CS 294S/294W
Human-Centered Software: AI meets Systems

Monica Lam
Computer Science Department
Stanford University
lam@cs.stanford.edu

With Giovanni Campagna, Michael Fischer, Mehrad Moradshahi, Rakesh Ramesh, Richard Socher, Silei Xu, Richard Yang
Sponsors: NSF, AVG, Google, HTC, Hitachi, ING Direct, Nokia, Samsung, Sony Ericsson, UST Global
History of the Course

- June 2007: iPhone arrived!
  - Opportunity of a new computing paradigm: Mobile
- Sept 2007: Research proposal
  - Programmable Open Mobile Internet 2020
  - 22 professors: Radio, networking, systems, security, education
  - 1st NSF Expedition funded in June 2008
- Sept 2008: The first CS 294S/W class
  - Small interdisciplinary project-based course
CS 294S/W (2008-2012)

• Questions:
  • What will the mobile phone replace? What do the new apps look like?
  • Who will the dominant owner of private information? Can we protect privacy?

• Projects:
  • Prototype of mobile app
    • Innovative (e.g. ride sharing)
    • More sophisticated apps with (optional) infrastructure support
  • Infrastructure
    • Musubi: Disintermediated interactive social feeds for mobile devices (top 2 papers in WWW '12)
CS 294S/W: A Project Course

• Not your traditional course — not lectures & assignments — cannot be a MOOC
• Alternative learning method: How to learn by experiments?
  • Undergrad: senior project + research theme
  • MS: a mini research minor
  • Prepare you for R & D (Research and Development) in universities / companies
• “Independent / cross-disciplinary study” projects by groups of 1-3:
  • Product prototype with user study —> research prototype that may fail
• What will you learn?
  • Choose a topic, literature search, build prototype, evaluate, refine
    • Key criterion for choosing a project: what can we learn?
  • Collaboration, presentations, discussions, brainstorming
  • Hands-on experience + other teams’ experience
How Do I Teach?

Apprenticeship

• Hardest part of a PhD: how to select a topic
• Provide context & tools for research and potential future products
• Group meetings outside class
  • to help define topics and make progress
  • to help prepare for group discussions
• Help find resources
# Course Schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Activity</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apr 2</td>
<td>Introduction</td>
<td>Brainstorming</td>
</tr>
<tr>
<td>Apr 9</td>
<td>Forming projects</td>
<td>Forming projects</td>
</tr>
<tr>
<td>Apr 16</td>
<td>Mini Hackathon</td>
<td>Project proposals</td>
</tr>
<tr>
<td>Apr 23</td>
<td>Project proposals</td>
<td>Project proposals</td>
</tr>
<tr>
<td>Apr 30</td>
<td>Tutorial</td>
<td>Tutorial</td>
</tr>
<tr>
<td>May 6</td>
<td>Student-led discussions</td>
<td>Student-led discussions</td>
</tr>
<tr>
<td>May 14</td>
<td>Mini Hackathon</td>
<td>Student-led discussions</td>
</tr>
<tr>
<td>May 21</td>
<td>Student-led discussions</td>
<td>Student-led discussions</td>
</tr>
<tr>
<td>May 28</td>
<td>Final project discussions</td>
<td>Final project discussions</td>
</tr>
<tr>
<td>Jun 4</td>
<td>Final project discussions</td>
<td></td>
</tr>
</tbody>
</table>

Jun 10: (1:30-3:15 pm)  Final project demo (3 minute video) and poster session  
Jun 11  Final report due
Course Grading

• 15% Class participation (compulsory attendance)
• 5% Homework
• 80% Project

• Definition of success:  
  how much do we learn?  
  how should we define our next research project?

• Satisfy expectations set up in individual meetings  
  will get an A for the project
CS 294S/294W (2015-now)

- 2014: Alexa arrived!
  - Opportunity of a new paradigm: ubiquitous computing with IoTs
  - Human-centered software: AI meets Systems
- Research results to date (http://almond.stanford.edu)
  - Almond: The architecture of an open, crowdsourced, privacy-preserving, programmable virtual assistant (WWW '17)
  - Controlling fine-grain sharing in natural language with a virtual assistant (UbiComp '18)
  - Genie: A generator of natural language semantic parsers for virtual assistant commands (PLDI '19)
  - Brassau: Automatically generating graphical user interfaces for virtual assistants (MobileHCI '18)
  - ImagineNet: Detail-preserving neural style transfer (submitted)
- 2019 April: NSF Proposal funded
  - Autonomy and privacy with open federated virtual assistants
  - 8 professors in Systems, AI, HCI, Security
What will the future hold?

- How ubiquitous will computing become?
- What will be the human-computer interface?
- What can virtual assistants do?
- Facebook sells 2 billion people’s private data: will the problem get better or worse?
- Can ML and privacy preservation co-exist?
Concepts & Infrastructure

• Natural language programming
• Tool to generate semantic parsers
• Sharing with access control
• Automatic generation of GUIs
Example: Asthma Patient

**Dr. Smith:**
- "if Bob’s peak flow-meter drops below 180L/min let me know"
- "when the air quality index is above 500 and Bob is running, warn him"

**Bob:**
- "Let my Dad know if I am at the hospital"
- "when I use my inhaler, record my GPS location in logfile on Box"

**Diagram:**
- Devices: "when I use my inhaler, record my GPS location in logfile on Box"
- Location: "Let my Dad know if I am at the hospital"
- Environment: "when the air quality index is above 500 and Bob is running, warn him"
Impact of Natural Language Programming

- Natural language: highest-level programming languages
- Gives the power of programming to end users!
- Today’s software cannot satisfy the long tail of user needs
Core Concepts

Natural Language → LUInet (Neural Network) → ThingTalk (Control Constructs) → Formal Virtual Assistant Language → Open Interoperable Web

+ Thingpedia (Full API Signatures)
Natural Language Programming

“When I use my inhaler, get my GPS location, if it is not home, write it to logfile in Box.”

• Event-driven program
• Multiple function calls
• Parameter passing
• Filters on values
"When I use my inhaler, get my GPS location, if it is not home, write it to logfile in Box."

```
monitor @Inhaler-use(), => @GPS(), location <> "home"
=> @Box-write(file="logfile", data=location)
```
### Thingpedia: Encyclopedia of Things

- **Interoperability**
  - API signatures + corresponding NL
  - Not just intents

- **Open repository**
  - Available to Alexa, Google Assistant, ...

> 60 devices / 200 functions

<table>
<thead>
<tr>
<th>Natural Language</th>
<th>API Signatures</th>
</tr>
</thead>
</table>
| **WHEN**
  - @Stanford tweets | Monitor (@home_timeline(), ...) author="Stanford") |
| **GET**
  - tweets matching "#Cardinal" | search(...), contains (hashtag, ...) |
| **DO**
  - tweet “Stanford won!” | post (status) |
THINGTALK Compound Statement

**WHEN** [FILTERS] → **GET** [FILTERS] → **DO**

**FILTERS:** =, <, >, <=, >=, <>, contains, starts with, ends with

- When I use my inhaler, get my location, save them to Dropbox
- If I get taken to a hospital, let my dad know.
- When the air quality index is above 500, and I am running, send me an SMS.
Real Natural Language Input

When my car is at home, and it is not plugged in, send me a reminder email.

Remind me if my car is not plugged in at home. If I am not charging my car when it is home, let me know. Remind me to plug in my car whenever I’m home.
Genie: A Semantic Parser Generation Tool

<table>
<thead>
<tr>
<th>WHEN</th>
<th>Natural Language</th>
<th>API Signatures</th>
</tr>
</thead>
<tbody>
<tr>
<td>GET</td>
<td>tweets matching &quot;#Cardinal&quot;</td>
<td>search(...), contains (hashtag, ...)</td>
</tr>
<tr>
<td>DO</td>
<td>tweet &quot;Stanford won!&quot;</td>
<td>post (status)</td>
</tr>
</tbody>
</table>

Real User Input — Dev Set

Natural Language

LUInet

Code

Training Data

Parameter & data augmentation

Paraphrase

Synthesize programs

ThingTalk Grammar

NL Templates

When <> get <>
Get <> when <>

Thingpedia
LUInet Results

- **Dataset (44 skills, 122 functions, 147 parameters)**
  - 664K synthetic, 24K paraphrased sentences
  - Training set after parameter expansion: 2.9 M sentences

- **Model: DecaNLP — MQAN (Multi-Task Question Answering Network):**
  - Encoder (bottom to top):
    - BiLSTM
    - Co-attention (context-question)
    - BiLSTM
    - Self-Attention
  - Decoder:
    - LSTM
    - Self-Attention
    - Encoder-Decoder Attention
    - Mixed pointer-generator layer

- **Accuracy:** 87% on paraphrases; 62% on real user data
Sharing with privacy

Let your virtual assistant help you share
General+Fine-Grain: ThingTalk Extension

Requester:

\[ \text{GET-PREDICATE [FILTERS]} \]

\[ \text{WHEN [FILTERS]} \rightarrow \text{GET [FILTERS]} \rightarrow \text{DO} \]

FILTERS: \( =, <, >, <=, >=, \text{contains}, \text{starts with}, \text{ends with} \)

Let Dr. Smith monitor my peak-flow-meter, if it drops below 180L/min
Let my father monitor my security camera for motion,
Let my secretary, whenever I am out of town, read email messages whose subject is marked urgent.
Let my daughter, from 6-8pm, watch NetFlix
Let my boyfriend get pictures from my dropbox, taken on Feb 14, and post them on Facebook
Almond: 1st Federated Virtual Assistant

Expressiveness:
Any ThingTalk command

Privacy:
Remote execution model
Owner executes requests
Returns need-to-know

Giovanni, Xu, Ramesh, Fischer, Lam, Ubicomp 2018
Remote Execution

“Ask @alice to notify me when her security camera detects motion.”

σ=@dad, c=SELF: monitor @security_camera.event(), has_motion=true ⇒ return

σ=@dad, c=SELF: monitor @security_camera.event(), has_motion=true && @phone.get_gps() {location≠home} ⇒ return

“@dad wants to get notified when any event is detected on your security camera and has motion is equal to true.”

(a) Request

(f) Return detected events

“Notification from monitor security camera: motion detected ...”

Policy Database

(b) Check

(e) Save

Alice’s Assistant

(c) Ask for permission

“Only if I’m not home.”

(d) Respond

Alice

Dad
Conformance of Access Control

Natural Language

2nd-Party Program Access Control

Satisfiability Modulo Theories (SMT)
Needs and Acceptance?
Do Consumers Need Access Control?

<table>
<thead>
<tr>
<th>Role-Based Permission</th>
<th>Attribute-Based Permission</th>
</tr>
</thead>
</table>
| **Teenage daughter** to use credit card | **With a $20 budget limit**  
For restaurants only |
| **Amazon courier** to unlock door | **If the package is over $1000**  
If your security camera is on |
| **Friends** to access cloud drive | **Photos with their faces in them**  
Photos in a specific folder |
| **Parent/kid** to see security cameras | **If you are not at home**  
Cameras facing the front yard/garage |
| **10-year-old kid** to use Netflix | **Between 7 PM to 9 PM**  
Free G or PG rated movies |

% People comfortable in giving permission (200 person survey)

% 100%
More Examples

Willingness to share doubles with attribute-based access control
Automatic Generation of GUI

- Linguistic user interface (LUI) is inadequate
- Future: LUI + GUI
Automatic GUI Generation

“monitor the camera in the baby room”

“monitor mails from my dad”

“adjust the volume of my speakers”

“translate ‘hello’ to Chinese”

“I’d like to order a coffee”

Fischer, Giovanni, Xu, Lam, MobileHCI 2018
Restyling in Real-Time

“monitor the camera in the baby room”
“monitor mails from my dad”
“adjust the volume of my speakers”
“translate ‘hello’ to Chinese”
“I’d like to order a coffee”

Bicentennial Print
Roy Lichtenstein, 1975