Connections

Eric Horvitz

ICMI 2015
November 12, 2015

Sustained Accomplishment Award Lecture
Toward Fluid Connectivity

Promise of Deeper Human-Machine Connection & Collaboration
An early connection

Engaging at NASA’s Mission Control Center on human-in-the-loop
<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fuel Press</strong></td>
<td>10.5</td>
<td>11.8</td>
<td></td>
</tr>
<tr>
<td><strong>Chamb Pres</strong></td>
<td>5.4</td>
<td>4.8</td>
<td></td>
</tr>
<tr>
<td><strong>He Pres</strong></td>
<td>17.7</td>
<td>14.7</td>
<td></td>
</tr>
<tr>
<td><strong>Delta v</strong></td>
<td>33.3</td>
<td>63.3</td>
<td></td>
</tr>
<tr>
<td><strong>Oxygen</strong></td>
<td>10.2</td>
<td>10.6</td>
<td></td>
</tr>
<tr>
<td><strong>Fuel Press</strong></td>
<td>12.8</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td><strong>Chamb Pres</strong></td>
<td>0.0</td>
<td>15.7</td>
<td></td>
</tr>
<tr>
<td><strong>He Pres</strong></td>
<td>32.3</td>
<td>63.3</td>
<td></td>
</tr>
</tbody>
</table>

**System**

**Display**

**Act, t**

**Delay**

**Utility**
Opportunity: Complementary Computing
Direction: Augment human cognition

Pillars
Pillar: Inferring Beliefs, Goals, Knowledge

Predictions about world

Diagram showing relationships between beliefs, goals, and knowledge.
Pillar: Inferring Beliefs, Goals, Knowledge

Predictions about world

Inferences on beliefs, goals, knowledge
Predictions about world

Inferences on beliefs, goals, knowledge

Pillar: Inferring Beliefs, Goals, Knowledge
Pillar: Inferring Beliefs, Goals, Knowledge

Predictions about world

Inferences on beliefs, goals, knowledge

Ideal actions under uncertainty
Pillar: Complementarity

Machine Intellect

Human Intellect
Pillar: Complementarity

Machine Intellect

Human Cognition

Opportunity: Complementary Computing
Direction: Augment human cognition
Pillar: Complementarity
Pillar: Complementarity

Memory
Attention
Judgment

Human Cognition
Pillar: Mix of Initiatives
Pillar: Mix of Initiatives
Pillar: Mix of Initiatives
Pillar: Coordination

Continuous process of contributing, signaling, monitoring

Clark, Duncan, Goffman, Goodwin, Kendon, et al.
Pillar: Coordination

Continuous process of contributing, signaling, monitoring

Clark, Duncan, Goffman, Goodwin, Kendon, et al.
Pillar: Coordination

Continuous process of contributing, signaling, monitoring

Clark, Duncan, Goffman, Goodwin, Kendon, et al.
Pillar: Coordination

Continuous process of contributing, signaling, monitoring

Clark, Duncan, Goffman, Goodwin, Kendon, et al.
Coordination in Open World

Situated interaction
Bohus, H. et al.
Coordination in Open World

Situated interaction
Bohus, H. et al.
Coordination in Open World

Situated interaction
Bohus, H. et al.
Coordination in Open World

Situated interaction

Bohus, H. et al.
Reflections on several research efforts
Efforts in Complementarity

Machine Intellect

Human Intellect
Handsfree Trauma Care System (1991)

Handsfree Trauma Care System

Handsfree Trauma Care System

Video

Lumiere Project (1993)

Goals, Understanding, and Uncertainty

Infer likelihoods of user’s goals, attention, understanding and take ideal actions

- User query
- User activity
- User location
- User profile
- Data structures
- Vision, speech, sound

Pr(Goals, Understanding)

Value-Focused Action

Slides from early days at Microsoft Research...
Lumiere Project (1993)

Big Picture

Events Synthesis → Inference about User, World → Computation of Ideal UI Action → Ideal Actions

Learning Models

Events

Control

New Perceptual Actions?
Lumière Project (1993)

Actions + Words $\rightarrow$ Goals
Learning about Assisting Computer Users
Wizard of Oz Studies

Expert peeks through a keyhole, plays assistant role

User with challenge task

Expert as Wizard of Oz “Agent”

Video
Evidential distinctions identified

- **Search**: e.g., exploring of multiple menus
- **Introspection**: e.g., sudden pause, slowing of command stream
- **Focus of attention**: e.g., selected objects
- **Undesired effects**: e.g., command/undo, dialogue opened and cancelled
- **Inefficient command sequences**
- **Syntactic / semantic content of file**
- **Goal-specific sequences of actions**
Building Bayesian user model

Building Bayesian user model

User Needs Assistance

Recent Menu Surfing

Pause after Activity

Building Bayesian user model

- User Expertise
- Difficulty of Current Task
- Recent Menu Surfing
- User Needs Assistance
- Pause after Activity
- User Distracted

Building a Bayesian user model involves considering various factors such as user expertise, difficulty of the current task, recent menu surfing, user needs, assistance required, user distraction, and the need for pausing activities.
Event Streams and Architectures

Sensed actions
Sensed actions
Sensed actions

Help Wizard's Best Guesses for Help:
- Autoformatting my document
- Format a chart
- Changing alignment
- Working with borders
- Working with fonts

Freely describe your problem:
how do i make this look prettier?

User's query
Prob. user desires assistance

Probabilities of relevant help topics given a request for help
Prob. user desires assistance
Video: Lumiere
Efforts with Mix of Initiatives

Lookout (1998)

In-stream supervision

Infer $p(\text{goal} = \text{schedule} \mid E)$
Lookout (1998)

In-stream supervision

- Do nothing?
- Engage user? *(and when?)*
- Just do it?

→ Infer $p(\text{goal} = \text{schedule} \mid E)$

Lookout (1998)

In-stream supervision

- Do nothing?
- Engage user? (and when?)
- Just do it?

→ Infer $p(\text{goal} = \text{schedule} \mid E)$

Lookout (1998)

In-stream supervision

→ Infer $p(\text{goal} = \text{schedule} \mid E)$

Getting the timing right

- Do nothing?
- Engage user? *(and when?)*
- Just do it?

Predictive Model

- **No Action**
- **Ask**
- **Action**

$P^*$

$u(A,D)$

$u(A,D)$

$u(\overline{A},D)$

$u(\overline{A},\overline{D})$
Lookout (1998)

In-stream supervision

→ Infer \( p(\text{goal} = \text{schedule} | E) \)

- Do nothing?
- Engage user? (and when?)
- Just do it?

Getting the timing right

![Graph showing dwell time vs. length of original message](image)

You will be busy then with the Lumiere project meeting... How about trying Friday at 3 PM.

```
From: Robert Croft  Sent: Tue 9/22/98 8:25 PM
To: Eric Horvitz
Cc: 
Subject: Planning for session

I'd like to catch up on plans for the conf. session on decision making. How about speaking by phone sometime around 2pm on Fri? Looking forward to chatting.

-- Robert
```
Dear Executive Council Members:

The meeting on Sunday will be held in the Sequoia Room on the upper level of Tressider Union at Stanford University. The meeting will run from 1:00-5:00 pm and a light lunch will be available at 12:30 pm. An agenda will be sent to you in a subsequent message. Currently, the following people plan to attend the meeting:

[Names of attendees]

Would you like to schedule an appointment?
Dear Executive Council Members:

The meeting on Sunday will be held in the Sequoia Room on the upper level of Tressider Union at Stanford University. The meeting will run from 1:00-5:00 pm and a light lunch will be available at 12:30 pm. An agenda will be sent to you in a subsequent message. Currently, the following people plan to attend the meeting:

- ...

---

How about tomorrow at 1 PM?

Dear Executive Council Members:

The meeting on Sunday will be held in the Sequoia Room on the upper level of Tressider Union at Stanford University. The meeting will run from 1:00-5:00 pm and a light lunch will be available at 12:30 pm. An agenda will be sent to you in a subsequent message. Currently, the following people plan to attend the meeting:

Dear Executive Council Members:

The meeting on Sunday will be held in the Sequoia Room on the upper level of Tressider Union at Stanford University. The meeting will run from 1:00-5:00 pm and a light lunch will be available at 12:30 pm. An agenda will be sent to you in a subsequent message. Currently, the following people plan to attend the meeting:
Video: Lookout system
TV commercial (1999)

Where do you want to go today?

Eric, I’d like to discuss those business plans we started talking about. How does meeting 2pm on Saturday sound?

Thanks... Let me know...
Architecture for conversation (1999)

Extending mixed-initiative interaction to hierarchies of contribution


Paek, H. *Conversation as Action Under Uncertainty*. UAI 2000
Bayesian Receptionist (2000)

"I need a ride."

User’s Goal

Goal 1

Goal n

Subgoal 11

Subgoal 1x

Subgoal 1x1

Subgoal 1xy

Level 0

Level 1

Level 3
Bayesian Receptionist (2000)

“*I need a ride.*”

User’s Goal

Goal 1

Subgoal 11
Subgoal 1x
Subgoal 1x1
Subgoal 1xy

Level 0

Level 1

Level 3

VOI

“Can you elaborate on that?”
Bayesian Receptionist (2000)

"I need a ride."

User’s Goal

Goal 1

Subgoal 11
Subgoal 1x
Subgoal 1x1
Subgoal 1xy

VOI

Subgoal 1

Level 0

Level 1

Level 3

"So you'd like a shuttle?"

I need a ride please.

DECL1

NP1

PRON1* "I"

VERB1* "need"

NP2

DETP1

ADJ1*

NOUN1* "ride"

AVP1

ADV1* "please"

CHAR1 ".,"
Bayesian Receptionist (2000)

User’s Goal

"I need a ride."

Goal 1

Subgoal 11

Subgoal 1x

Subgoal 1x1

Subgoal 1xy
Bayesian Receptionist (2000)
Advances in Core Capabilities
Core Fabric: Multisensory Fusion

Fuse vision, acoustics, activity with computer (Seer, ICMI 2002)

→ Representation, learning, inference

Core Fabric: Selective Perception

Guide computation to where it counts (S-SEER, ICMI 2003)
Compute expected value of information

![Diagram showing on and off states of Audio Classification, Video Classification, Sound Localization, and Keyboard/Mouse over time.]

Core Fabric: Selective Perception

Guide computation to where it counts (S-SEER, ICMI 2003)
Compute expected value of information

DC: Distant Conversation
NP: Nobody Present
O: Other
P: Presentation
FFC: F-F Conversation
WC: Working on computer
PC: Phone Conversation

Core Fabric: Tools for Multisensory Prototypes
Core Fabric: Tools for Multisensory Prototypes

Video: Sensing Smartphone (2000)

Hinckley, Pierce, Sinclair, Horvitz. Sensing Techniques for Mobile Interaction, UIST 2000
Core Fabric: Tools for Multisensory Prototypes

Video: Surface computing (2004)

Wilson, Sarin. BlueTable: Connecting Wireless Mobile Devices on Interactive Surfaces Using Vision-Based Handshaking, GI 2007
Olwal, Wilson. SurfaceFusion: Unobtrusive Tracking of Everyday Objects in Tangible User Interfaces, GI 2008
Core Fabric: Probabilistic Fusion of Signals

Core Fabric: Probabilistic Fusion of Signals

Core Advances in Perception
Core Advances in Perception
Core Advances in Perception

Core Advances in Perception

Core Advances in Perception

Power of data + CNNs

Conversational Speech: *Switchboard* challenge

WER %

Human-level

2009
Understanding Human Cognition
Studies of Attention & Memory

Memory
Attention
Judgment

Human Cognition

Opportunity: Complementary Computing
Direction: Augment human cognition
Models of Attention

Predict Cost of Interruption

H., Apacible, Learning and Reasoning about Interruption, ICMI 2003
Predict Cost of Interruption

H., Apacible. *Learning and Reasoning about Interruption*. ICMI 2003

Leveraging Models of People
Priorities (1999)

Learn to sort & route email by urgency

Priorities (1999)
Learn to sort & route email by urgency

Notification Platform (2000)

Generalize to multiple sources & endpoints

Context Sources

Notification Preferences

Notification Manager

Context Whiteboard

XML Notification Schema

XML Device Schema

Email
Messenger
Telephone
News
Financial
DocWatch
Background Query
Lookout
Error Messages

Desktop Office
Desktop Home
Pocket PC
Cell Phone
Voicemail

Notification Platform (2000)

Generalize to multiple sources & endpoints


Video: Gates keynote & demo, CHI 2001
Models of Surprise

Models of Memory

Human Cognition

Memory
Models of Memory Landmarks


---


Models of Memory Landmarks

*Lifebrowser (2004)*

Models of Memory Landmarks


**Video: Lifebrowser**

Rich timeline of predicted memory landmarks:
- Meetings
- Photos/videos
- Activities
- Locations

*Multimodal training data*
Models of Memory Landmarks


Video: Lifebrowser
Forgetting & Reminding: *Jogger*

\[ p(\text{Forget } x_i | E) \]

Predict user forgets \( x \)

\[ p(\text{Relevant } x_i | E) \]

Predict \( x \) is relevant

\[ p(\text{Cost at } t_o | E) \]

Predict cost of notification at \( t \)

Exp. value of reminder \( x \)

Reminders

Toward Deeper Collaborations
In our Lifetimes?
Situated Interaction

Planning
Understanding
Contributions
Engagement

Beliefs, intentions, plans
Activities & events
Actors, objects in space & time
Situated Interaction Project

Study: Tasks undertaken by receptionists
Situated Interaction Project

Entities, relations, intentions over time

Track conversational dynamics
Make turn-taking decisions

Active

Suspended

Other interaction

System

User

Active interaction

Suspended interaction
Behavioral control

Dialog management & Interaction Planning

Speech Synthesis

Output Management

Conversational Scene Analysis

Behavioural control

Dialog management & Interaction Planning

Tracker

Speech Recognition

Avatar Synthesis

multi-core PC

WORLD

actions

context

AGENT’S OBSERVATIONS

time

A_t

C_t

O_t
The Receptionist

Video: The Receptionist

Studies of Engagement

Video: Multiparty engagement

Studies of Engagement ...in the Open World

Studies of Engagement ...in the Open World

Decisions about Turns in Multiparty Collaboration

Decisions about Turns in Multiparty Collaboration

Looking to the Future: Directions

- Machine Intellect
- Human Intellect

Directions:
- Planning
- Understanding
- Contributions
- Engagement

Beliefs, intentions, plans
Activities & events
Actors, objects in space & time
New Applications & Services
The Assistant

- Face ID
- Vmail
- Calendar
- Room acoust.
- Email
- Location

Multiparty Engagement & Dialog

Prediction about *presence*
Prediction of *cost of interruption*
Prediction about *forgetting*
Prediction of *message urgency*
The Assistant

Video: Approaching the Assistant
The Assistant ...in the Open World

Video
Ecosystem of Collaborating Intelligences

New Types of Coordination
Ideal Fusion of Human & Machine Intellect

Example: Labeling Sloan Digital Sky Survey

Ideal Fusion of Human & Machine Intellect

Example: Labeling Sloan Digital Sky Survey

Machine perception

Human perception

Machine learning, prediction, action

Ideal fusion, routing, stopping

Mix of Initiatives on Physical Tasks
Mix of Initiatives in Surgery

Padoy and Hager. "Human-machine collaborative surgery using learned models." ICRA 2011
Mix of initiatives on road
Example: Tesla Autosteer
Mix of initiatives on road

Example: Tesla Autosteer

Autosteer is no longer steering Model S confidently.
Take over steering immediately.
Advances in Perceptual Capabilities, Competencies, and Pipelines
Advances in Perceptual Capabilities & Pipelines

Figure 1. An illustrative example of our approach’s pipeline.

Advances in Perceptual Capabilities & Pipelines

Figure 1. An illustrative example of our approach’s pipeline.

Advances in Perceptual Capabilities & Pipelines

Figure 1. An illustrative example of our approach’s pipeline.

New Interactive Sensing & Capabilities

Example: Real-time hand tracking

Video: Recognizing Subtleties of Hand Pose
Toward Fluid Natural Dialog & Coordination
Challenge: Timing of Dialog Actions

Dialog decisions under uncertainty

Bohus, H. Decisions about Turns in Multiparty Conversation: From Perception to Action, ICMI 2011
Challenge: Natural Backchannel

Video

Pejsa, Bohus, Cohen, Saw, Mahoney, H. *Natural Communication about Uncertainties in Situated Interaction*, ICMI 2014
A Grand AI Challenge:
General Situated Collaboration
General Situated Collaboration

General Situated Collaboration

Domain commonsense
Social skills
Models of human cognition
Natural language processing
Dialog
Machine Learning
Inference
Planning
Machine vision
Speech recognition
Acoustical Analysis
General Situated Collaboration

Domain commonsense
Models of human cognition
Dialog
Inference
Planning
Speech recognition
Acoustical Analysis
Machine vision
Machine Learning
Natural language processing
Social skills
General Situated Collaboration
General Situated Collaboration

- Coordination
  - Domain commonsense
  - Social skills
  - Models of human cognition
  - Natural language processing
  - Dialog
- Machine Learning
- Inference
- Planning
- Machine vision
- Speech recognition
- Acoustical Analysis
- Reflection & Learning
Critical Enablers:
Tools, Platforms, Infrastructure
Enabling Fast-Paced Prototyping

Video: Hackathon project on sight for the blind

Tools enable fast-paced exploration and prototyping
Getting off the ground
Critical Connections & Collaborations
Andrea (and Monica & Chad)

Michael Cohen  James Mahoney