Should Intelligent Agents Listen and Speak to Us?

James A. Larson
Larson Technical Services
jim@larson-tech.com
Expectations for Human Agents

- Use multiple means of communication
- Recognize me
- Learn my preferences and choices
- Understand my requests
- Resolve ambiguous requests
- Identify required resources
- Perform my requests
Expectations for Human Agents

Automated

• Use multiple means of communication
• Recognize me
• Learn my preferences and choices
• Understand my requests
• Resolve ambiguous requests
• Identify required resources
• Perform my requests
Use Multiple Means of Communication

- Speak and listen
- Text
- Illustrations and pictures
Text Capture

Speech Recognition

Keyboarding

Hand Writing Recognition

Normal alphanumeric gestures

Punctuation mode - tap to enter, backspace to exit

Extended character mode - prefix with...
Pickers and Pallets
Pattern Recognition

UPC Code

QR Code
Multiple Ways to Present Information

Maps

Charts

Video

Text Documents
Use Multiple Means of Communication

• Speak and listen
• Text
• Illustrations and pictures

Use the appropriate mode (modes) of communication
Recognize Me

• User identification devices
  – Something the user has
  – Something the user knows
  – Something about the user

• Combination increases security
Recognize Me:
Something the User Has
Recognize Me: Something the User Knows

• Account names and passwords
  – Complex strings of digits and characters
  – Must be changed frequently
  – Difficult for user to remember

• Challenge dialog:
  – System: what is your mother’s maiden name?
  – User: Lewis
  – System: What is the name of your first pet?
  – User: Fido
  – System: Apply your secret function to the number 7
  – User: 4
  – System: Approved to access
Recognize Me: Something About the User

- Signature Recognition
- Speaker Recognition
- Fingerprint Recognition
- Face Recognition
Learn My Preferences and Choices

- User Profile
- History
Understand My Requests: Semantic Interpretation

• To create smart voice user interfaces, we need to extract the semantic information from speech utterances

• Example:
  – Utterance: “I want to fly from Boston to Chicago”
  – Semantic Interpretation:
    
    ```
    {
        origin: "BOS"
        destination: "OHR"
    }
    ```
Semantic Interpretation

Grammar with Semantic Interpretation Scripts

ASR
Semantic Interpretation Processor

Application
VoiceXML Interpreter

<submit>

I want to fly from Boston
I want to fly from Boston

<Application>

<VoiceXML Interpreter>

<submit>

<item>
from Boston
<tag>$\.origin= "BOS";\</tag>
</item>

<Grammar with Semantic Interpretation Scripts>

<ECMAScript object>

<ASR>
Semantic Interpretation Processor

<text>
Semantic Interpretation

I want to fly from Boston

Semantic Interpretation Processor

Text

ECMAScript object

VoiceXML Interpreter

Application

Grammar with Semantic Interpretation Scripts

{ origin: "BOS" }

from Boston
<tag>$.$.origin= "BOS";</tag>
</item>
Semantic Interpretation

Grammar with Semantic Interpretation Scripts

ASR
Semantic Interpretation Processor

ECMAScript object

{ origin: "BOS" }

To Chicago

<submit>

Application

VoiceXML Interpreter

text

To Chicago

{ origin: "BOS" }
Semantic Interpretation

Grammar with Semantic Interpretation Scripts

To Chicago

Application

<submit>

VoiceXML Interpreter

text

ECMAScript object

ASR

Semantic Interpretation Processor

{ origin: "BOS" }

To Chicago
Semantic Interpretation

```xml
<item>
  from Boston
  <tag>$.$from="OHR" tag>
</item>
```

Grammar with Semantic Interpretation Scripts

```
{ origin: "BOS"
  destination: "OHR"
}
```

To Chicago

```
{ origin: "BOS"
  destination: "OHR"
}
```

Application

VoiceXML Interpreter

ASR Semantic Interpretation Processor

ECMAScript object

text
Resolve Ambiguous Requests

• Consult my user profile and histories
• Ask me if you are uncertain
Resolve Ambiguous Requests: Consult My User Profile and History

User Profile
- Preferred airline: United

Travel History
- From Chicago to Portland, Maine
- From Chicago to Portland, Oregon

I want to fly from Chicago to Portland
Do you want to fly to Portland, Maine or Portland, Oregon?
Resolve Ambiguous Requests: Ask Me if You Are Not Certain

I want to fly to Austin

Do you want to fly to Austin or Boston?
Identify Required Resources

Local Resources

Remote Resources
Perform Required Tasks

• Carry out instructions accurately and efficiently
Expectations for Human Agents

Automated

• Use multiple means of communication
• Recognize me
• Learn my preferences and choices
• Understand my requests
• Resolve ambiguous requests
• Identify required resources
• Perform my requests
Should intelligent agents listen and speak to us?

Yes!

And a whole lot more.