Conversational commerce:

Emerging architectures for smart & useful voice and chatbots

Victoria Livschitz
Founder and CTO
About the speaker

- Founder & CEO of Grid Dynamics from 2006, transition to CTO role in 2015
- Principal engineer at Sun; lead architect of SunGrid, world’s first public cloud, 1997-2006
- HPC engineer at Ford Labs 1994 - 1997
About the company: experts in digital transformation thru emerging technologies

Open Source  Cloud-ready  Scalable  Automated

Digital commerce technologies
Re-platforming from legacy/on-premise to microservices on the cloud

Big data & real time analytics infrastructure & applications

ML applications: NLP, AI, voice, image recognition, predictive analytics
Thesis: voice commerce is the next big thing after mobile

Voice devices entering households
• 9M smart voice devices shipped in Q1 2018
• 50% household penetration expected by 2022
• IOT makes voice control an expected feature

Technology advances by leaps and bounds
• Speech2text approaches human quality
• Deep learning-based NLU / NER enters mainstream
• Plethora of AI platforms on the market

Massive investment
• From tech giants - over $1B IBM Watson, etc.
• From VC - over $700M in 2017
• 20 AI acquisitions - siri, api.ai, viv

Better interface for many digital interactions
• Reduces engagement barrier
• Perfect for search, Q&A
• Interactive clarification of intent
## Types of CUI experience

### Conversational commerce:
- Deep, branched conversation
- Discovery, selection, recommendation
- Large result sets, ambiguity, comparisons
- Deep integration with search, catalog, ontology
- Customer-trained NLU
- Smart dialog management to arbitrate and orchestrate multiple models

### Other “deep” conversational applications

### Infobot:
- Fact-oriented
- Few intent-reply patterns

### Virtual assistant:
- Task-oriented
- Few intent-entity-reply patterns

Shallow, linear conversation

### Smalltalk:
entertainment, empathy, human mimicry
## Types of CUI experience

**Conversational commerce:**
- Deep, branched conversation
- Discovery, selection, recommendation
- Large result sets, ambiguity, comparisons
- Deep integration with search, catalog, ontology
- Customer-trained NLU
- Smart dialog management to arbitrate and orchestrate multiple models

**Other “deep” conversational applications**

<table>
<thead>
<tr>
<th>Infobot:</th>
<th>Virtual assistant:</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Fact-oriented</td>
<td></td>
</tr>
<tr>
<td>• Few intent-reply patterns</td>
<td></td>
</tr>
<tr>
<td>• Task-oriented</td>
<td></td>
</tr>
<tr>
<td>• Few intent-entity-reply patterns</td>
<td></td>
</tr>
</tbody>
</table>

Shallow, linear conversation

**Smalltalk**: entertainment, empathy, human mimicry
Grid Genie: Conversational demo by Grid Dynamics

R&D project from Grid Labs
Why Grid Genie project?

• Demonstrate technical feasibility of current state-of-the-art
• Explore the limits of current technologies
• Learn to develop, test, deploy CUI applications
• Develop reference designs and tooling for all facets of conversational application development
• Accelerate development of new CUI applications
Conversational demo design principles

• **100% open and free**: written using open source technologies in pluggable architecture
• **Multiple devices**: supports Google Home, Alexa, Siri, etc. via device adapters
• **Multiple channels**: voice-based, text-based, web-based via channel adapters
• **Multiple skills**: specific dialog skills can be added (promo, checkout, order tracking)
• **Multiple models**: mix, match and use best-of-breed ML/NLU models for specific tasks
• **Integrated with backend services** via layers of platform services and enterprise adapters
• **Containerized**: for seamless on premise or cloud deployment
Conversational commerce blueprint

channel adapter

dialog manager

context

dialog skills

domain skills

intent API

NLU API

NLG API

search

factoid

recommend

notifications

smalltalk

dialog skills

enterprise adapters

search

catalog

recommend

ontology

enterprise services

data mining / ML

search

discovery

promo

order

loyalty

checkout

enterprise services
Let’s chat about cameras

Why cameras?

• Diverse category of products: about 1,000 items in catalog
• Wide range of subcategories, from cheap point-n-shoot to high-end professional cameras
• Technical products with vastly different features
• Some folks know exactly what they want; others don’t know much about them

What we are looking to prove?

• Device interoperability
• One dialog manager arbitrating between multiple (pluggable) skills
• Taxonomy-based domain knowledge
  Better dialog thru better taxonomy
• Know the difference between advisory & order taking
• Fluidity of conversation, keeping it natural, human-like
Grid Genie Demo
Conversational commerce blueprint

Channel adapter
- dialog manager
  - context
- dialog skills
- domain skills

Enterprise adapters
- search
- catalog
- recommend
- ontology
- enterprise services

APIs
- intent API
- NLU API
- NLG API
- discovery
- factoid
- promo
- recommend
- order
- notifications
- loyalty
- smalltalk
- checkout

Data mining / ML
Microservices architecture

Dialog Services
- Intent Classifier
- NLG
- Ontology Service

Platform Services
- Auth Service
- Shipment Service
- Order Search Service
- Order Indexing Service
- Search Service
- Catalog Service
- Catalog Ingestion Public API

Adapters
- Auth Service Adapter
- Shipment Service Adapter
- Order API Adapter
- Order API
- Order API

Store
- OAuth2.0
- Shipment API
- Order API
- Order API
- Bulk Catalog API

Channel Adapters
- Google Home
- Amazon Alexa
- Facebook

Agents
- WISMO Agent
- Discovery Agent
- Small Talk Agent

Platform Storage
- MongoDB
  - Context
- Elastic Search
  - Product
  - Order

Platform Services
- Catalog ETL
- Order API
- Shipment Service
- Auth Service

Platform Services
- Catalog Ingestion Public API
- Order Indexing Service
- Order Search Service
- Shipment Service
- Auth Service
Lessons learned (so far)?

• Reference architecture works (by and large)
• Closed platforms like Google dialog flow, are easy to start with, but hard to grow with
• It is possible to write open, cross-device applications!
• But... lack of standards leave you at mercy of (frequently) changing Google/Alexa APIs
• So, continuous testing (and full true CICD) is key to spot regressions quickly
• Google and Alexa publishing standards help test your application
• Creating pluggable taxonomies is hard
• Finding good dialog writers is hard
• Best practices for writing general-purpose, cross-platform dialog managers are immature

• Conversational applications are ready for prime time
• It’s fun and practical to write them, although not (yet) cheap or easy
# Conversational platform components

<table>
<thead>
<tr>
<th>Service</th>
<th>Description</th>
<th>Responsibilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Channel adapters</td>
<td>Connect the platform to edge devices</td>
<td>Account linking, Auth, SSO, user management</td>
</tr>
<tr>
<td>Context</td>
<td>Context keeps dialog state</td>
<td>Shares context between devices and interactions</td>
</tr>
<tr>
<td>Dialog manager</td>
<td>Controls conversation flow</td>
<td>maintains conversation context, detects the customer’s broader intent, delegates tasks to an appropriate dialog agent</td>
</tr>
<tr>
<td>Pluggable dialog agents</td>
<td>Implement narrow conversation flow</td>
<td>execute particular use case, communicate with NLU, enterprise adapters to fulfill the actions, use NLG to form natural language response</td>
</tr>
<tr>
<td>Dialog services</td>
<td>NLU/NLG/NER service adapters and implementations</td>
<td>Provide NLG/NLU/NER capabilities</td>
</tr>
<tr>
<td>Enterprise adapters</td>
<td>Connect agents to data and services</td>
<td>Provide access to enterprise APIs</td>
</tr>
<tr>
<td>Domain ontology</td>
<td>Capture domain knowledge</td>
<td>deep understanding of business domain, its terminology, slang, and the relationship between terms</td>
</tr>
</tbody>
</table>
Domain specific ontology

Implementation:

- Neo4J

Content sources:

- BabelNet
- WordNet
- DBpedia
Intent classification

Can you suggest me an action camera?

GloVe/fasttext word embedding

Matrix representation

CNN: Conv layers

Pooling

Fully Connected

<table>
<thead>
<tr>
<th>CLASS</th>
<th>PROB</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCEPT</td>
<td>0.01</td>
</tr>
<tr>
<td>DENY</td>
<td>0.001</td>
</tr>
<tr>
<td>NEXT</td>
<td>0.05</td>
</tr>
<tr>
<td>PREVIOUS</td>
<td>0.012</td>
</tr>
<tr>
<td>REQUEST</td>
<td>0.011</td>
</tr>
<tr>
<td>RESET</td>
<td>0.03</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>WISMO</td>
<td>0.05</td>
</tr>
<tr>
<td>DISCOVERY</td>
<td>0.81</td>
</tr>
<tr>
<td>SELECTION</td>
<td>0.1</td>
</tr>
<tr>
<td>ORDER</td>
<td>0.05</td>
</tr>
<tr>
<td>SMALLTALK</td>
<td>0.02</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

CNN based model for intent classification
To learn more

- Subscribe to our blog: blog.griddynamics.com
Named entity recognition: BIO tagging

<table>
<thead>
<tr>
<th>I</th>
<th>need</th>
<th>diving</th>
<th>camera</th>
<th>under</th>
<th>two</th>
<th>hundred</th>
<th>bucks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>B-purpose</td>
<td>B-device_name</td>
<td>B-relational</td>
<td>B-quantity</td>
<td>I-quantity</td>
<td>B-unit</td>
</tr>
</tbody>
</table>

BIO tags (Beginning, Inside, Outside)

![Diagram showing the process of named entity recognition using GloVe/fasttext word embedding, matrix representation, BiLSTM, RNN, and softmax for entity tagging.]

BiLSTM based model for entity tagging
Named entity recognition: tag grouping

<table>
<thead>
<tr>
<th>I</th>
<th>need</th>
<th>diving</th>
<th>camera</th>
<th>under</th>
<th>two</th>
<th>hundred</th>
<th>bucks</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>O</td>
<td>B-purpose</td>
<td>B-device_name</td>
<td>B-relational</td>
<td>B-quantity</td>
<td>I-quantity</td>
<td>B-unit</td>
</tr>
</tbody>
</table>

- **to predict**
  - unrelated
  - related
  - related
  - related
  - related

Linear XGBoost model is trained to make binary classification for token pairs
About Grid Dynamics

Founded in 2006, Grid Dynamics is an engineering services company built on the premise that cloud computing is disruptive within the enterprise technology landscape. Since that time, we’ve had the privilege to help companies like Microsoft, eBay, PayPal, Cisco, Macy’s, Yahoo, ING, Bank of America, Kohl’s, among others, to re-architect their core mission-critical systems, develop new cloud services, accelerate innovation cycles, increase software quality, and automate application management.

Grid Dynamics has multiple locations in the USA and Europe, and employs over 1,000 expert engineers worldwide.
Thank you!