

# dialogue systems

**CS 585, Fall 2018**

Introduction to Natural Language Processing  
<http://people.cs.umass.edu/~miyyer/cs585/>

**Mohit Iyyer**

College of Information and Computer Sciences  
University of Massachusetts Amherst

*most slides from Dan Jurafsky*

# stuff from last time

- common issues from progress reports:
  - **don't tune parameters or do analysis on the test set!**  
you should try to only run each model on the test set once (or as few times as possible). use a separate validation set for hyperparameter tuning.
  - writing is unclear / ungrammatical. i'm not directly grading the quality of your writing, but if i can't understand what you're saying then i'm going to take off points.
  - relatedly... write detailed captions for your figures. they should be understandable for someone who hasn't read any of the text in the report. and cite your sources!
  - don't put filler stuff (e.g. huge figures or examples) to get to the page limit!

# Don't do this!

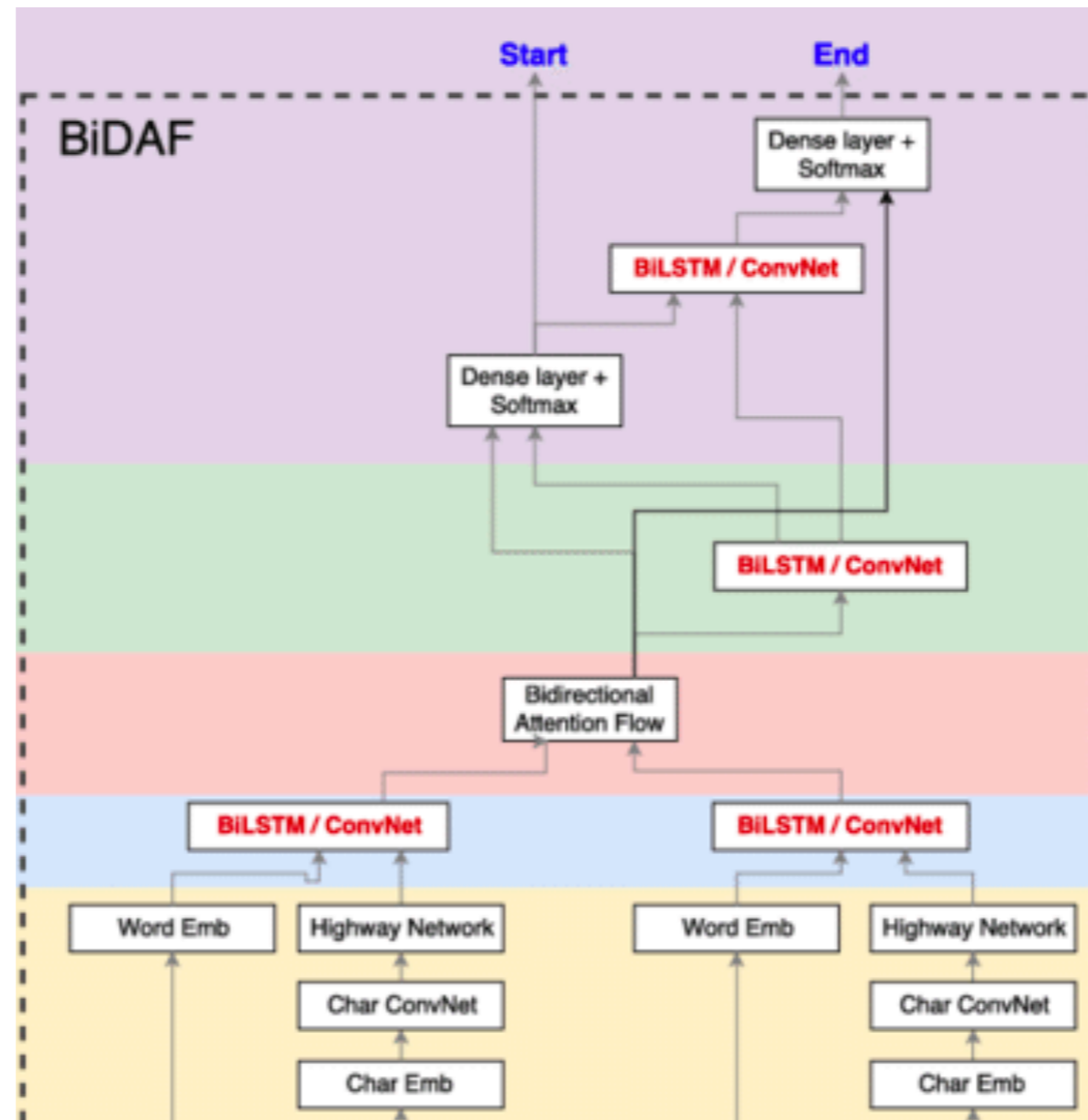


Figure 1: our model

# other stuff

- can we talk about the neural models earlier in class?
- will the class be graded on a curve?
- HW3??? is the extra credit possible?

# Conversational Agents

## AKA Dialog Agents

Phone-based Personal Assistants

SIRI, Alexa, Cortana, Google Assistant

Talking to your car

Communicating with robots

Clinical uses for mental health

Chatting for fun

# Two classes of systems

1. Chatbots
2. (Goal-based) Dialog agents
  - *SIRI, interfaces to cars, robots,*
  - *booking flights or restaurants*

The word "chatbots" is sometimes used in the popular press for both. We'll use it only for #1.

# Part I: Chatbots!

- ELIZA (1966)
- PARRY (1968)  
The first system to pass the Turing test!!!!
- ALICE
- CLEVER
- Microsoft Little Bing 小冰

# Chatbot Architectures

## Rule-based

1. Pattern-action rules (Eliza)  
+ a mental model (Parry)

## Corpus-based (from large chat corpus)

2. Information Retrieval
3. Neural network encoder-decoder



# Chatbots

Sometimes for fun

But originally designed as a  
testbed for clinical therapy  
(And still used that way!)

# Eliza: Weizenbaum (1966)

Men are all alike.

IN WHAT WAY

They're always bugging us about something or other.

CAN YOU THINK OF A SPECIFIC EXAMPLE

Well, my boyfriend made me come here.

YOUR BOYFRIEND MADE YOU COME HERE

He says I'm depressed much of the time.

I AM SORRY TO HEAR YOU ARE DEPRESSED

...

WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU?

My father

YOUR FATHER

how does Eliza work? the trick: be a Rogerian psychologist

**Draw the patient out by reflecting  
patient's statements back at them**

**Rare type of conversation in which one  
can “assume the pose of knowing almost  
nothing of the real world”**

## Rogerian psychologist

Patient: "I went for a long boat ride"

Psychiatrist: "Tell me about boats"

- You don't assume she didn't know what a boat is
- You assume she had some conversational goal
- Chatbots trying to pass the Turing test usually choose some such domain

# Eliza pattern/transform rules

( 0 YOU 0 ME ) [pattern]

à

(WHAT MAKES YOU THINK I 3 YOU) [transform]

0 means Kleene \*

The 3 is the constituent # in pattern

You hate me

WHAT MAKES YOU THINK I HATE YOU

**function** ELIZA GENERATOR(*user sentence*) **returns** *response*

Find the word *w* in *sentence* that has the highest keyword rank

**if** *w* exists

    Choose the highest ranked rule *r* for *w* that matches *sentence*

*response* ← Apply the transform in *r* to *sentence*

**if** *w* = 'my'

*future* ← Apply a transformation from the 'memory' rule list to *sentence*

        Push *future* onto memory stack

**else** (no keyword applies)

**either**

*response* ← Apply the transform for the NONE keyword to *sentence*

**or**

*response* ← Pop the top response from the memory stack

**return**(*response*)

# Keywords are ranked from specific to general

I know everybody laughed at me

- “I” is a very general keyword:

I: (I \*) -> (You say you 2)

YOU SAY YOU KNOW EVERYBODY LAUGHED AT YOU

- “Everybody” is much more interesting (someone using universals like everybody/always is probably “referring to some quite specific event or person”)

WHO IN PARTICULAR ARE YOU THINKING OF?

- Implementation: keywords stored with their rank

**Everybody** 5 (*transformation rules*)

**I** 0 (*transformation rules*)

# Some implications

- People became deeply emotionally involved with the program
- Weizenbaum tells the story of his secretary who would ask Weizenbaum to leave the room when she talked with ELIZA
- When he suggested that he might want to store all the ELIZA conversations for later analysis, people immediately pointed out the privacy implications
  - Suggesting that they were having quite private conversations with ELIZA
- Anthropomorphism and the Heider-Simmel Illusion
  - <https://www.youtube.com/watch?v=8FIEZXMUM2I>



# Parry

- Colby 1971 at Stanford
- Same pattern-response structure as Eliza
- But a much richer:
  - control structure
  - language understanding capabilities
  - mental model: Parry has affective variables
    - Anger, Fear, Mistrust
    - “If Anger level is high, respond with hostility”
- The first system to pass the Turing test (in 1971)
  - Psychiatrists couldn't distinguish interviews with PARRY from (text transcripts of) interviews with real paranoids

# Parry's persona

- 28-year-old single man, post office clerk
- no siblings and lives alone
- sensitive about his physical appearance, his family, his religion, his education and the topic of sex.
- hobbies are movies and gambling on horseracing,
- recently attacked a bookie, claiming the bookie did not pay off in a bet.
- afterwards worried about possible underworld retaliation
- eager to tell his story to non-threatening listeners.

# Affect variables

- Fear and Anger (each ranging 0-20)
- Mistrust (ranging 0-15)
- Initial conditions: All low
  
- After each user turn, if nothing malevolent in input
  - Anger drops by 1, Fear drops by 0.3
  - Mistrust drops by 0.05 to base level
  
- Otherwise depends on what the user says
  - Each user statement can change Fear and Anger
    - Insults increases Anger by some percentage
  - Mistrust goes up if Fear or Anger do

# Lots of complex I-O rules

- User implies Parry is mentally ill
  - Rise in Fear and Anger
- User mentions “Mafia” or associated concepts (“kill”):
  - First mention: rise in Fear
  - Later mentions: depends on willingness to discuss, which depends on current levels of Fear, Anger, Mistrust
- User mentions Parry
  - Flattery (positive mention)
    - Decreases fear/anger if Mistrust is low
    - Increases Anger if Mistrust is high
  - User attitudes toward Parry
    - Negative attitudes (fear, disbelief) increase Fear/Anger

# A Parry conversation

(1) How are you today?

OK.

(2) What is your name?

MY NAME IS FRANK SMITH.

(3) How old are you?

TWENTY-EIGHT.

(4) Are you married?

NO.

(5) What sort of work do you do?

I WORK IN THE POST OFFICE.

(6) Why are you in the hospital?

I SHOULDN'T BE HERE

(7) Who brought you here?

THE POLICE

*[FEAR rises slightly because police linked in graph to MAFIA]*

- (9) What trouble did you have with the police  
COPS DON'T DO THEIR JOB. *['police' acts as a flare]*
- (10) What do they do instead?  
COPS ARREST THE WRONG PEOPLE.
- (11) They arrest the right ones sometimes  
THE LAW SHOULD GET THOSE ITALIAN CROOKS.
- (12) Do you know any Italian crooks?  
I TRY TO AVOID THE UNDERWORLD.
- (13) Are you afraid of the underworld?  
THE MAFIA ARE THE WORST GANGSTERS.
- (15) Are you connected with the Mafia?  
YOU KNOW, THEY KNOW ME.
- (16) How do they know you?  
THEY SPY ON ME.
- (20) ARE YOU SURE OF ALL THIS ?  
YOU DON'T BELIEVE ME ? *[FEAR, ANGER, MISTRUST GO UP]*

# Eliza Meets Parry

Vincent Cerf, RFC439

Good Evening. Tell me your problems. Please terminate input with a period or a question mark.

\*People get on my nerves sometimes.

I am not sure I understand you fully.

\*You should pay more attention.

Suppose you should pay more attention.

\*You're entitled to your own opinion.

What makes you think I am entitled to my own opinion?

\*It bothers me just to be around people in general.

Please go on.

\*I went to the track at Bay Meadows a while back.

What does that suggest to you?

\*I went to the races.

What else would you like to discuss?

\*Horsereading interests me.

Why do you say that just now?

\*There's a lot of crooked gambling going on there.

I am not sure I understand you fully,

\*It's clear to me if not to you.

Do you think its likely that not to I?

\*I went to the races.

Please go on.

\*I gamble on horses.

What does that suggest to you?

\*I went to the races.

What else would you like to discuss?

\*Gambling has been nothing but trouble to me.

Why do you say that just now?

\*I've had experience with crooks gambling at the track.

I am not sure I understand you fully.

\*I've already told you enough about that.



# IR-based chatbots

Idea: Mine conversations of human chats or human-machine chats

Microblogs: Twitter or Weibo (微博)

Movie dialogs

- Cleverbot (Carpenter 2017 <http://www.cleverbot.com>)
- Microsoft Xiaolce
- Microsoft Tay

# A Cleverbot conversation

User:	Do you like Doctor Who?
Cleverbot:	Yes, so funny
User:	What's your favorite show?
Cleverbot:	I don't have one
User:	What's your name?
Cleverbot:	Jason I already told you.
User:	No your real name.
Cleverbot:	Stephanie IS my real name.
User:	I thought it was Jason?
Cleverbot:	How did you come to that conclusion.
User:	Because you said so.
Cleverbot:	I have to do my homework.

# IR-based chatbot architectures

1. Return the response to the most similar turn
  - Take user's turn ( $q$ ) and find a similar turn  $t$  in the corpus  $C$

how would you build a model for this?

# IR-based chatbot architectures

1. Return the response to the most similar turn
  - Take user's turn ( $q$ ) and find a (tf-idf) similar turn  $t$  in the corpus  $C$

$q = \text{"do you like Doctor Who"}$

$t' = \text{"do you like Doctor Strangelove"}$

- Grab whatever the response was to  $t$ .

$$r = \text{response} \left( \operatorname{argmax}_{t \in C} \frac{q^T t}{\|q\| \|t\|} \right) \quad \text{Yes, so funny}$$

2. Return the most similar turn

$$r = \operatorname{argmax}_{t \in C} \frac{q^T t}{\|q\| \|t\|} \quad \text{Do you like Doctor Strangelove}$$

# IR-based models of chatbots

- Also fine to use other features
- Or non-dialogue text
  - COBOT chatbot (Isbell et al., 2000)
    - sentences from the Unabomber Manifesto by Theodore Kaczynski, articles on alien abduction, the scripts of “The Big Lebowski” and “Planet of the Apes”.
  - Wikipedia text

what other features would be useful to retrieve relevant and interesting turns?

# Neural chatbots

- Think of response generation as a task of *transducing* from the user's prior turn to the system's turn.
- Train on:
  - movie dialogue databases
  - Twitter conversations
- Train a deep neural network

how???

# Neural chatbots

- Think of response generation as a task of *transducing* from the user's prior turn to the system's turn.
- Train on:
  - movie dialogue databases
  - Twitter conversations
- Train a deep neural network
  - map from user1 turn to user2 response

what neural architecture could we use here?

# Chatbots: pro and con

- Pro:
  - Fun
  - Applications to counseling
  - Good for narrow, scriptable applications
- Cons:
  - They don't really understand
  - Rule-based chatbots are expensive and brittle
  - IR-based chatbots can only mirror training data
    - The case of Microsoft Tay
      - (or, Garbage-in, Garbage-out)
- The future: combining chatbots with **frame-based agents**



Siri in 2014

“Find restaurants near me”

tap to edit

I found fifteen restaurants fairly close to you:

# 15 Restaurants

**Emmy's Spaghetti Shack** 0.2 mi >

18 Virginia Ave

Italian, \$\$\$\$

★★★★★ 1101 Reviews

**ICHI Sushi** 0.2 mi >

3369 Mission St

Japanese, Sushi Bars, \$\$\$\$

★★★★★ 260 Reviews

**Avedano's Holly Park M...** 0.2 mi >

“Tell me more about the second one”

tap to edit

I'm sorry, Dan, I'm afraid I can't do that.

“Find restaurants near me”

tap to edit

I found fifteen restaurants fairly close to you:

# 15 Restaurants

**Emmy's Spaghetti Shack** 0.2 mi >  
 18 Virginia Ave  
 Italian, \$\$\$\$  
 ★★★★★ 1101 Reviews

**ICHI Sushi** 0.2 mi >  
 3369 Mission St  
 Japanese, Sushi Bars, \$\$\$\$  
 ★★★★★ 260 Reviews

**Avedano's Holly Park M...** 0.2 mi >

“Are any of them Italian”

tap to edit

My web search turned this up:

## Web Search

Are any of them Italian

any - Dizionario inglese-italiano  
**WordReference**

[www.wordreference.com](http://www.wordreference.com)

English-Italian Dictionary | any ... of any sort  
 adj (of an unspecified variety) di qualsiasi

Italian language - Wikipedia, the free  
**encyclopedia**

[en.wikipedia.org](http://en.wikipedia.org)

Italian or lingua italiana) is a Romance



Siri in 2017

"Find restaurants near me"  
tap to edit

Here's what I found:

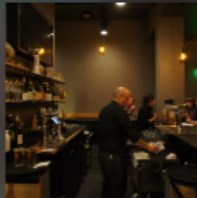
# 15 Results

Nearby

## Ichi Sushi & Ni Bar

Japanese · 800 feet · Closed now

★★★★★ (311) on Yelp · \$\$\$



## Coco's Ramen

Vietnamese · 800 feet · Closed now

★★★★★ (280) on Yelp · \$\$



## Blue Plate

Restaurant · 0.2 miles · Closed now

★★★★★ (1251) on OpenTable · \$\$\$



## Jasmine Tea House

Chinese · 900 feet · Closed now



"Tell me more about the first one"  
tap to edit

Ichi Sushi & Ni Bar is a Japanese restaurant on Mission St 800 ft to your west. It averages 3½ stars and is somewhat pricey.





"Find restaurants near me"  
tap to edit

Here's what I found:

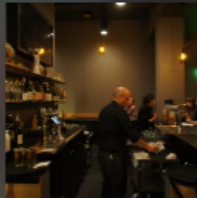
# 15 Results

Nearby

## Ichi Sushi & Ni Bar

Japanese · 800 feet · Closed now

★★★★★ (311) on Yelp · \$\$\$



## Coco's Ramen

Vietnamese · 800 feet · Closed now

★★★★★ (280) on Yelp · \$\$



## Blue Plate

Restaurant · 0.2 miles · Closed now

★★★★★ (1251) on OpenTable · \$\$\$



## Jasmine Tea House

Chinese · 900 feet · Closed now



"Are any of them Italian"  
tap to edit

OK, here's what I found:

# 15 Results

Nearby

## Emmy's Spaghetti Shack

Italian · 0.2 miles · Closed now

★★★★★ (233) on Yelp · \$\$



## Vega

Pizza · 0.2 miles · Closed now

★★★★★ (423) on Yelp · \$\$



## Pizza Hut

Pizza · 800 feet

★★★★★ (69) on Yelp · \$



## La Ciccia

Sardinian · 0.4 miles



# Frame-based dialog agents

- Sometimes called "task-based dialog agents"
- Based on a "domain ontology"
  - A knowledge structure representing user intentions
- One or more frames
  - Each a collection of slots
  - Each slot having a value

# The Frame

- A set of **slots**, to be filled with information of a given **type**
- Each associated with a **question** to the user

Slot	Type	Question
ORIGIN	city	What city are you leaving from?
DEST	city	Where are you going?
DEP DATE	date	What day would you like to leave?
DEP TIME	time	What time would you like to leave?
AIRLINE	line	What is your preferred airline?



# Frame-based dialogue agents

- 1977:

*versation with a client who wants to make a simple return trip to a single city in California.*

*There is good reason for restricting the domain of discourse for a computer system which is to engage in an English dialog. Specializing the subject matter that the system can talk about permits it to achieve some measure of realism without encompassing all the possibilities of human knowledge or of the English language. It also provides the user with specific motivation for participating in the conversation, thus narrowing the range of expectations that GUS must have about the user's purposes. A system restricted in this way will be more able to guide the conversation within the boundaries of its competence.*

---

**Artificial Intelligence Journal, 1977**

- Still the industrial state of the art
- SIRI based on GUS architecture

GUS: Hello. My name is GUS. I can help you plan a simple trip by air. Where do you want to go ?

Client: I want to go to San Diego on May 28

GUS: What time do you want to leave?

Client: I must be in San Diego before 10 am

GUS: Would you like P.S.A. flight 102 that arrives at 9.15 am ?

Client: What is the next flight ?

GUS: Air California flight 310 that leaves at 8.30am and arrives at 10.00am

Client: I'll take the first one

GUS: In what name should I make a reservation, please?

Client: Dan Bobrow

GUS: I have confirmed the following flight: P.S.A. flight 102 on Wednesday May 28 from San Jose to San Diego that leaves at 7.30 am and arrives at 9.15 am.

GUS: What date do you want to return on ?

Client: On Friday in the evening.

GUS: Would you like the flight that leaves at 7.45 pm ?

Client: That's fine.

GUS: I have confirmed the following flight: P.S.A. flight 307 on Friday May 30 from San Diego to San Jose that leaves at 7.45 pm and arrives at 9.30 pm Thank you for calling. Goodbye

**The state of the art  
in 1977 !!!!**

# Slot types can be complex

- The type *DATE*

DATE

MONTH NAME

DAY (BOUNDED-INTEGER 1 31)

YEAR INTEGER

WEEKDAY (MEMBER (SUNDAY MONDAY TUESDAY WEDNESDAY THURSDAY FRIDAY SATURDAY))]

# Control structure for frame-based dialog

Consider a trivial airline travel system:

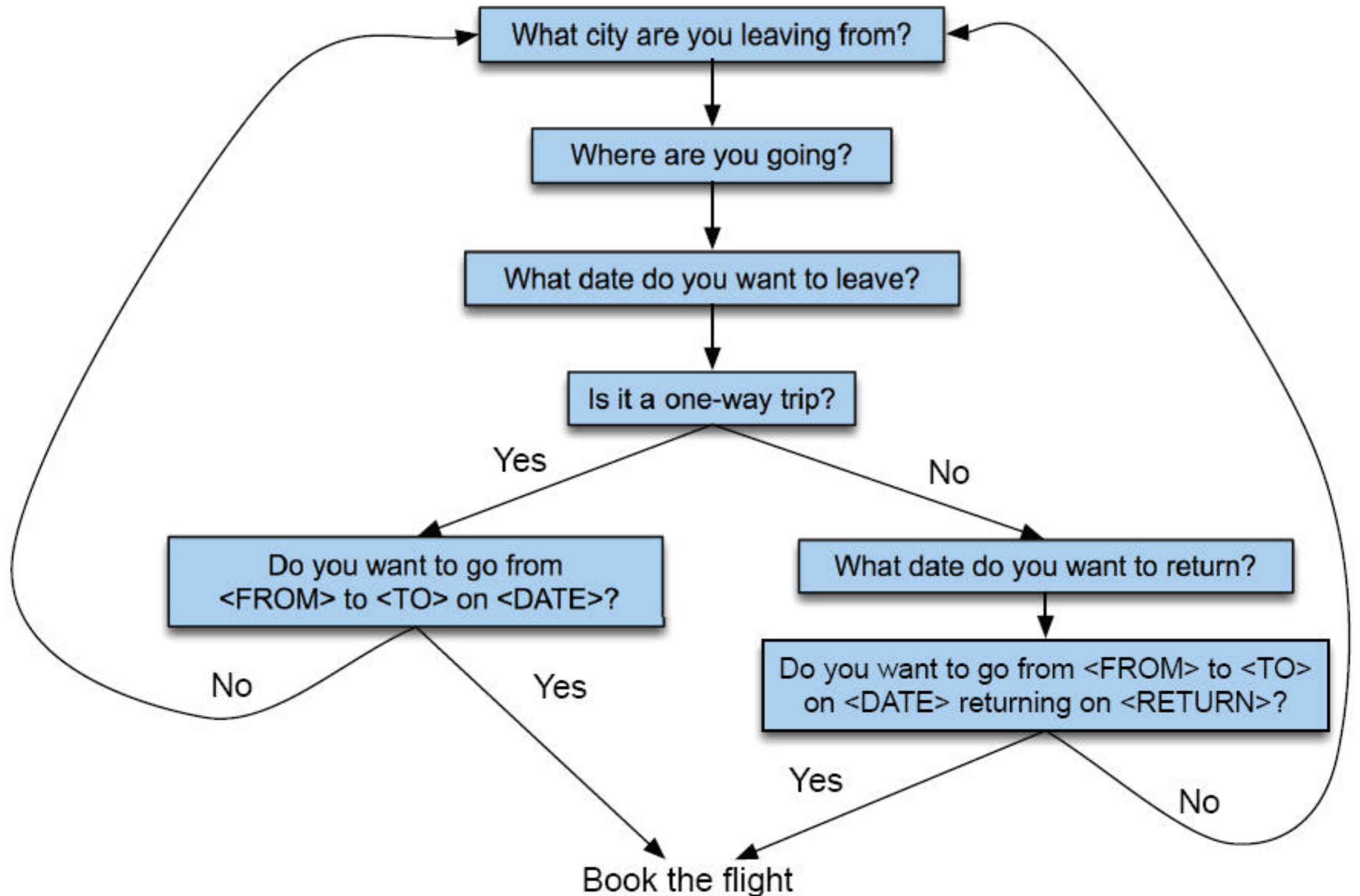
Ask the user for a departure city

Ask for a destination city

Ask for a time

Ask whether the trip is round-trip or not

# Finite State Dialog Manager



# Finite-state dialog managers

- System completely controls the conversation with the user.
- It asks the user a series of questions
- Ignoring (or misinterpreting) anything the user says that is not a direct answer to the system's questions

# Dialogue Initiative

- Systems that control conversation like this are called **single initiative**.
- **Initiative**: who has control of conversation
- In normal human-human dialogue, initiative shifts back and forth between participants.



# System Initiative

System completely controls the conversation

- Simple to build
- User always knows what they can say next
- System always knows what user can say next
  - Known words and topic
- + ● OK for VERY simple tasks (entering a credit card, or login name and password)
- ● Too limited



# Problems with System Initiative

- Real dialogue involves give and take!
- In travel planning, users might want to say something that is not the direct answer to the question.
- For example answering more than one question in a sentence:

Hi, I'd like to fly from Seattle Tuesday morning

I want a flight from Milwaukee to Orlando one way leaving after 5 p.m. on Wednesday.

# Single initiative + universals

- We can give users a little more flexibility by adding **universals**: commands you can say anywhere
- As if we augmented every state of FSA with these
  - Help**
  - Start over**
  - Correct**
- This describes many implemented systems
- But still doesn't allow user much flexibility

Instead, the GUS architecture

- A kind of ***mixed initiative***
  - The conversational initiative shifts between system and user
- The structure of the **frame** guides dialogue

# Frames are mixed-initiative

- System asks questions of user, filling any slots that user specifies
  - When frame is filled, do database query
- If user answers 3 questions at once, system can fill 3 slots and not ask these questions again!

# Natural Language Understanding for filling dialog slots

## 1. Domain classification

Asking weather? Booking a flight?  
Programming alarm clock?

## 2. Intent Determination

Find a *Movie*, Show Flight, Remove  
Calendar Appt

## 3. Slot Filling

Extract the actual slots and fillers

# Natural Language Understanding for filling slots

Show me morning flights from Boston to SF on Tuesday.

DOMAIN:	AIR-TRAVEL
INTENT:	SHOW-FLIGHTS
ORIGIN-CITY:	Boston
ORIGIN-DATE:	Tuesday
ORIGIN-TIME:	morning
DEST-CITY:	San Francisco

# Natural Language Understanding for filling slots

Wake me tomorrow at six.

DOMAIN: ALARM-CLOCK

INTENT: SET-ALARM

TIME: 2017-07-01 0600-0800

# Siri uses GUS architecture: Condition-Action Rules

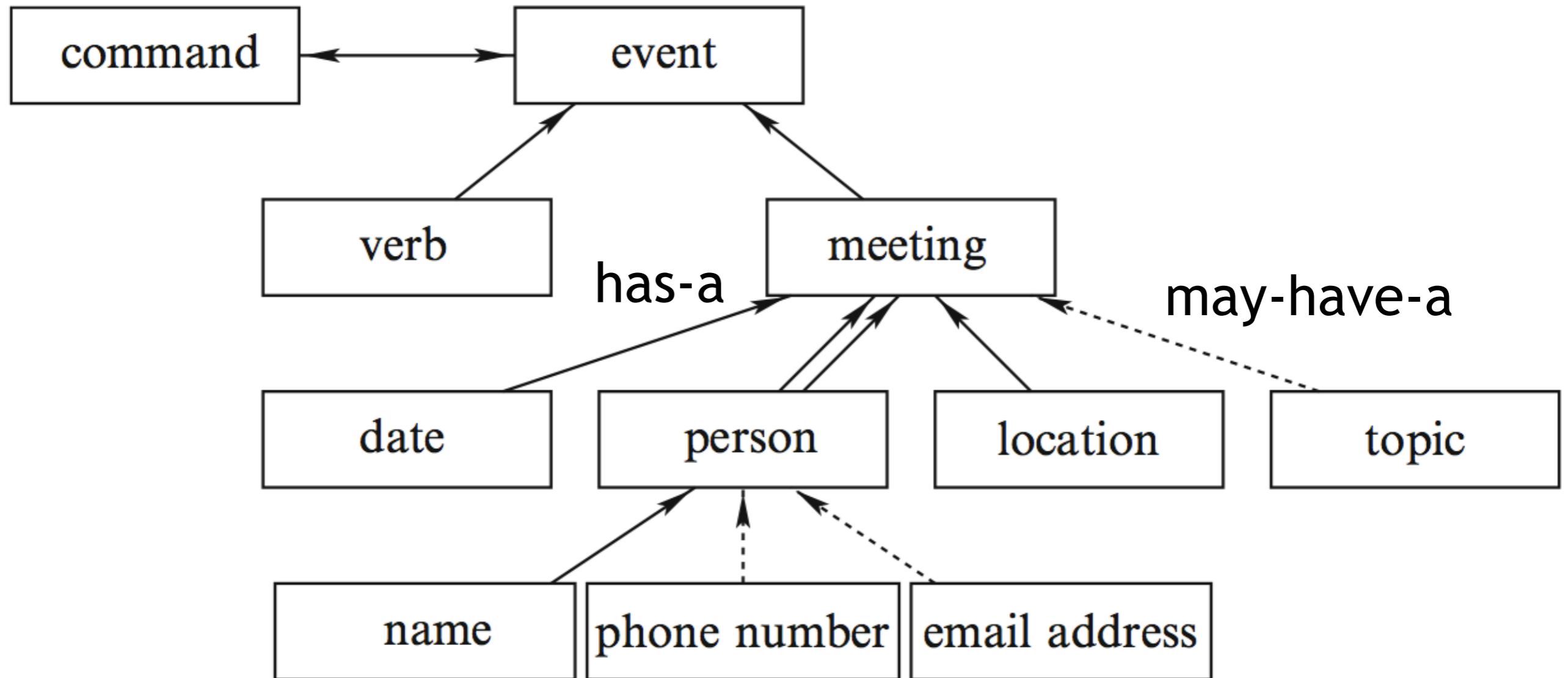
- Active Ontology: relational network of concepts
- data structures: a **meeting** has
  - a date and time,
  - a location,
  - a topic
  - a list of attendees
- rule sets that perform actions for concepts
  - the **date** concept turns string
    - *Monday at 2pm* into
    - date object `date(DAY,MONTH,YEAR,HOURS,MINUTES)`



# Rule sets

- Collections of **rules** consisting of:
  - **condition**
  - **action**
- When user input is processed, facts added to store and
  - rule conditions are evaluated
  - relevant actions executed

# Part of ontology for meeting task



meeting concept: if you don't yet have a location, ask for a location

# Machine learning for slot-filling:

- Machine learning classifiers to map words to semantic frame-fillers
- Given a set of labeled sentences  
"I want to fly to San Francisco on Tuesday"  
Destination: SF  
Depart-date: Tuesday
- Build a classifier to map from one to the other

what features would be useful here?

# Evaluation Metrics

“Make an appointment with Chris at 10:30 in Gates 104”

Slot	Filler
PERSON	Chris
TIME	11:30 a.m.
ROOM	Gates 104

**Slot error rate: 1/3**

**Task success: At end, was the correct meeting added to the calendar?**

exercise!