EE596: Conversational Artificial Intelligence Project

A Chatbot that walks you through your job search

Overview

1. Motivation
2. System Architecture
3. Dialog Design and State Machine
4. Knowledge Database
5. Data Processing
6. Sample Dialogs
7. Observations
8. Takeaways
9. Future Work

Conversational Search

The challenge: *Context*

> *“It didn’t use contextual information so there was no way to expand on the previous answer it gave me."*

Participant commenting on their experience.
(Vtyurina el al.'s 2017 work on Conversational Search)

Formulating a full question takes effort and is unnatural.

Motivation

Team Vision:

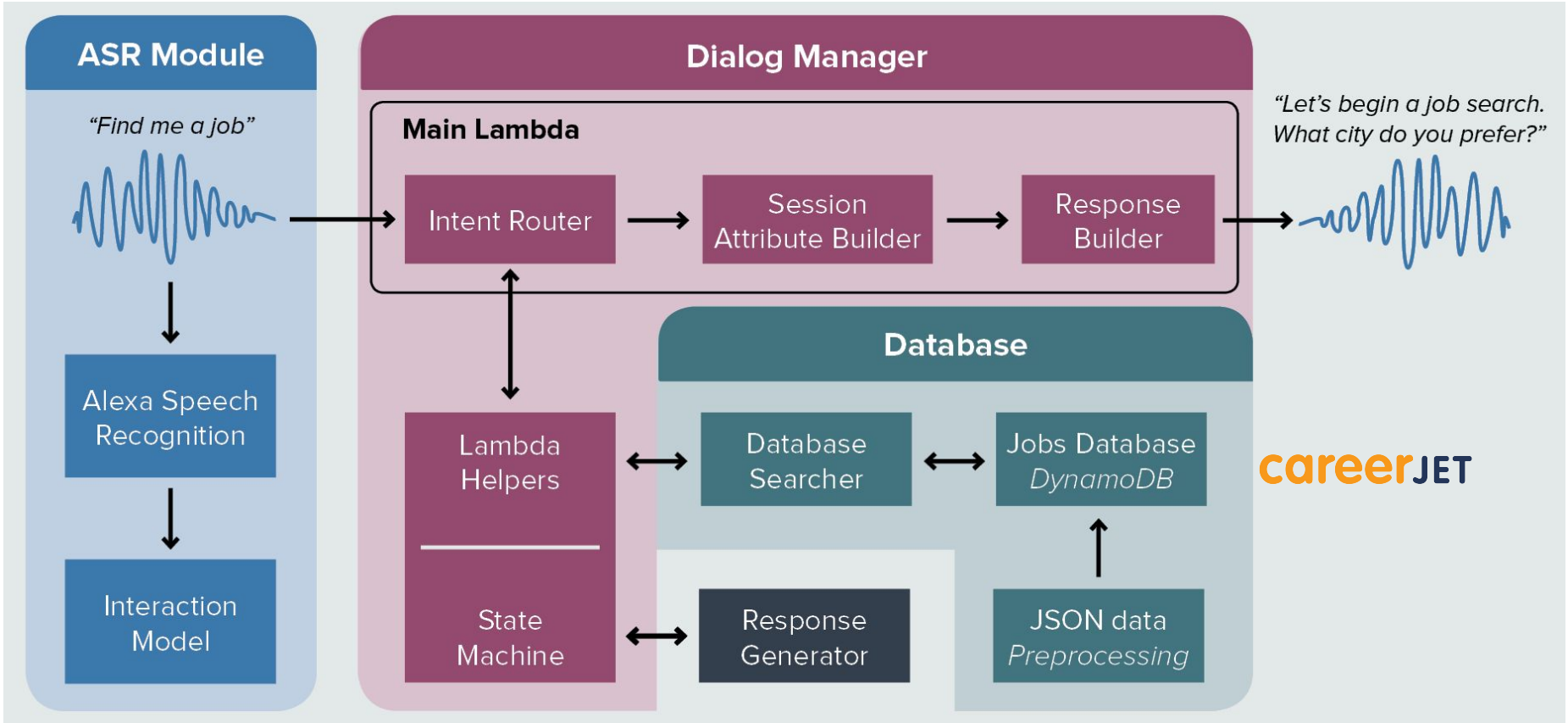
- Job Data challenging (yet tractable)
- Talk the user through a job search
- Develop a conversational system

Target Users: New Job Seekers, people changing careers or job seekers who are new to the industry

Team aims:

- Tech Learning Opportunity (AWS)
- Conversational Knowledge Base Navigation
- Context Dependant Response

System Architecture



Dialog Design

Mega-states and their allowed actions:

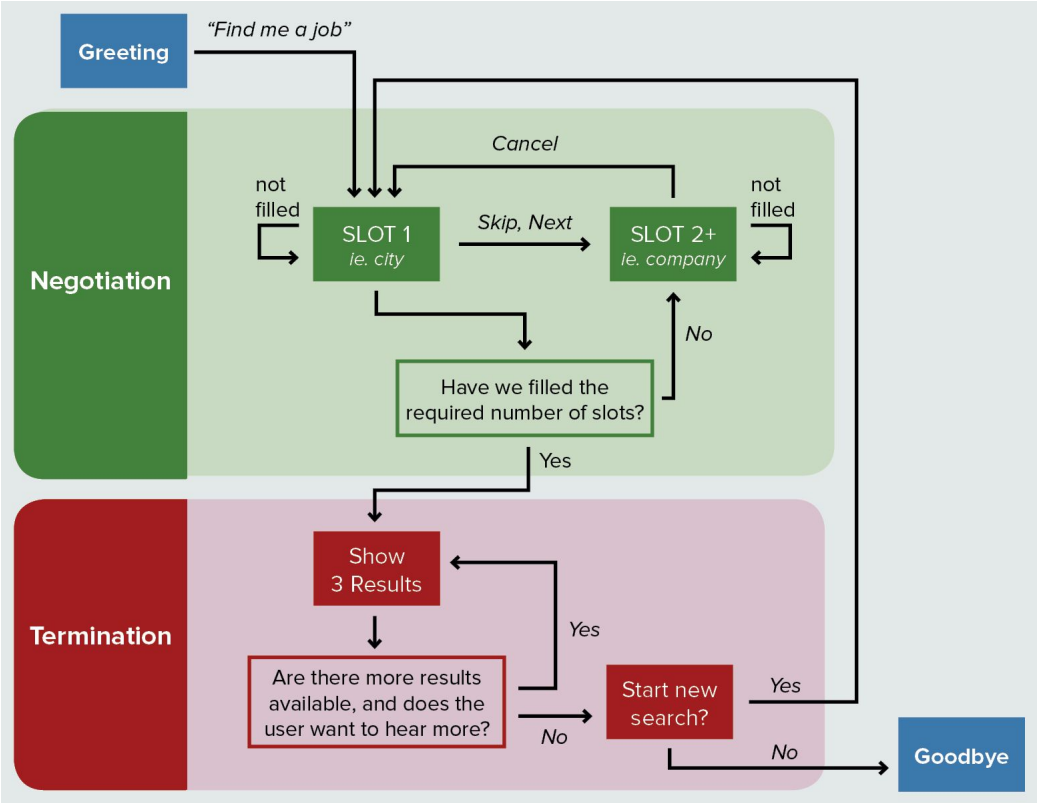
- Greeting - Conversation w/o context slot filling, Help, Confirm and Decline
- Negotiation - Conversation with slot filling, Confirm, Decline, Backup, Reset
- Termination - Conversation with slot filling, Confirm, Decline, Backup, Reset

Action Description:

  - Confirm: <yes words/sentences>
  - Decline: <no words/sentences>
  - Backup: Skip/Next and Confirm/Decline
  - Reset(Cancel): Reverts the conversation state to the filling the first context slot

Dialog driver: data-defined context slots

State Machine



Job Knowledge Base

- Small data set for prototyping:
  400 jobs (20 industries x 20 jobs)
- Rich for dialog experimenting
- Modelled as Limited Domains

```
SLOT_NAMES= [
    "city",
    "company",
    "title",
    "state",
    "skills"
]
```

```
SLOT_STATES = [
    'EMPTY',
    'FILLED',
    'IN_PROGRESS',
    'PARTIALLY_FILLED'
]
```

jobpost	date	Title	Company	Announceme	Term	Eligibility	Audience	StartDate	Duration	Location	JobDescripti	JobRequirme	RequiredQua	Salary	ApplicationP
AMERIA	5-Jan-04	Chief Financi	AMERIA Inve	NA	NA	NA	NA	NA	NA	Yerevan, Arn	AMERIA	- Supervises	To perform	NA	To apply for
Internationa	7-Jan-04	Full-time Cor	International	NA	NA	NA	NA	NA	3 months	IREX	NA	NA	Bachelor's	NA	Please
Caucasus	7-Jan-04	Country Coo	Caucasus En	NA	NA	NA	NA	NA	Renewable	Yerevan, Arn	Public	- Working			
Manoff	7-Jan-04	BCC Specialis	Manoff Grou	NA	NA	NA	NA	NA	NA	Manila, Phili	The LEAD	- Identify			
Yerevan	10-Jan-04	Software De	Yerevan Brar	NA	NA	NA	NA	NA	NA	Yerevan, Arn	NA	- Renderi			
Boutique	10-Jan-04	Saleswoman	Boutique "Ag	NA	NA	NA	NA	NA	NA	Yerevan,	Saleswoman	NA			
OSI	11-Jan-04	Chief Accour	OSI Assistanc	NA	NA	NA	NA	NA	NA	Yerevan, Arn	The	NA			
Internationa	13-Jan-04	Non-paid par	International	NA	NA	NA	NA	NA	6 months	IREX	NA	NA			
Yerevan	13-Jan-04	Assistant to I	Yerevan Brar	NA	NA	NA	NA	NA	NA	Yerevan, Arn	NA	-			
American	13-Jan-04	Program Ass	American	NA	NA	NA	NA	NA	NA	NA	The	NA			
Internationa	13-Jan-04	Short-Term T	International	NA	NA	NA	NA	NA	NA	NA	NA	NA			
Internationa	13-Jan-04	Non-paid par	International	NA	NA	NA	NA	NA	6 months	IREX	NA	NA			
Institute for	13-Jan-04	Chief of Part	Institute for	NA	NA	NA	NA	NA	5 year	Tashkent, Uz	ISC seeks an	NA			
Food	14-Jan-04	Community I	Food Securit	NA	NA	NA	NA	NA	NA	Ijevan town,	Food	- Assist th			
Teleplus LLC	14-Jan-04	General Man	Teleplus LLC	NA	NA	NA	NA	NA	NA	Yerevan, Arn	NA	- Manage			
NetCall	15-Jan-04	Network Adr	NetCall Com	NA	NA	NA	NA	NA	NA	Yerevan, Arn	NA	- Networ			
SOC.Stockh	15-Jan-04	Utopian Wor	SOC.Stockho	NA	NA	NA	NA	NA	NA	NA	NA	NA			
United	15-Jan-04	Country Ecor	United Natio	NA	NA	NA	NA	NA	3 months init	NA	The United	The			
Counterpart	16-Jan-04	Driver/ Logis	Counterpart	NA	NA	NA	NA	NA	NA	Yerevan, Arn	Driver/	- Maintai			
Xalt LLC	16-Jan-04	Graphic Desi	Xalt LLC	NA	NA	NA	NA	NA	NA	Yerevan, Arn	The position	Graphic			

```
query={
    'location'   : '',
    'keywords'   : 'accounting',
    'affid'      : '36e406480df97
    'user_ip'    : '11.22.33.44',
    'url'        : 'http://www.exa
    'user_agent' : 'Mozilla/5.0 (
}
```

Data Processing

Current scope of system:

- Clustering: simple grouping by domain
- String normalization: **e, e, è, é, ê, ë**
- Data Standardization: analysis driven towards matching user utterance
  punctuation, abbreviations: mostly unchanged
  written: **St Louis**
  spoken: **Saint Louis**
- Search: using exact token matching

3_gram	freq	%ge	cum%
salt_lake_city	4	33.33%	33.33%
augusta_richmond_county	3	25.00%	58.33%
port_saint_lucie	3	25.00%	83.33%
louisville_jefferson_county	1	8.33%	91.67%
west_palm_beach	1	8.33%	100.00%

Future scope:

Reverse indices, Named-entity resolution, Entity linking/resolution, typo-fixing

Context-based Search Example

"Top skills": Context Driven

context = {}		context = {city=California}	
'risk management'	20	'policy development'	4
'key account management'	19	'budget oversight'	3
'p l management'	19	'corporate security'	3
'project planning'	19	'cost accounting'	3

Sample dialog

A end-to-end dialog sample shown here

Language Generation

Rule-based Response Generator: directs to apt response based on the turn goal

```
ok_words = ['Ok', 'Okay', 'Cool', 'Alright', 'Great', 'Alright then', 'Sounds good

prompt_suffixes = ['are you interested in?', 'interests you?',
                   'are you considering?'  , 'do you prefer?',
                   'are you thinking of?'  , 'appeals to you?']

start_prompts = ['Let\'s start a job search.', 'You are starting a job search.',
                 'Let\'s begin a job search.', 'Let\'s begin searching for a job.'

reprompt_suffixes = ['you\'d like to include in the search.',
                     'you would like to include in the search.',
                     'you want to include in the search.']

thanks_fors = ['Thanks for', 'Hope you enjoyed', 'Thank you for']

goodbyes    = ['Goodbye.', 'See you later.', 'See you around.',
               'Until next time.', 'Bye for now.', 'Goodbye for now.']

proposals   = ['Would you like', 'Do you want', 'Do you wish',
               'Would you be interested', 'Want']

restarts    = ['would you like to search for more jobs?', 'would you like to searc
               'do you want to start another search?', 'want to search for a diffe
```

Observations

- Alexa intent recognition algorithm is heavily dependent upon:
  - Number of words in the utterances
  - The utterance vocabulary
  - Number of samples in each slot
- Alexa ASR **does not**: recognize sentence boundaries, perform well with named entity words that are commonly referred to with abbreviations, i.e., AT&T.
- Built-in intents handle disfluencies well when supplemented with sufficient sample utterance data, but do not handle backup and corrections, i.e., “I want to work in Portland uh I mean Seattle”
- A slot-filling approach is a naive but effective approach to building an information retrieval chatbot
- Wake words such as ‘Alexa’ cannot be recognized in any utterance

Suggested Evaluation

Naive approach:

- % of search restarts
- % of turns per slot
- % of session “quits” without entering Termination state

Future Scope: Formal methods such as the PARADISE framework

DEMO

Takeaways

- Alexa ASR is robust and easy to work with
- Dialog model could be made more flexible
- Decouple NLU from interaction model
- Data model has great potential for research

Future Scope

Dialog Policy: transform system to mixed-initiative system; incorporate Markov processes to maintain information state.

NLU Robustness: allow user to fill multiple context slots in a single utterance; map user preferences using context slot clustering in database; use RNNs to identify user utterances and map them to intents.

NLG Improvement: change current rule-based system to a neural sequence model.

Entity Resolution: disambiguate between named entities.

Additional Features: allow user to change existing context slot values at any point; send email to user containing job application links

Thanks for Listening!

Q & A

Team HireMe

Alvin Cao	EE
Ayushi Aggarwal	CL
Henry Lu	EE
John Greve	CL
Kevin Hsu	EE
B. Kevin Ramada	EE