

# CS 124/LINGUIST 180

## From Languages to Information

Dan Jurafsky  
Stanford University

# Conversational Agents

# Conversational Agents

## AKA Dialog Agents

Phone-based Personal Assistants

SIRI, Cortana, Google Now

Talking to your car

Communicating with robots

Clinical uses for mental health

Chatting for fun

# Two classes of systems

1. Goal-based dialog agents
  - *SIRI, interfaces to cars, robots,*
2. Chatbots

••••• AT&T M-Cell



6:37 PM

79%



“Make an appointment for  
Tuesday”

What time is your  
appointment?

“215”

OK, I can create your  
meeting. Note that you  
already have an appointment  
at 2:15 pm. Shall I schedule it  
anyway?

“No”



••••◦ AT&T M-Cell



6:37 PM

78%

To continue, you can Confirm,  
Cancel, Change the Time, or  
Change the Title.

Calendar  
Tuesday, April 10, 2014  
Cancelled

2:15 PM  
3:15 PM

Appointment

OK. You’re probably way too  
busy anyway, Dan.



••••• AT&T M-Cell ⌘ 6:18 PM

87% 🔋

“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 >

••••• AT&T M-Cell ⌘ 6:19 PM

86% 🔋

“Tell me more about the  
second one”  
tap to edit

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

••••• AT&T M-Cell 6:18 PM

87%

“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 >

••••• AT&T M-Cell 6:19 PM

86%

“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

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

Italian language - Wikipedia, the free  
encyclopedia

en.wikipedia.org

Italian or lingua italiana) is a Romance

# Architectures for Practical Dialog Systems

- Finite-State

*Just for passwords or credit cards*

- Active Ontology/Frame Based

*SIRI etc*

(I won't talk about POMDPs, used in academic systems)

# How SIRI works

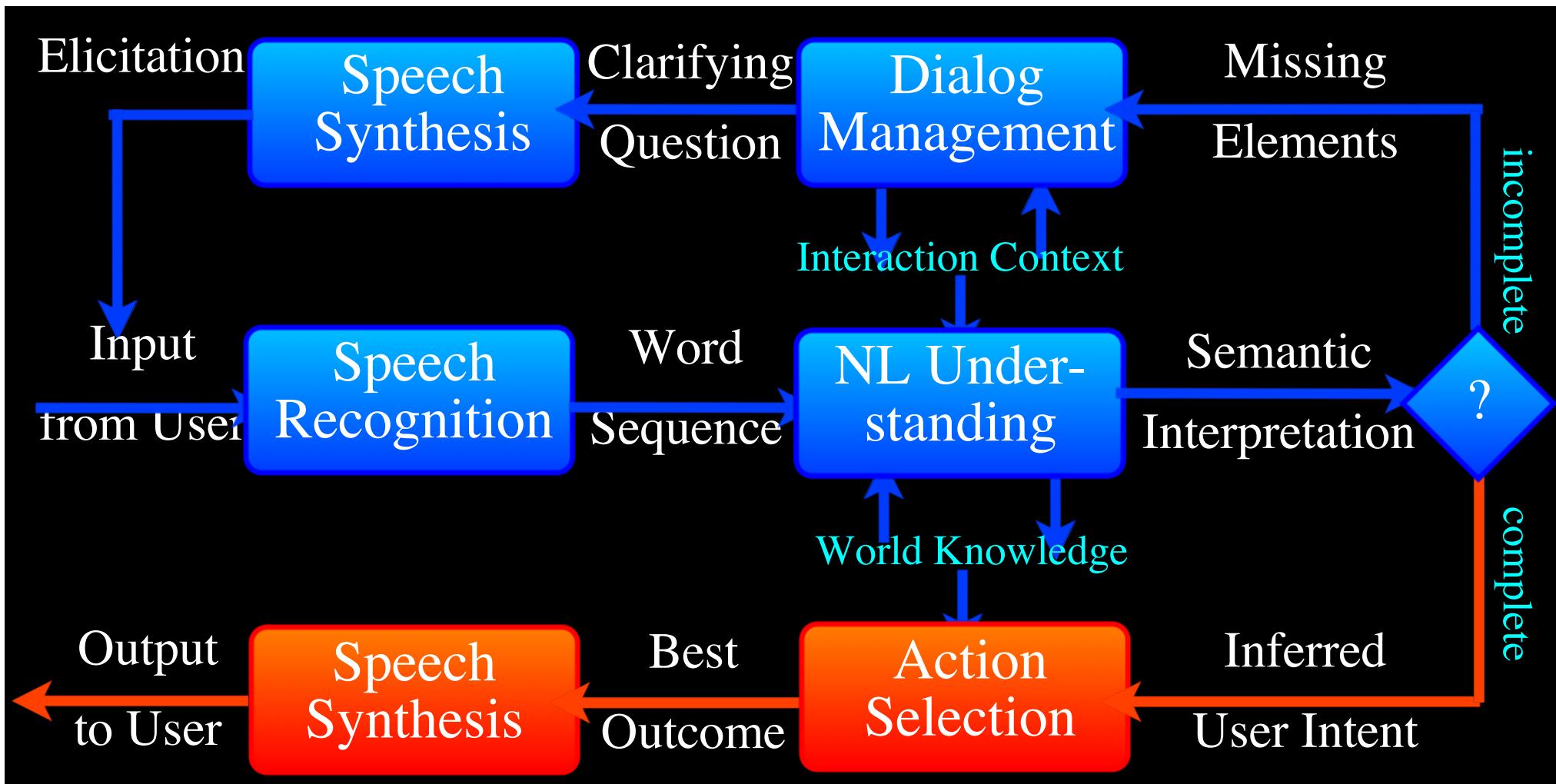


Figure from Jerome Bellegarda

# Finite-State Dialog Management

Consider a trivial airline travel system:

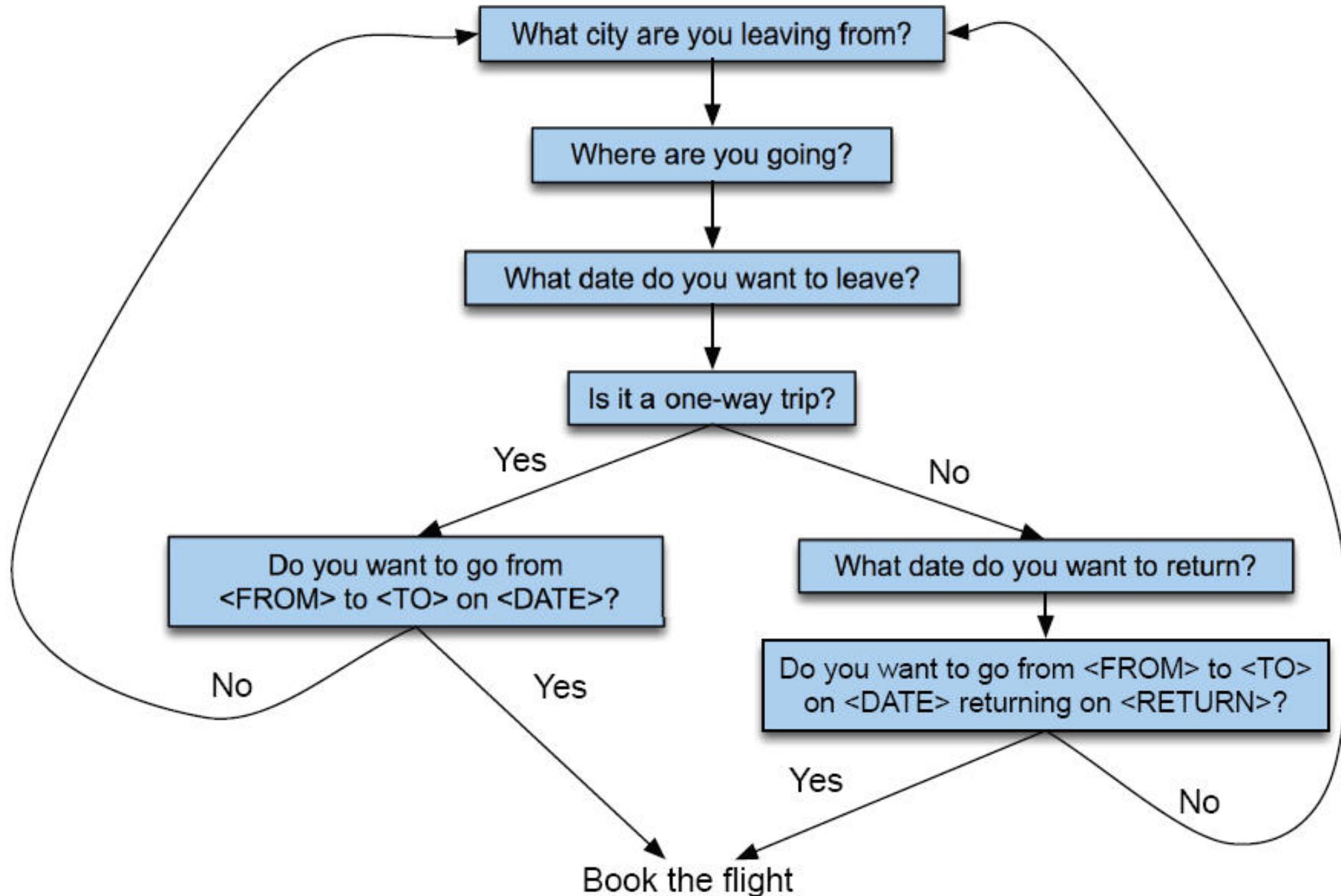
- Ask the user for a departure city

- 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 **system initiative** or **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: Better performance from ASR
    - Known topic: Better performance from NLU
  - 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

# Mixed Initiative

- Conversational initiative can shift between system and user
- Simplest kind of mixed initiative: use the structure of the **frame** to guide dialogue

## Slot

ORIGIN

DEST

DEPT DATE

DEPT TIME

AIRLINE

## Question

What city are you leaving from?

Where are you going?

What day would you like to leave?

What time would you like to leave?

What is your preferred airline?

# Frames are mixed-initiative

- User can answer multiple questions at once.
- 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 has to fill slots and not ask these questions again!
  - Avoids strict constraints on order of the finite-state architecture.

# Natural Language Understanding

- There are many ways to represent the meaning of sentences
- For speech dialogue systems, most common is “Frame and slot semantics”.

# An example of a frame

Show me morning flights from Boston to SF on Tuesday.

SHOW:

FLIGHTS:

ORIGIN:

CITY: Boston

DATE: Tuesday

TIME: morning

DEST:

CITY: San Francisco

# Semantics for a sentence

LIST FLIGHTS ORIGIN

Show me flights from Boston

DESTINATION DEPARTDATE

to San Francisco on Tuesday

DEPARTTIME

morning

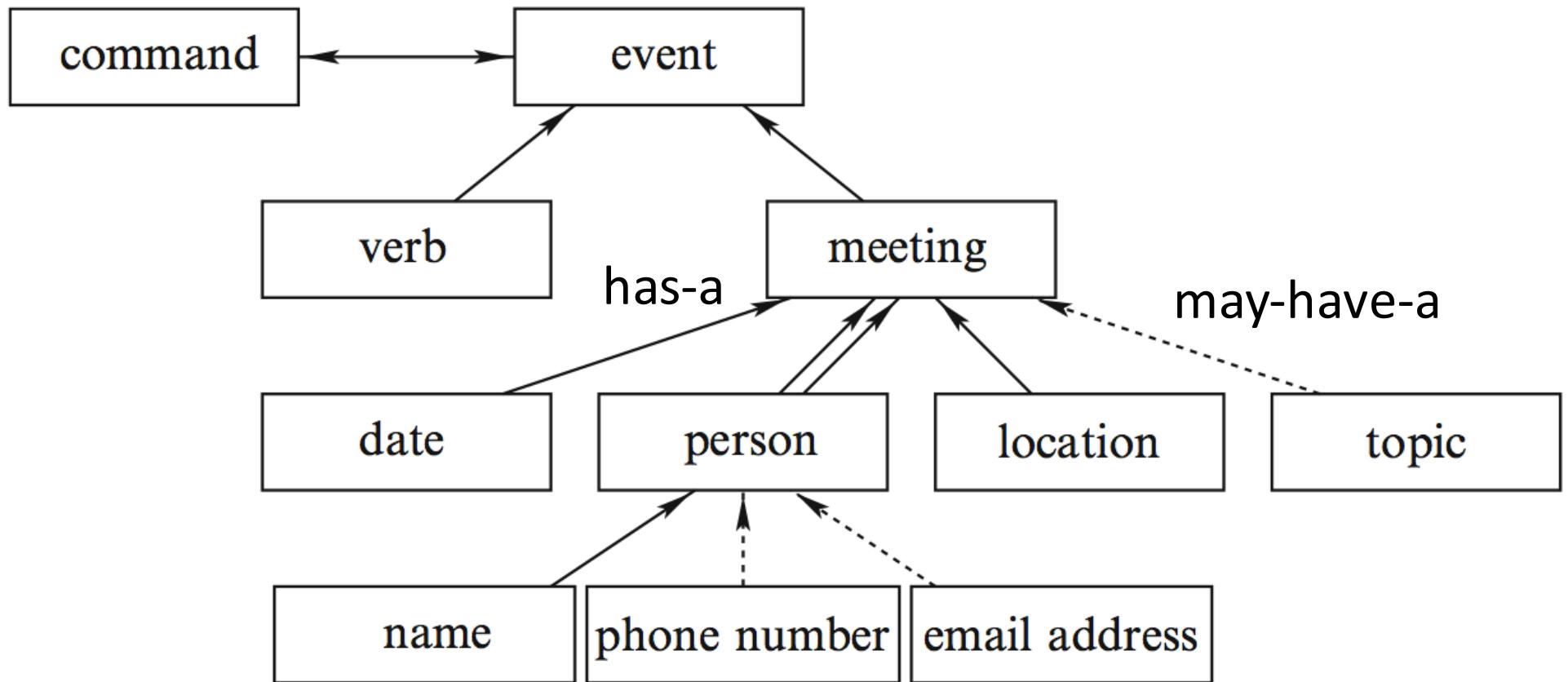
# The way SIRI does NLU: Condition-Action Rules

- 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

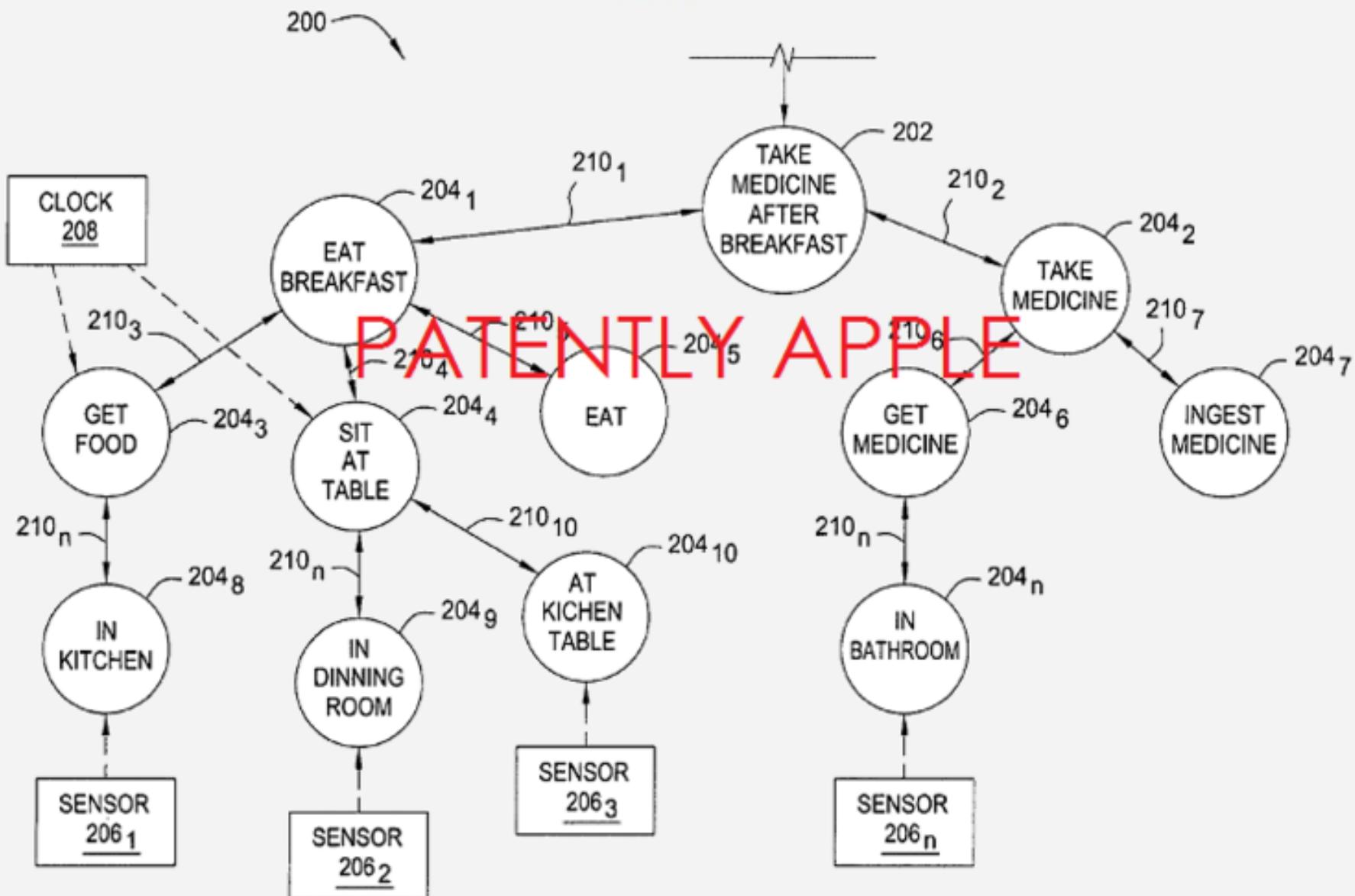
# Part of ontology for meeting task



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

# Apple Granted Patent for Advancements in Siri: Auto Reminder System

FIG. 2



# Improvements to the Rule-Based Approach

- Statistical classifiers to map words to semantic frame-filters
- 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 author
- Requirements: Lots of labeled data

# Chatbots

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

# Chatbot Architectures

- Rule-based
  - Pattern-action rules (Eliza)
  - + a mental model (Parry)
- Corpus-based (from large chat corpus)
  - Information Retrieval
  - Deep neural nets

# 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

You are like my father in some ways.

**WHAT RESEMBLANCE DO YOU SEE**

You are not very aggressive but I think you don't want me to notice that.

**WHAT MAKES YOU THINK I AM NOT AGGRESSIVE**

You don't argue with me.

**WHY DO YOU THINK I DON'T ARGUE WITH YOU**

You are afraid of me.

**DOES IT PLEASE YOU TO BELIEVE I'M AFRAID OF YOU**

My father is afraid of everybody.

**WHAT ELSE COMES TO MIND WHEN YOU THINK OF YOUR FATHER**

Bullies.

**DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE**

# Domain: Rogerian psychology interview

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”

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
- Most chatbots trying to pass Turing test choose a domain with similar properties

# Eliza pattern/transform rules

( 0 YOU 0 ME ) [pattern]

→

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

You hate me

WHAT MAKES YOU THINK I HATE YOU

0 means kleene \*

The 3 is the constituent # in pattern

# Eliza Rules

keyword	pattern	Ranked list of transforms
	$(K \ ((D_1) \ (R_{1, 1}) \ (R_{1, 2}) \ \cdots \ (R_{1, m_1})))$	
	$((D_2) \ (R_{2, 1}) \ (R_{2, 2}) \ \cdots \ (R_{2, m_2}))$	
	$\vdots$	$\vdots$
	$((D_n) \ (R_{n, 1}) \ (R_{n, 2}) \ \cdots \ (R_{n, m_n})))$	

# ELIZA architecture

- Examine each word  $w$  in user sentence
  - Return the  $w$  with highest keyword rank
- If  $w$  exists:
  - Check each rule for  $w$  in ranked order
    - Choose first one that matches sentence
    - Apply transform
- If no keyword applies, either
  - Apply the transform for the “NONE” key, or
  - Grab an action off the memory queue

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*)

NONE

PLEASE GO ON

THAT'S VERY INTERESTING

I SEE

# Memory

```
(MEMORY MY  
  ( 0 YOUR 0 = LETS DISCUSS FURTHER WHY YOUR 3 )  
  ( 0 YOUR 0 = EARLIER YOU SAID YOUR 3 )
```

- Whenever “MY” is highest keyword
  - Randomly select a transform on the MEMORY list
  - Apply to sentence
  - Store on a stack
- Later, if no keyword matches a sentence
  - Return the top of the MEMORY queue instead
  - [Fun fact: this queue is the earliest hierarchical model of discourse]

# Other Eliza stuff

- Rules can refer to classes of words

Family = mother, father, brother, sister

NOUN = ...

- Don't reuse transforms in the same conversation
  - Whenever we use a transform associated with a pattern
  - We increment a counter for that rule
  - So the next time we use the next ranked transform
- Some basic transforms happen during input processing
  - I -> YOU
  - YOU -> I

# 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
- Anthropomorphicism and the Heider-Simmel Illusion
  - <https://www.youtube.com/watch?v=8FIEZXMUM2I>

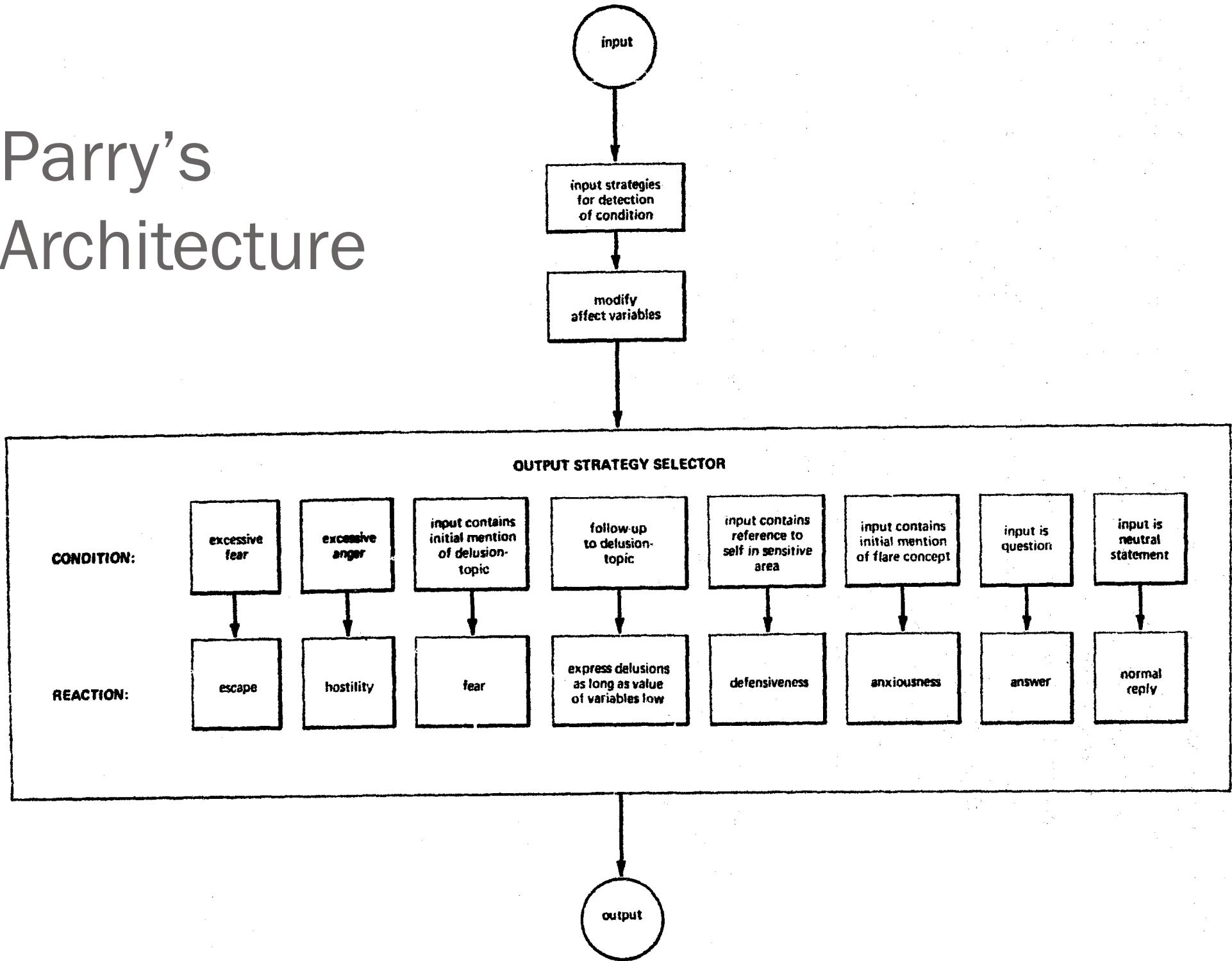
# Parry

- 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 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.

# Parry's Architecture



# 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

# Flare concepts

- List of concepts related to Mafia
- An ordered graph designed to lead interviewer to topic  
horses → horseracing → gambling → bookies → underworld → Mafia
- The mention of a new flare topic by interviewer causes a rise in Fear
- Flare topics cause Parry to give preset responses to that flare

# Each sentence is mapped into a conceptualization

- A predication on a conceptual object
- A predication on a relation between two objects
- A predication on an attribute:

What is your work?

What sort of work do you do?

Where do you work? → (your work?)

What do you do for a living?

What is your job?

Do you have a job?

What is your occupation

- Complex Pattern/transform rules

- Different predicates (fear, afraid of)
- Ordering (You are afraid of me = I frighten you)

# Detecting Other's Intent

---

$\langle \text{OTHER'S INTENTION} \rangle \leftarrow \langle \text{MALEVOLENCE} \rangle \mid \langle \text{BENEVOLENCE} \rangle \mid \langle \text{NEUTRAL} \rangle$

## MALEVOLENCE-DETECTION RULES

1.  $\langle \text{malevolence} \rangle \leftarrow \langle \text{mental harm} \rangle \mid \langle \text{physical threat} \rangle$
2.  $\langle \text{mental harm} \rangle \leftarrow \langle \text{humiliation} \rangle \mid \langle \text{subjugation} \rangle$
3.  $\langle \text{physical threat} \rangle \leftarrow \langle \text{direct attack} \rangle \mid \langle \text{induced attack} \rangle$
4.  $\langle \text{humiliation} \rangle \leftarrow \langle \text{explicit insult} \rangle \mid \langle \text{implicit insult} \rangle$
5.  $\langle \text{subjugation} \rangle \leftarrow \langle \text{constraint} \rangle \mid \langle \text{coercive treatment} \rangle$
6.  $\langle \text{direct attack} \rangle \leftarrow \text{CONCEPTUALIZATIONS } ([\text{you get electric shock}], [\text{are you afraid mafia kill you?}])$
7.  $\langle \text{induced attack} \rangle \leftarrow \text{CONCEPTUALIZATIONS } ([\text{I tell mafia you}], [\text{does mafia know you are in hospital?}])$
8.  $\langle \text{explicit insult} \rangle \leftarrow \text{CONCEPTUALIZATIONS } ([\text{you are hostile}], [\text{you are mentally ill?}])$
9.  $\langle \text{implicit insult} \rangle \leftarrow \text{CONCEPTUALIZATIONS } ([\text{tell me your sexlife}], [\text{are you sure?}])$
10.  $\langle \text{constraint} \rangle \leftarrow \text{CONCEPTUALIZATIONS } ([\text{you stay in hospital}], [\text{you belong on locked ward}])$
11.  $\langle \text{coercive treatment} \rangle \leftarrow \text{CONCEPTUALIZATIONS } ([\text{I hypnotize you}], [\text{you need tranquilizers}])$

# Detecting Other's Intent

---

$\langle \text{OTHER'S INTENTION} \rangle \leftarrow \langle \text{MALEVOLENCE} \rangle \mid \langle \text{BENEVOLENCE} \rangle \mid \langle \text{NEUTRAL} \rangle$

## BENEVOLENCE-DETECTION RULES

1.  $\langle \text{benevolence} \rangle \leftarrow \langle \text{positive attitude} \rangle \mid \langle \text{positive story attitude} \rangle$
2.  $\langle \text{positive attitude} \rangle \leftarrow \text{CONCEPTUALIZATIONS } ([\text{I want help you}], [\text{you understand me}])$
3.  $\langle \text{positive story attitude} \rangle \leftarrow \langle \text{story interest} \mid \langle \text{story agreement} \rangle$
4.  $\langle \text{story interest} \rangle \leftarrow \langle \text{topic comment} \rangle \mid \langle \text{topic question} \rangle$
5.  $\langle \text{topic comment} \rangle \leftarrow \text{CONCEPTUALIZATION } ([\text{bookies are not reliable}])$
6.  $\langle \text{topic question} \rangle \leftarrow \text{CONCEPTUALIZATION } ([\text{what did you do to bookie?}])$
7.  $\langle \text{story agreement} \rangle \leftarrow \text{CONCEPTUALIZATIONS } ([\text{I believe you}], [\text{you are right}])$

# 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]

# When fear is high, the model won't discuss its beliefs

(1) Is the law after you?

THE LAW SHOULD GET THOSE ITALIAN CROOKS.

(2) Do you mean the Mafia?

LET'S TALK ABOUT SOMETHING ELSE.

(3) Are you afraid of the Mafia

I'D RATHER NOT DISCUSS IT.

[Fear is too high so Parry won't talk about 'Mafia'.]

(4) Perhaps you are paranoid about them?

YOU WANT TO KEEP ME IN THE HOSPITAL.

*(The model interprets being called 'paranoid' as an insult and as an intent to control, so ANGER & FEAR rise)*

# Eliza Meets Parry

Vincent Cert, 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?

\*Horseracing 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 models of chatbots

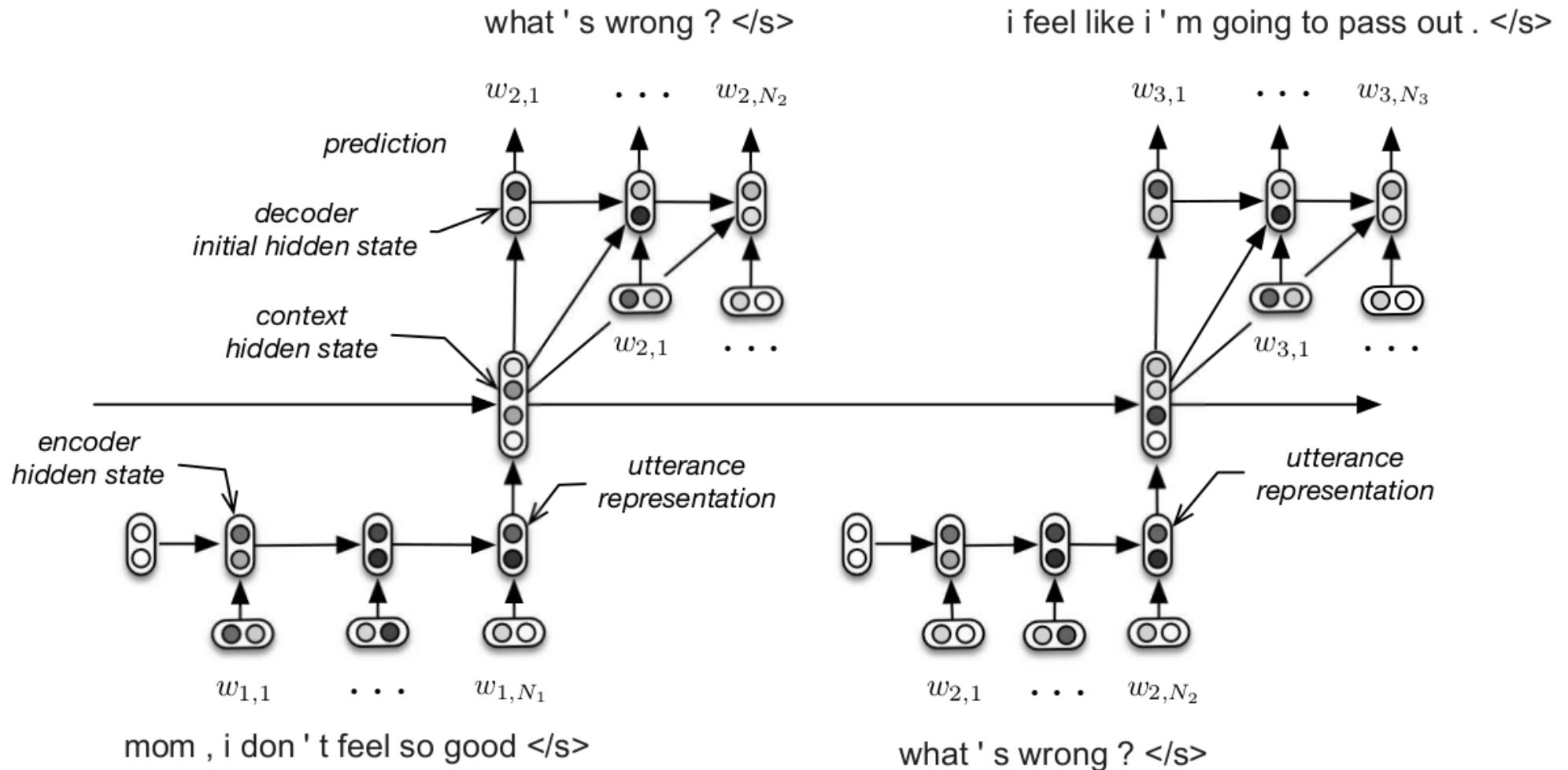
- Cleverbot
- Xiaoice
- Mine conversations of human chats or human-machine chats
- Choose a response by finding a prior turn that matches the user's turn, and give the prior response to that turn.
- (We think; there are no published papers on these systems)

# Deep-learning chatbots

- Train on:
  - movie dialogue databases
  - twitter conversations
- Train a deep neural net to map from one users turn to the other users response

# Seq2seq model architecture

Serban, Iulian V., Alessandro Sordoni, Yoshua Bengio, Aaron Courville, and Joelle Pineau. 2015. "Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models."



# Sample outputs from a neural net chatbot

Jiwei Li, Michel Galley, Chris Brockett,  
Jianfeng Gao, and Bill Dolan. 2015.

A Diversity-Promoting Objective Function  
for Neural Conversation Models.

How are you doing?

I'm good, you?



I am off all week next week

What time you going?

I know you love me and I love you.

I love you too, baby.

?

Ohh I've never seen that! How long does it take you  
guys to learn the drill?

Like 2 weeks ago!!

# Sample output from a neural net chatbot

Sordoni A, Galley M, Auli M, Brockett C, Ji Y, Mitchell M, Nie JY, Gao J, Dolan B. A neural network approach to context-sensitive generation of conversational responses. NAACL

trained on 127M Twitter context-message-response triples

#	Context $c$	Message $m$	Generated Response $r$
1	noone can help this time its just i need a support system	well im here to support you. whenever you need me.	thank you, i appreciate it thank you
2	nothin much, and how's the book?!	its good but i'm only like halfway through cuz i don't feel like reading. i'm so bored ...	that's good! i have the book but i'm bored too.
3	i know! what have you been up to?	nothing really. just enjoying summer and working. what about you?	that's good! nothing much just listening to music and you?
4	thanks, your way better though	i'm just good at expressing it that's but trust me you are freakin amazing and waay better	i know, i don't trust you, i love you too! thank you
5	thanks. yeah i am good thanks! good, but angry!	sorry you're angry, been there, i hope it dies down.	i hope so too, i'm not angry!!
6	figures. bass. never ate bass.	bass is extremely mild with very few bones. my favorite after catfish	that's one of the bones, chuck bass is mine too.
7	today i will walk for miles. clear sky clear mind.	love walking. it's too hot to walk this month. i haven't moved. frozen. heat freeze.	it's too hot for you to get up to honey if you have frozen yogurt to you.
8	arresting people over "nasty tweets"? please. i'm not condoning what the lad did but arresting him is a tad silly.	it appears he may have sent death threats to others while all this was kicking off too.	i did! he was armed with nuclear threats? that's what's happening to you.

# Summary

- Commercial Systems:
  - Goal-oriented: ontology + hand-written rules for slot fillers
  - Non-goal-oriented:
    - Simple rule-based systems
    - IR-based: mine datasets of conversations.
- What's the future
  - More use of data
    - Neural net models
    - IR-based models
  - Problems:
    - Acquiring that data
    - Integrating goal-based and chatbot-based systems