A Handheld Flexible Phrase Translator

Melvyn Hunt & Yoon Kim

Overview

- The central idea
  - large set of basic phrases with many variants
- Principles in designing the phrases
  - make sure the response will be comprehensible
- Some technical details
- Grammar of the variants for one basic phrase
- Extension to two-way translation
- Live Demo
The Central Idea

- To equip a traveler with a handheld device providing voice-to-voice translation of a large set of phrases for specific scenarios
- But the user can’t be expected to know what the phrases are
- Solution: make the phrases highly flexible
  - map the many forms to a single “base form”
  - translate the base form

Mode of Operation

- User speaks a phrase in his/her native language
- The translator displays up to 5 base phrases
- The user selects the phrase that’s the best translation of the underlying intent
  (if the confidence on the top choice is high enough, the translator may carry on without confirmation)
- The translator accesses the translated phrase corresponding to the selected base phrase
  - and plays out a spoken recording of the translation
Some Constraints on the Phrases

- No use translating a question if the user can’t understand the answer!

- So translated phrases must be carefully constructed:
  - “please point to the restrooms”
  - and not “where are the restrooms, please?”
  - “please write down the cost of the room for three nights”
  - “please point to the bus stop on this map”

Technical Background

- Novauris has developed compact technology for spoken access to very large lists
  - works on smartphones, PDAs and other mobile devices
  - especially effective when spoken items are long

- We have extended the technology to allow efficient processing of optional components (e.g. “please”)
- and to embed sets of alternatives in an item
  (e.g. “I’d like an apple / an orange / a tomato … juice [please]”)

- For more technical details, please see our talk:
  Recognition from Large Lists in the Technical Track session

  “Effective use of speech technology” 2:45 to 4:00 this afternoon
Grammar for the base phrase: “Do you have a reservation?”

- [and] do you have a >reservation [please]
- [and] have you (got | made) a >reservation [please]
- [and] >RI >if you (have [got] | [have] made) a >reservation [please]
- >RI >if you (have [got] | made) a >reservation [please]
- >RI >if >you've (got | made) a >reservation [please]

Macro definitions:
>RI* = [please] (can | would | could) you [please] tell me [please] | (may | might | could | can) I ask [you]
>if = (if | whether)
>reservation = (reservation | booking)
“RI” = “Request Information"

NB The ASR is quite robust, and phrases outside the anticipated set are frequently correctly classified

Extension to Two-Way Mode

- Novauris ASR is speaker-independent
  - so if ASR is provided in both languages, both parties can speak in their native language
  - but the phrases from the experienced party must be designed to evoke simple responses from the inexperienced party
    - for example, an immigration officer should say:
      - “Is the purpose of your visit business or pleasure?”
        NOT “What’s the purpose of your visit?”
      - Anticipated responses:
        “Business”, “Pleasure”, “Both”, “Neither” and variants
Demonstration

English ↔ Korean