USING ASR WITH HUMAN ASSISTED UNDERSTANDING IN MOBILE APPLICATIONS

Yoryos Yeracaris (aka “yy”)

CTO
Interactions Corp Profile

- Incorporated in 2004: Based in MA; Operations Centers in IN and TX
- 10 years of live operational experience with a patented Human Assisted Understanding (HAU) system
- Offering: Enterprise Virtual Assistant Solutions
- Business Model: Private Cloud Service Provider
  - MPLS, VOIP, Web API, REST, WebRTC*
- Providing Voice and Text based application for Customer Care
  - 30+ Fortune 500 production customers
An Human Assisted Understanding (HAU) system uses a combination of automation, such as ASR and NLP, and a symbiotic combination of human and machine to recognize fragments of conversations.

- What it is not, is listening and tracking and rescuing conversations that are “off track”

Mobility environment presents a set of issues, some new, but mostly old:

- We analyzed thousand of calls from primarily mobile voice application, where the ASR was having trouble recognizing or recognized with low-confidence:
  - Noise is surprisingly higher than land-line approximately 2 X
  - Poor voice quality: background noise, speaker phones, Bluetooth microphones, microphone quality (in-vehicle voice), etc.
- Recognition was improved through acoustic model adaptation, but there was still a gap
ASR AND MOBILITY CONDITIONS

• Is it the Codec?
  – More data should result in better results!
  – According to a recent article in the IETE, the ability of the ASR with G.722 Codec and G.711 is virtually the same. Recognition rates for TIMIT tests are more or less the same.
  – G.711 is 64 Kbps, G.722 is ~16 Kbps (Kbps can be variable)
ASR AND MOBILITY CONDITIONS

- Mobility offers a rich set of modes to communicate:
  - Audio, Visual, Text
- Speech in mobile applications:
  - WALKING
    - It is hard to text and walk
    - Looking at a smart device while walking is sort of Darwinian, especially in Manhattan
  - DRIVING
    - Currently 41 states and D.C. band texting while driving
    - Currently 12 States and D.C. band hand held devices while driving
RECOGNITION ACCURACY FROM AN OPENING “OPEN-ENDED PROMPT”

APPLICATION with ASR alone
HUMAN adapts to COMPUTER

"HOW CAN I HELP YOU?"
• What is acceptable?
• False Acceptance? Re-prompting?
• Out of grammar? Noise
• Humans to adapt to ASR

WHAT ASR UNDERSTANDS
with lower confidence
~68%

WHAT ASR UNDERSTANDS
with high confidence
~45%

APPLICATION with HAU
COMPUTER adapts to HUMAN

"HOW CAN I HELP YOU?"
HAU Understanding
• Improved Understanding Accuracy
  Ears can discriminate voice from noise
• Built-in Classification capability
• Flexible to human behavior

WHAT ASR UNDERSTANDS CORRECTLY
with high confidence
~95%

WHAT ASR UNDERSTANDS CORRECTLY
with lower confidence
~45%

Errors

WHAT PEOPLE SAY

• Improved Understanding Accuracy
• Ears can discriminate voice from noise
SO WHERE CAN WE GO WITH UNDERSTANDING AND MOBILITY?

• A HAU system helps with understanding, making mobility applications more conversational, enabling more automation of customer service transactions

• Automation and HAU can be dialed to optimize service requirements and costs
  – It is not just making a better application thru better understanding, but making an application more like a “real virtual assistant”

• HAU does not listen to, nor track the conversation, humans and computers work symbiotically to help with Natural Language Understanding
  – Google uses, retrospectively, thousands of humans to assist their search-engine system, combining software with human assistance to improve their search capabilities!
SO WHERE CAN WE GO WITH UNDERSTANDING AND MOBILITY?

• Today, for speech, a VUI forces users to adapt their speech to what they expects a computer can capture and understand. For text do we want the same?

• Challenges that face mobility applications
  – Quickly and unpredictably changing products and services with new terminology and tasks
  – Unconstrained recognitions
  – Accuracy and Understanding – is completely automated recognition, “The Holy Grail” really in sight?

• HAU system has benefits beyond improved accuracy, including faster time to market:
  – More conversational systems, i.e. more like people talking to people, not computers
  – Starting up requires virtually no training and tuning
  – Allows for more effective tuning and training
  – Ability to quickly respond to changing/shifting environments
HUMAN ASSISTED UNDERSTANDING SYSTEM

• Where can automated understanding technology go on its own?
  – More data and context can help with recognition
  – Personalization, such as personal acoustic models, personal preference context

• When all is said and done…
  – There is still significant gap in recognition
  – HAU Systems make the conversation with your customer successful:
    • Significant increase of recognition accuracy
    • Recognize multiple pieces of information, like address collection all at once!
    • Collections of email address? i.e. the problems of sound, “ee” sound, “ay” sound, etc.
    • Help with multiple hypotheses resolution
    • More interactions can be automated, reducing costs
    • Your customer time, your customer satisfaction and their fallout are important to customer service applications
HOW PEOPLE FEEL ABOUT HAU SYSTEMS

• “Instead of having to go through an interminable series of push-button choices or stick to overly simple verbal commands, you can talk just as you would to a human representative—and, surprisingly, it actually works.” - Rachel Metz, Technology Review

• It is so much easier than getting a hold of a person and then telling them the account number, it is just simple it is great. Just what you needed to do at [Interactions Client]. Don't change it now that you have got it right. – Customer Survey Verbatim

• [Interactions Client] automated system is very HER-like. Scary. - Twitter

• Whoever is providing [Interactions Client] automated telephone system is offering a great system. First one I’ve used that is relatively usable - Twitter