat Psoriasis in children:
- Anthralin
- Calcipotriene
- Corticosteroids

What are the types of Psoriasis in children?
What drug treats this condition?
Building conversational systems for domain specific knowledge bases

- Semantic understanding of the knowledge base data
  - Entities, relationships, hierarchies
- Design and automatic population of conversational artifacts
  - User intents: queries that can be answered by the knowledge base
  - Entities: domain vocabulary
  - Dialog: patterns of interaction with the user

**Goal:** Develop a principled methodology to build conversation system for querying domain-specific KBs, in a domain-agnostic way
Ontology-based system architecture

Offline Process

Dialoge -> Intent -> Entity

Query Template Generation

Well-formed Query (SQL)

Ontology

Knowledge Base

Human in-the-loop

Automated

Online Process

Structured Query (SQL)

Knowledge Base

Conversational Interface

Dialogue

Intent

Entity

Structured Query (SQL) bootstrap

Ontology

Knowledge Base

Conversational Interface

Dialogue

Intent

Entity

Bootstrap
Ontologies to describe the domain schema

- Ontologies capture the semantics of the domain schema of the knowledge base
  - Concepts
  - Data Properties
  - Relationships
- Provide a rich and expressive data model
- Powerful-object-oriented paradigm
  - Capture real world relationships: Inheritance, Union, Functional

Domain knowledge captured in the ontology + SME feedback -> Enable bootstrapping the conversation space
Bootstrapping the conversation space

**Intents:**
- Intents express the purpose or goal expressed in the user query/input
- System uses ML Classifiers/Deep NNs to identify intents

**Entities:**
- Represent real world objects relevant in the context of a user query

**Dialog:**
- Uses discovered intents, entities and context from the application to respond to the user

**Knowledge Base data:**
- Interaction with knowledge base data through structured queries
- Stored under different data models: Json, Relational
Bootstrapping the conversation space: Intent generation

- **Key concepts**
  - Stand on their own
  - Represent domain entities that a user is interested in
  - Identified using centrality analysis and statistical segregation

- **Dependent concepts**
  - Concepts in the neighborhood of key concepts
  - Use data statistics to identify categorical attributes
  - Union and Inheritance have special semantics

- **Intents: query patterns**
  - Described as subgraphs around key and dependent concepts

Leveraging the ontology structure
Bootstrapping the conversation space: Intent generation

**Query Patterns**

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Show me the <strong>Precautions</strong> for &lt;@Drug&gt;?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Concept</td>
<td>Key Concept</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Query</th>
<th>Show me the <strong>Precautions</strong> for Benazepril?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instance of Key Concept</td>
<td></td>
</tr>
</tbody>
</table>

**Look-up Pattern**

- Finds information about the key concept with reference to a dependent concept.

**Look-up Pattern Special Semantics**

- Handles dependent concepts with special semantics:
  - Union
  - Inheritance

**Additional Patterns**

- Show me the **Risks** associated with <@Drug>?
  - Dependent Concept: Union
  - Key Concept

- Show me the **Contra-Indications** associated with <@Drug>?
  - Dependent Concept
  - Key Concept

- Show me the **Black Box Warnings** associated with <@Drug>?
  - Dependent Concept
  - Key Concept
Bootstrapping the conversation space: Intent generation

Pattern 1: What Drug treats <@Indication>?

Query 1: What Drug treats Fever?

Pattern 2: What Indications are treated by <@Drug>?

Query 1: What Indications are treated by Aspirin?

Direct Relationship Pattern

The pair of identified key concepts are connected via at least one direct (one-hop) relationship between them.

Pattern 1: Give me the Drug and its Dosage that treats <@Indication>

Query 1: Give me the Drug and its Dosage that treats Fever

Pattern 2: Give me the Dosage for <@Drug> that treats <@Indication>

Query 2: Give me the Dosage for Aspirin that treats Fever

In-Direct Relationship Pattern

The pair of identified key concepts are connected via multiple hops of relationships.
Bootstrapping the conversation space: Generating training samples

Query patterns viewed as subgraphs over the domain ontology
Entities and corresponding data instances used to automate generation of training examples

Augmentation of training examples using SMEs and prior user experience
Bootstrapping the conversation space: Structured query template generation

- Each intent mapped to a structured query template
- Template populated with identified entities to generate structured query
- Structured query used to retrieve results from the KB

**Lookup Query Pattern**

**Show me the Precautions for [@Drug]?>**

<table>
<thead>
<tr>
<th>Initial Phrase</th>
<th>Dependent Concept</th>
<th>Key Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>Give me the Precautions for Ibuprofen?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Structured Query(SQL)**

SELECT oPrecautions.description FROM Precautions oPrecautions INNER JOIN Drug oDrug WHERE oPrecautions.for=oDrug.DrugID AND oDrug.name = 'Ibuprofen'

**Structured Query Template**

SELECT oPrecautions.description FROM Precautions oPrecautions INNER JOIN Drug oDrug WHERE oPrecautions.for=oDrug.DrugID AND oDrug.name = '<@Drug>'
Bootstrapping the conversation space:
Entity extraction

- All concepts from the ontology
- Categorical attributes
  - Data instances added from KB
- Domain specific synonyms for
  - Ontology concepts
  - Data instances

**Entities:**
- Concepts: Drug, Precautions, Dosage, Indication
- Risk: Contra-Indication, Black Box Warning
- Drug Interaction: DrugFood Interaction, DrugLab Interaction
- Drug: Aspirin, Ibuprofen, Citicoline, Pancreatin
- Indication: Fever, Headache, Bronchitis, Diabetes
- Contra-Indication: Cardiovascular disease, Breast carcinoma

**Entity Synonyms:**
- Adverse Effect: Side effect, adverse reaction, adverse event, AE
- Condition: disease, finding, disorder
- Drug: medicine, meds, medication, substance
- Precaution: caution, safe to give
- Dosage: dosing
- Dose adjustment: dose modification, dosing modification, dose reduction
Bootstrapping the conversation space: Building dialog

- Dialog tree
  - Defines the space of user utterances the system can recognize and respond to
  - Enables interactive experience for the user
  - Responses conditioned on
    - Combination of intents and entities identified in the user utterance
    - Context captured from previous utterances

Designed to handle both domain specific requests and general conversation management
IBM Micromedex (MDX)

An evidence-based clinical decision support application

- Summary and expanded content from the world’s biomedical literature
- Includes drug information, toxicology, diseases and conditions, and alternative medicine

MDX ontology

- 59 Concepts, 178 Properties, and 58 Relationships

Conversation space

- 36 intents
  - 14 lookup and 8 relationship patterns
  - 14 intents for conversation management
  - 52 Entities and corresponding data instances
Conversational Micromedex Works Great!!

Success rate per intent provided by users for top 10 intents

Success rate per intent provided by SMEs for top 10 intents

Average success rate across all intents 96.3%

Top 10 intents account for 75% or workload

Overall F1-Score of Intent identification 85%
Conclusions

• Our ontology-based framework provides sufficient semantic information for building an effective conversation system
• Our bootstrapping mechanism creates a rich conversation space
• Access to prior user queries and SME knowledge further improves the precision
• Our results show that the overall conversational system is very effective in exploring knowledge bases
  • Average success rate across all intents is 96.3%