Using Natural Language to Enhance Mission Effectiveness

Anna C. Trujillo
NASA Langley Research Center

Erica Meszaros
Eastern Michigan University
Outline

- Introduction and background
- Acceptance of using voice
- Intent inference
Introduction & Background

- Current interaction with a drone requires the operator to understand the specifics of the controller and the drone’s dynamic behavior
- Not a natural and higher level teaming relationship
  - Increased workload
  - Decreased situation awareness
  - Decreased trust
- Natural language may increase collaborative teaming
Options and commands were either voice or mouse/keyboard input:
- CMU Sphinx4
- Defined dictionary

Measured:
- Input correctness
- Input time
- Workload ratings
- Subjective ratings
- Subject comments
**Options Input Times**

- *Option* voice input took slightly more time

- Speech recognition system's parsing time for longer phrases may have increased overall times
• *Command* voice input took longer than mouse input
  – Longer phrases took more time

• Use mouse or touch input for mission critical or safety related commands
Voice input lower overall workload
Especially for:
- Mental
- Effort

Frustration about equal
- Time for voice command to register
- No indication on screen until something changed that system was parsing voice command
  - Lacking intent information
Subjective Ratings and Comments

- Slight overall preference for mouse input
  - Except for Responsiveness
    - “Tedious to move the mouse around”

- Largest differences are for Commands
  - Use mouse or touch input for mission critical or safety related commands
Initial Voice Usability Experiment Summary

- Voice took longer to input information
  - Longer phrases took longer
- Voice has slightly lower workload
  - Frustration about equal with mouse input → No indication that voice system was working
- Subjective preferences indicated mouse input preference
  - Critical input commands had lowest preference in using voice input

- Voice input acceptable to non-critical input
- Mouse/keyboard-touchscreen preferred for critical input

Inferring Commander’s Intent by machine may further increase teaming
Predicting Commander’s Intent

• Most users (of computers, autonomous systems, and technology in general) verbalize while working with machines
  – Especially true for members of teams with multiple humans
• Often verbalizations take the form of imprecise questions
  – “What’s it doing now?”
• Can we predict the *Commander’s Intent* and provide desired information on UAV behavior based only on such simple verbal questions?
Latent Semantic Analysis (LSA)

- Well-established tool in computational linguistics
- Determines the degree of semantic relationship between two pieces of language (documents, verbal utterances, etc.)

Methodology:
- Create term-document matrix of all words and utterances in the corpus
- Decompose using singular value decomposition to produce a similarity matrix
- Use multidimensional scaling to plot these similarity values graphically
- The closer two documents are, the more closely semantically they are related
LSA Example

Utterance 1: “One fish two fish red fish blue fish”
Utterance 2: “Black fish blue fish old fish new fish”
Utterance 3: “This one has a little star this one has a little car”
Utterance 4: “Say! What a lot of fish there are”

<table>
<thead>
<tr>
<th></th>
<th>U1</th>
<th>U2</th>
<th>U3</th>
<th>U4</th>
</tr>
</thead>
<tbody>
<tr>
<td>one</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>fish</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>two</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>red</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>blue</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>black</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>old</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>new</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>this</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>one</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>has</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>a</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>little</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>star</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>car</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>say</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>what</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>lot</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>of</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>there</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>are</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Generating the Semantic Map

- Observed language used by human operators while working with UAVs at NASA Langley’s Autonomy Incubator
- Analyzed data to produce a semantic map for UAV operation
  - A predefined semantic space enables better predictions
  - Semantic map can be continually trained
- LSA carried out in R Statistical Programming Language
Semantic Map of the Autonomy Incubator
Semantic Map of the Autonomy Incubator
Clustering on the semantic map is defined by different languages.

New documents or utterances can be mapped to an existing semantic map. The semantic context of the new utterance can be predicted based on which cluster is closest to the newly mapped utterance.

The ability to predict the semantic area of an utterance can be applied to the prediction of the content area of questions, such as “What’s it doing now?”

LSA allows for the prediction of Commander’s Intent for UAV operations.
Predicting Commander’s Intent – Planned Research

1. User’s verbal interactions with UAV transcribed using CMU Sphinx4
2. System triggered when the user asks a question
3. Question and immediate verbal context mapped to existing semantic map
4. Determine closest cluster to the newly mapped information
   - Information associated with this closest cluster provided to user
5. Correctly interpreted utterances added to existing semantic map to further define semantic sphere

• Autonomous agent able to answer back appropriately to the question “What are you doing”

   Based on:
   - Mission context
   - Previous utterances