Dialogue Systems and Human Machine Interaction – The User

Dr. Robert Ross
Applied Intelligence Research Centre
Dublin Institute of Technology
A Cognitive Perspective

Artificial Intelligence

Machine Learning

Deep Learning
The subset of machine learning composed of algorithms that permit software to train itself to perform tasks, like speech and image recognition, by exposing multilayered neural networks to vast amounts of data.

Any technique that enables computers to mimic human intelligence, using logic, if-then rules, decision trees, and machine learning (including deep learning).

A subset of AI that includes abstruse statistical techniques that enable machines to improve at tasks with experience. The category includes deep learning.
When I Started with Dialogue
A Wasted Youth!

Early 2000s
Speaking Autonomous Intelligent Devices

Late 2000s
Shared Control via Dialogue

2010s
Situated Intelligent Systems
Situated Interactive Systems

Definition

*Systems immersed in a real or virtual environment which can perceive, reason on, and act on that environment*

Examples

- Automotive systems
- Virtual characters
- Assistance systems
- Service robots

- Graphical, textual and tactile communication insufficient
- Importance of the verbal channel
Architectures of natural language processing components which can be used to enable spoken interfaces to applications

Prototypical spoken dialogue system architecture

Dialogue manager responsibilities
- Link language technology and domain component logic
- Integrate user utterances (dialogue actions)
- Plan system dialogue actions
Exciting Times!!!
Speech Recognition

Scene from Burn After Ready 2008
"Dialogue Systems are Academic Suicide"

-- said by a visiting Human Computer Interaction professor to me 6 months before I submitted my dissertation
Supervised Learning

Training Data Pairs → Learning Process → Model

- Input
- Output

- Classification
- Regression
### Alexa at Large!

<table>
<thead>
<tr>
<th>Message</th>
<th>Time</th>
<th>Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown</td>
<td>Today at 5:53 AM</td>
<td>Living Room echo dot</td>
</tr>
<tr>
<td>reply</td>
<td>Today at 5:53 AM</td>
<td>Living Room echo dot</td>
</tr>
<tr>
<td>alexa search briefing</td>
<td>Today at 5:27 AM</td>
<td>Kitchen Echo</td>
</tr>
<tr>
<td>alexa play some classical music</td>
<td>Yesterday at 5:26 PM</td>
<td>Kitchen Echo</td>
</tr>
<tr>
<td>Text not available. Click to play recording.</td>
<td>Yesterday at 5:26 PM</td>
<td>Living Room echo dot</td>
</tr>
<tr>
<td>alexa app</td>
<td>Yesterday at 5:25 PM</td>
<td>Kitchen Echo</td>
</tr>
<tr>
<td>alexa plush briefing</td>
<td>Yesterday at 5:19 PM</td>
<td>Kitchen Echo</td>
</tr>
<tr>
<td>alexa place briefing</td>
<td>Yesterday at 5:18 PM</td>
<td>Kitchen Echo</td>
</tr>
<tr>
<td>alexa play b. b. king of limbs</td>
<td>Yesterday at 9:48 AM</td>
<td>Kitchen Echo</td>
</tr>
<tr>
<td>wash briefing</td>
<td>Yesterday at 9:37 AM</td>
<td>Kitchen Echo</td>
</tr>
</tbody>
</table>
Features / Problems
- Slow Serialized Communication
- Fuzzy Meanings & Imprecise Knowledge
Interpretation Challenges: Reference Frames
“I’m goin’ down to check my thing”
From Interpretation to Production
## Qualities of Produced Content - Grice

<table>
<thead>
<tr>
<th>Maxim Name</th>
<th>Maxim Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quantity</td>
<td>Say neither more nor less than the discussion requires</td>
</tr>
<tr>
<td>Relevance</td>
<td>Be relevant to the discourse and conversational goals</td>
</tr>
<tr>
<td>Manner</td>
<td>Be orderly and coherent</td>
</tr>
<tr>
<td>Quality</td>
<td>Do not lie or make unsupported claims</td>
</tr>
</tbody>
</table>
• Tactics & Models
  • Psycholinguistic Mechanisms: Interactive Alignment, others.
  • Multi-Functionality in expression... **Communicative Function**
<table>
<thead>
<tr>
<th>Command and Control</th>
<th>Voice based Information Retrieval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fluid Conversational Systems</td>
<td>Situated Task Oriented Dialogue</td>
</tr>
</tbody>
</table>

Do we need to worry about this?
Alignment Hypothesis

- Interlocutors in HHI studies show alignment effects
- Alignment frequent in task-oriented dialogues (Reitter et al 2006a/b)
- Alignment is associated with task success
  - Tutorial system dialogues (Ward 2007)
  - Route instruction dialogues (Reitter et al 2007)
- People can treat machines like people
- Empirical evidence of HCI alignment
  - Loudness & Pause Latency (Suzuk et al 2007)
Design Consequences

- **Natural Conversational Agents**
  - By definition must employ alignment and adaptation models

- **Practical Dialogue Systems**
  - Alignment may enhance user acceptance
  - Language Shaping
  - Success Monitoring
- Bi-directionality of internal language resources
- Track partner & self language use
- Make production processes sensitive to extended context
Where is the User Model?

System → Data about user → User Modelling → User Model → Adaptation → Adaptation effect

User
Design Focus

- Bi-directionality of internal language resources
- Track partner & self language use
- Make production processes sensitive to extended context
Challenges

Model Integration

Training Data
My wife asked me why I carry a gun with me around the house. I told her I think the CIA is spying on us. She laughed, I laughed, our Amazon Echo laughed. I shot the Echo.
This presentation contains many images - some of which were pulled from the internet many years ago – I can’t remember where / when.

All images are however used in the spirit of Fair Use for Education purposes. If you are a content holder, please don’t hesitate to get on to me to have your image removed if you don’t believe it is being used under fair use practices.

R. Ross