CS 294S/294W
Democratizing Virtual Assistants
A Social-Good Research Project Course

Monica Lam

Stanford University
lam@cs.stanford.edu

Why a Remote Research Course?

A welcomed change from Zoom lectures.

Expose students to the exciting world of research.

This Class

1. Introduce an exciting research agenda

2. Explain the course design

3. Overview of the new methodology

4. Suggest research topics

5. Gather initial interest / Get to know each other

Exciting Time to Do CS Research

Computers get a new interface: Voice!

Talking Wikipedia	**Pervasive Dialogue Agents**	**End-user NL Programming**
General knowledge Q&A in all languages	A new software development toolset	Consumers/professionals automate their tasks
Add meaning to pretrained NL models	20M web developers → 20M NL developers!	Long-tail programming

OVAL: An Open-Source Initiative

The New York Times

Stanford Team Aims at Alexa and Siri With a Privacy-Minded Alternative
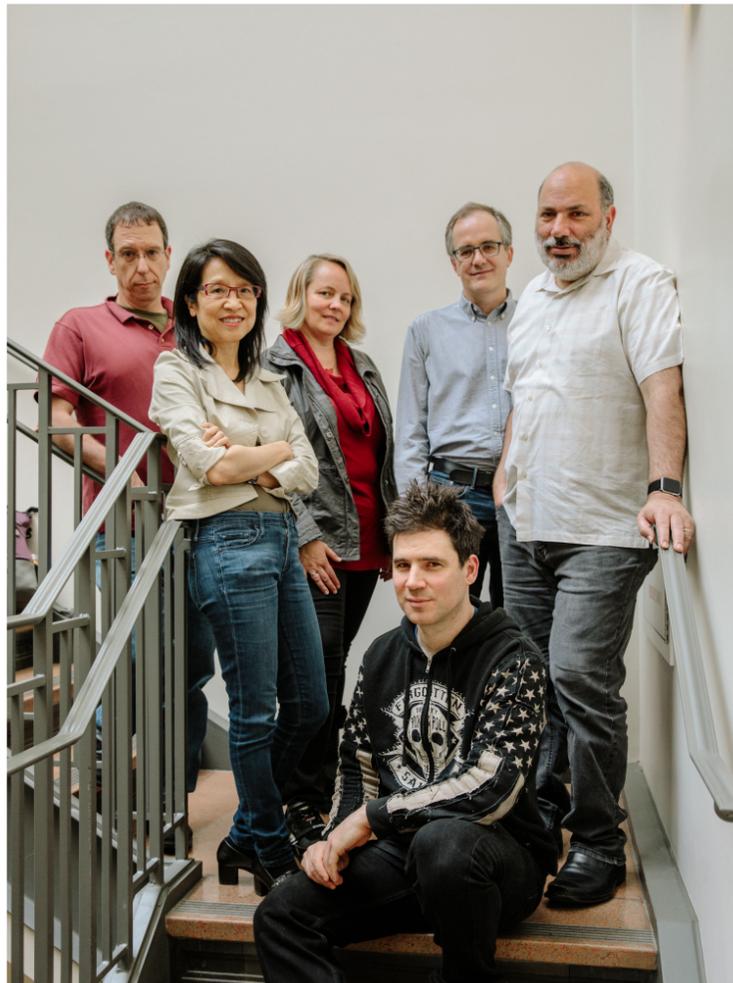
Sponsors

NSF   Alfred P. Sloan Foundation   Stanford Human-centered AI

Computer Science Faculty

Michael Bernstein	Dan Boneh	Monica Lam	James Landay
Fei-fei Li	Chris Manning	David Mazieres	Chris Re

Philanthropy & Digital Society   **Internet & Society Center**

Lucy Bernholz   Jen King

Students

Giovanni Campagna	Michael Fischer	Ranjay Krishna	
Mehrad Moradshahi	Sina Semnani	Silei Xu	Jackie Yang

An Open-Source Virtual Assistant Platform

GENIE

Virtual Assistant 2.0 Tools

Today:

Affordable only by the largest companies (Alexa: 10K employees)

Goal:

Democratize with affordable methodology & effective toolsets

THINGPEDIA

Crowdsourced Skill Repository

Today:

Proprietary voice web (Alexa: 100K 3rd party skills)

Goal:

Inter-operable skills open to all virtual assistants

ALMOND

Privacy-protecting assistant

Today:

Virtual assistants are ultimate surveillance tools

Goal:

A federated virtual assistant architecture that allows local execution.

Opportunities for many AI, HCI, Systems Research Projects

This Year's Infrastructure Goal

- An open privacy-preserving virtual assistant with the top 10 skills

  - Experimental research platform

  - An alternative for consumers (like Firefox)

  - To be released in June 2021

This Class

1. Introduce an exciting research agenda

2. Explain the course design

3. Overview of the new methodology

4. Suggest research topics

5. Gather initial interest / Get to know each other

A Research Course for Beginners

- Hardest part of a PhD: how to select a topic

  - Apprentice under a thesis supervisor

- A true and tried technique for junior researchers

  - Work with a professor, senior graduate students in a small group

  - Choose from an identified research project: meaningful and doable

    - Or suggest a new topic

- Groups of 2 or 3

Course Design

- Background
  - Lectures on basic technology and hands-on experience (2 homeworks)
- Project proposal (Discussions)
  - Proposed research projects in Google docs (on the website)
  - Your ideas are welcome
- 5-week projects
  - Due Mondays: Weekly status updates
  - Tuesday class: small group feedback
  - Thursday class: students give mini-lectures on their research topic (an important part of research training)
- Final project presentation and report

A Tentative Schedule

Week	Tuesday	Thursday	Due (10:30am)
Sep 15, 17	Course Introduction	Schema → Q&A (HW)	9/17: Student profile
Sep 22, 24	Schema → Dialogues	Project Discussions	9/24: HW due
Sep 29, Oct 1	Project Discussions	NL Primer	
Oct 6, 8	Proposals	Proposals	10/ 6: Project Proposal
Oct 13, 15	Group Meetings	Students' Mini-lectures	
Oct 20, 22	Group Meetings	Students' Mini-lectures	10/19: Weekly Update
Oct 27, 29	Group Meetings	Students' Mini-lectures	10/26: Weekly Update
Nov 3, 5	Group Meetings	Students' Mini-lectures	11/ 2: Weekly Update
Nov 10, 12	Group Meetings	Students' Mini-lectures	11/ 9: Weekly Update
Nov 17, 19	Final Project Presentation	Final Project Presentation	11/20: Project Report

Grading

- Attendance is mandatory
  - please let us know if you can't make it to class

- In-class participation: 15%

- Homework: 15%

- Final project: 70%

This Class

1. Introduce an exciting research agenda

2. Explain the course design

3. Overview of new methodology

4. Suggest research topics

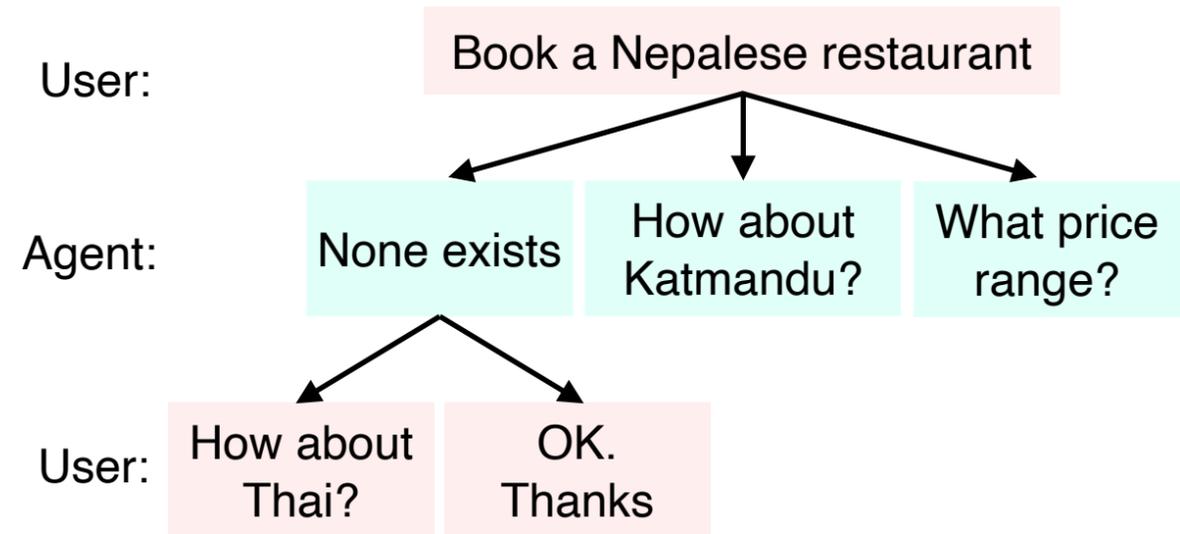5. Gather initial interest / Get to know each other

Paradigm Shift

Existing approach

1. Hand-annotated training data

   - Coverage, compositionally, cost, correctness
   - Alexa: 10,000 employees

User: | Book a Nepalese restaurant |

Agent: | None exists | How about Katmandu? | What price range? |

User: | How about Thai? | OK. Thanks |

2. Brittle dialogue trees
   Intent classifier per utterance

Virtual Assistant 2.0

1. Mostly synthesized training data, using pretrained language models

2. High-level programming
   One contextual neural network

Virtual Assistant 2.0

Dialogues + ThingTalk Annotations

Schema

Name	Price	Cuisine	...

+

Field Annotations

Genie

Can you help with information regarding a food place? I need to book at 15:45.

How about the restaurant with name La Tasca and Italian food?

Can you find something which serves

What date are you looking for?

Thursday please.

How about the Copper Kettle?
It is a food place with seafood food.

What is the price range and the area?

The Copper Kettle is a moderately priced restaurant in the north of the city.
Would you like a reservation?

No, thanks.

Can I help with you anything else?

Thank you, that will be it for now.

Train

NL→ThingTalk
Semantic
Parser

Dialogue
Agent

Iterative Refinement

Contextual Pure-Neural Semantic Parser
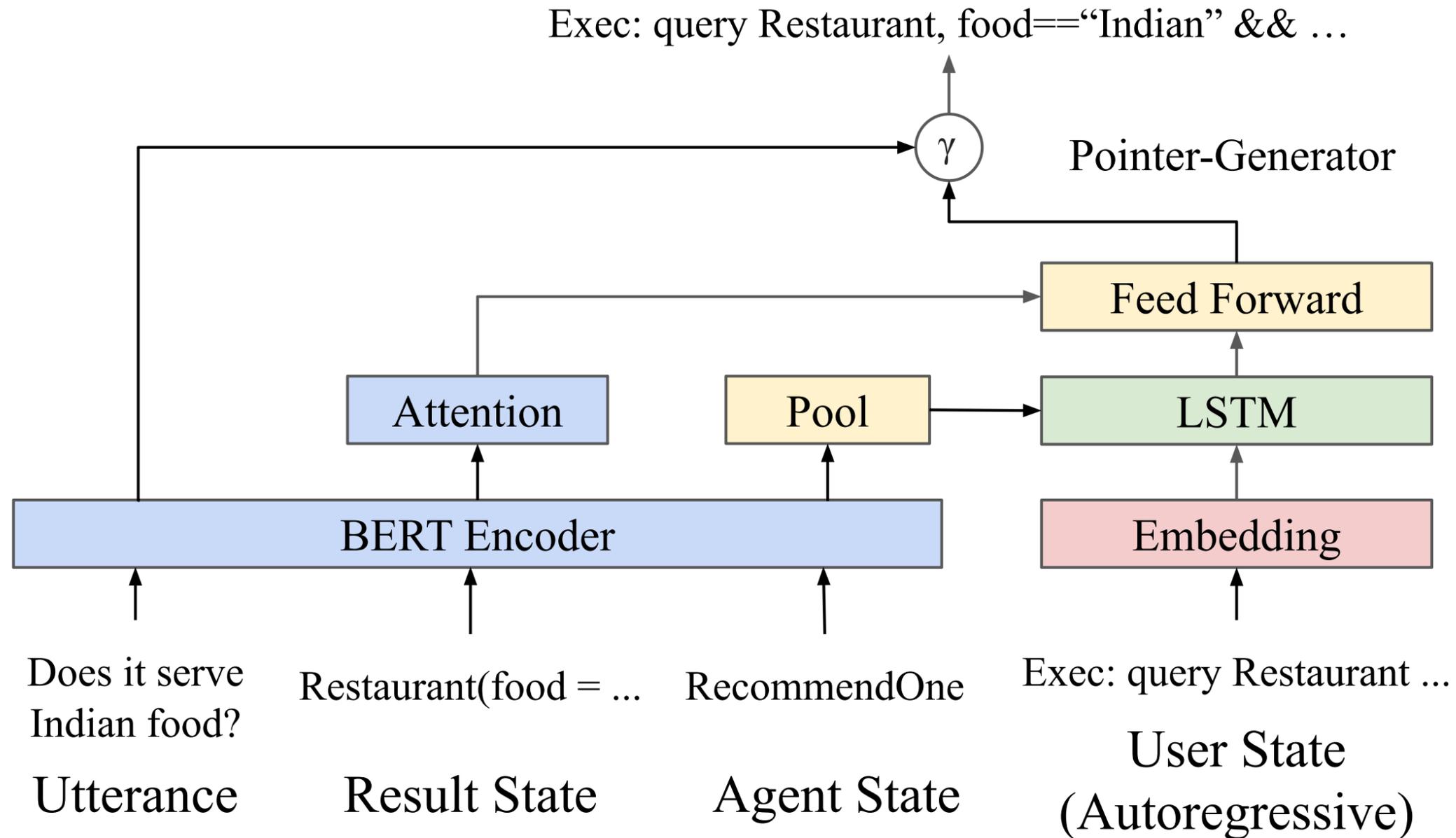
Dialogue State Tracking

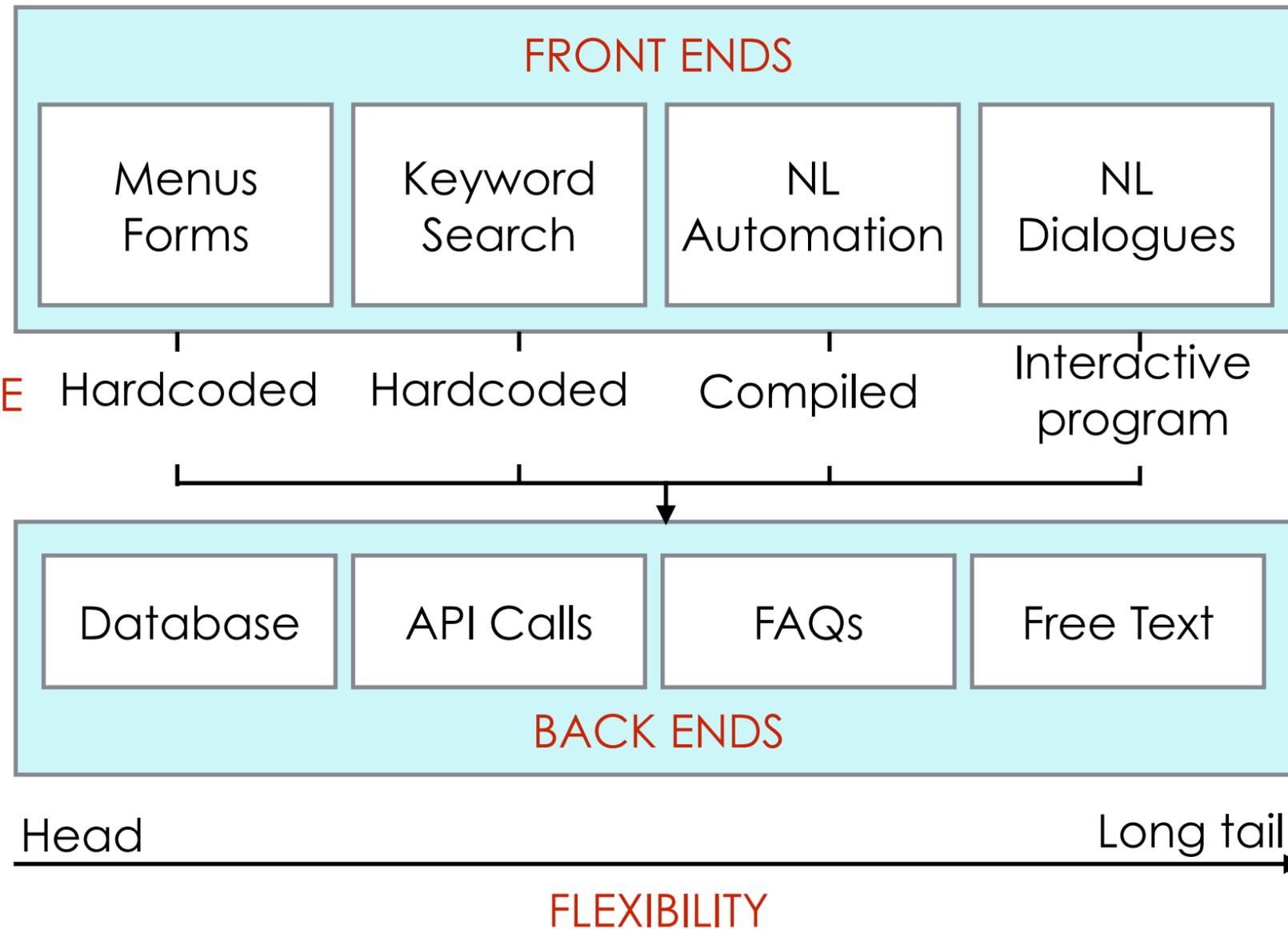| Model	Accuracy
Joint Accuracy (MultiWOZ 2.1)	
TRADE (Wu et al., 2019)	45.6
SUMBT (Lee et al., 2019a)	46.7
DSTQA (Zhou and Small, 2019)	51.2
DST-Picklist (Zhang et al., 2019a)	53.3
SST (Chen et al., 2020)	55.2
TripPy (Heck et al., 2020)	55.3
SimpleTOD (Hosseini-Asl et al., 2020)	55.7
Turn-By-Turn Accuracy (Cleaned Test Set)	
Genie	**71.1**

Genie

- Trained with only synthesized data

  - Perfect annotations

  - Validate and test with real data

- Need to track only the user state, one turn at a time

Answering Complex Questions

Queries	Alexa	Google	Siri	Genie
Show me restaurants rated at least 4 stars with at least 100 reviews				✓
Show restaurants in San Francisco rated higher than 4.5				✓
What is highest rated Chinese restaurant in Hawaii?	✓		✓	✓
How far is the closest 4 star and above restaurant?				✓
Find a W3C employee that went to Oxford				✓
Who worked for both Google and Amazon?				✓
Who graduated from Stanford and won a Nobel prize?		✓		✓
Who worked for at least 3 companies?				✓
Show me hotels with checkout time later than 12PM				✓
Which hotel has a swimming pool in this area?		✓		✓

New-Generation HCI: Voice

FRONT ENDS

Menus Forms	Keyword Search	NL Automation	NL Dialogues

INTERFACE

Hardcoded   Hardcoded   Compiled   Interactive program

Database	API Calls	FAQs	Free Text

BACK ENDS

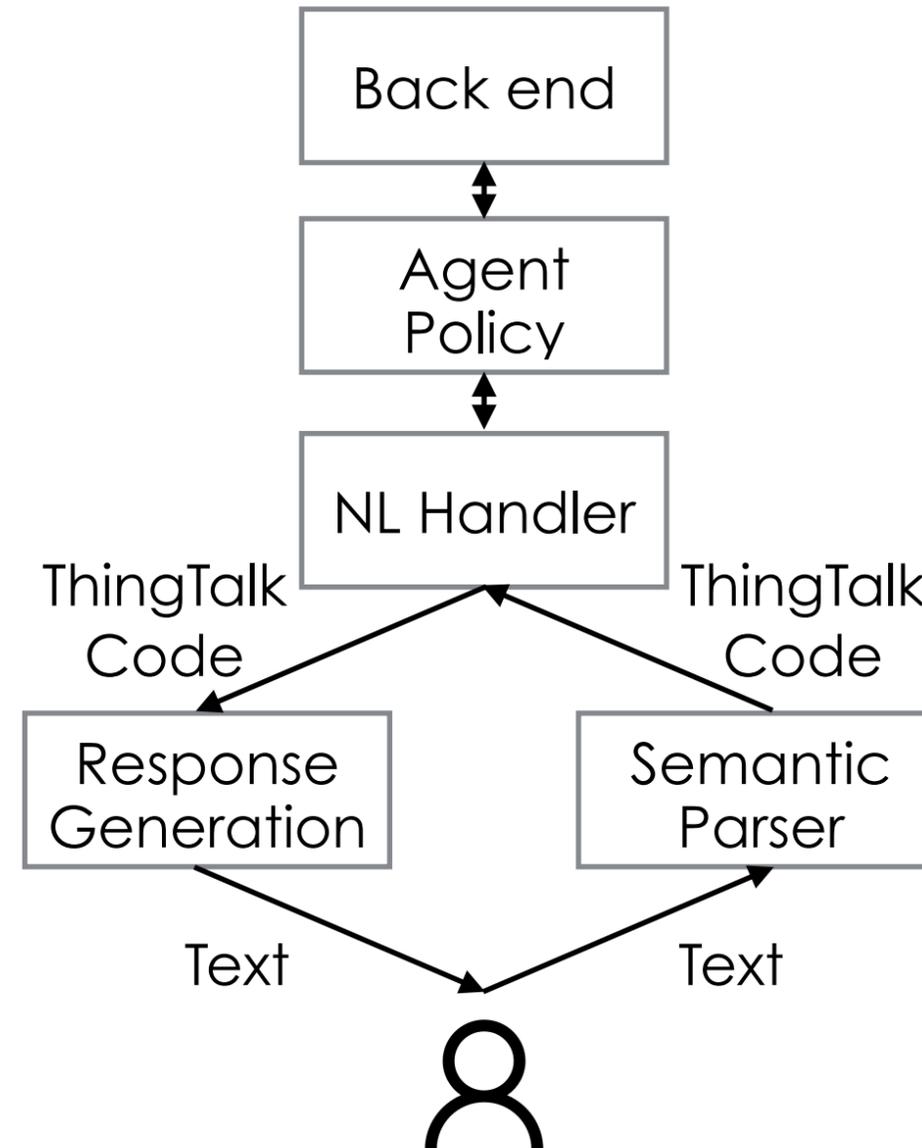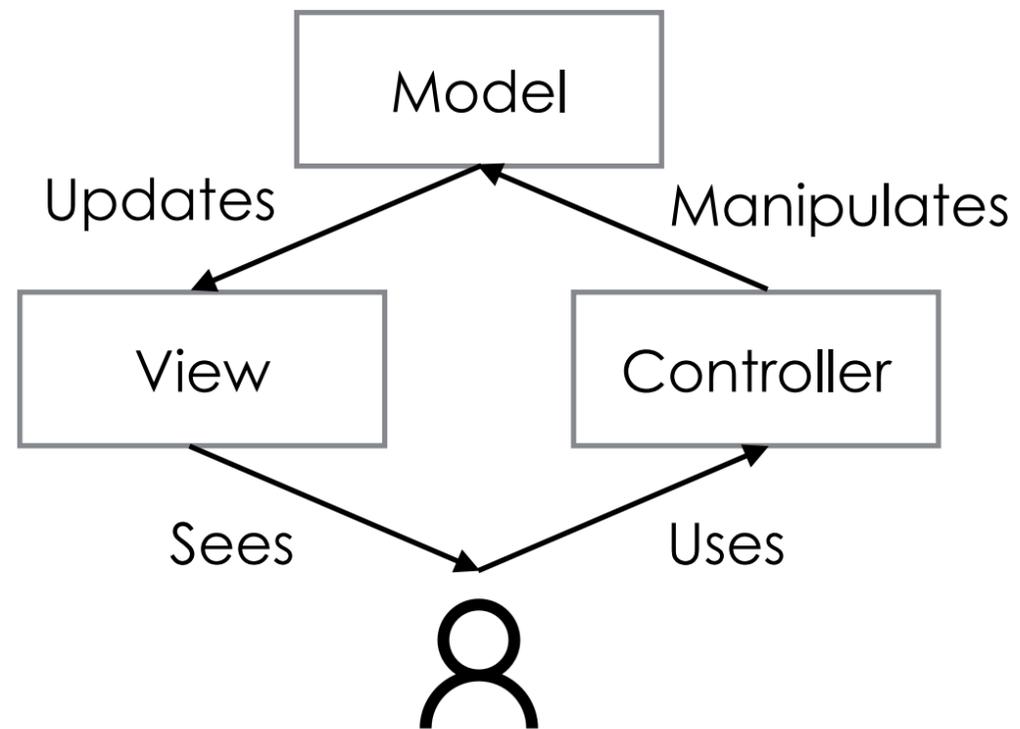Head → Long tail

FLEXIBILITY

NL Automation (User driven)

- Turn on the lights
- When apple stock drops to $100, buy 3 shares
- Find a Spanish restaurant that is open at 10pm in Palo Alto

NL Dialogues

- User-driven: reservations
- 2-way: doctor appts
- Agent-driven: Online teaching

MVC (Model View   Controller)
→ MRP (Model Response Parser)

Back end

Agent Policy

Model

NL Handler

Updates   Manipulates

ThingTalk Code   ThingTalk Code

View   Controller

Response Generation   Semantic Parser

Sees   Uses

Text   Text

This Class

1. Introduce an exciting research agenda

2. Explain the course design

3. Overview of the new methodology

4. Suggest research topics

5. Gather initial interest / Get to know each other

Research Projects

Problem	Area	Goal	Examples
Wikidata in NL	Systems	Scalability	Develop methodology & tools to cover Wikidata
	AI	Scalability	Zero-shot learning using type information
Usable Dialogue Agents (Transactions)	AI	Breadth	Generalize a contextual neural network from 5 (Multiwoz) to 11 domains (SGD)
		Accuracy	Named entity disambiguation in the wild (Bootleg)
		Error detection	Neural network to identify likely correct components
		Response fluency	Use Bart to generate fluent responses
		Multilingual: Localization	Use machine translation with entities in target languages (Chinese Multiwoz, CrossWoz)
	HCI	Usability	Conversational Q&A dialogue design for music, movies, etc
		Design	Dialogue to support function discovery
		Multimodal	Combining the best of voice and text in assistants
	Systems	Knowledge	Representation (time, location)

Multi-disciplinary Research Projects

Problem	Examples
Advanced Agents	Generic FAQ dialogue models
	Personalized agents with users' history & profile (e.g. ordering food)
End-user programming	A gentle way to introduce end-users to creating skills: cron jobs, monitors, comparison shopping
	Automate end-user routines with demonstrations (e.g. workout assistants)
End-to-end skills	Home Automation. IoTs for 1000 devices (with tens of abstract devices) Almond is the voice interface for Home Assistant
	News, sports, radios, podcasts: Listening + asking questions
	Safe voting, legal advice, personal finance