Natural Language Processing for Joint Fire Observer Training

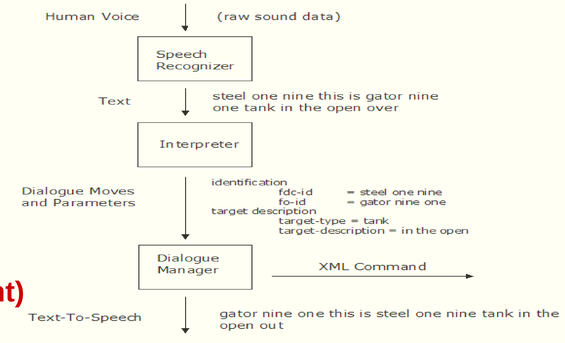
Antonio Roque, Kallirroi Georgila, Ron Artstein, Kenji Sagae, David R. Traum
Institute for Creative Technologies, University of Southern California

JFETS (Joint Fires and Effects Training Simulator at Fort Sill)

- Three training modules: Open Terrain, Urban Terrain, Close Air Support.
- Trains Forward Observers and Joint Terminal Attack Controllers.
- Supports joint Call for Fire (CFF) and Close Air Support (CAS) missions.

IOTA (Intelligent Operator Training Assistant)

- Automatically plays the role of the operator.
- Currently only handles CFF missions.



Overview of IOTA components

Differences between CFF and CAS dialogues

- CFF missions follow a controlled structure.
- CFF and CAS missions vary in vocabulary (see highlighted words in example dialogues).
- CAS missions contain a richer syntactic and semantic structure, requiring sophisticated techniques for automatic extraction of information.

Our research answers the following questions:

1. How well does automatic speech recognition (ASR) perform on CFF and CAS dialogues?
2. How can we tell if ASR is good enough for linguistic analysis?
3. What techniques can be used to interpret CAS utterances?

Example CFF dialogue

Trainee: M T O kilo alpha four **rounds target number** alpha bravo one out.

IOTA: **Shot** over.

Trainee: **Shot** out.

IOTA: **Splash** over.

Example CAS dialogue

Soldier: Contact the road **west** coming out of the body of water.

Operator: Uh I see a road leading off to the **southwest**.

Soldier: Call contact on uh two buildings **north** and **south** of the road.

Operator: Uh I contact uh two buildings uh **north** and **south** of the road uh both are near the uh **southwest** corner of the lake.

1. ASR performance evaluation

The following speech recognizers were used:

Cambridge HTK family: HVite, HDecode, AVite, Julius.

CMU Sphinx family: Sphinx 4, Pocket Sphinx.

Speech recognition results (Word Error Rate, lower values are better)

Non-real time	HVite		HDecode		Sphinx4	
	Dev	Test	Dev	Test	Dev	Test
CFF	10	15	11	12	-	-
CFF+CAS	66	57	49	39	76	-

Real time	Julius		AVite		PktSphx	
	Dev	Test	Dev	Test	Dev	Test
CFF	17	14	12	-	7	10
CFF+CAS	61	42	-	-	55	47

2. ASR error simulation

There is a shortage of data for training and evaluating ASR in new applications.

ASR error simulation allows testing of the natural language interpreter and the rest of the system under various error conditions.

Original utterance

I have target two buildings north of the road

above target to buildings forth road

Scrambled utterance with simulated errors

Our algorithm generates simulated ASR errors with a distribution very similar to the distribution of errors observed with a real speech recognizer.

3. Syntactic and semantic analysis of CAS dialogues

Utterance: *Once you get to that village you see a uh almost looks like a martini glass at the south end of the lake.*

```
[ MAIN CLAUSE :
  [ VERB : see
    SUBJECT : you
    OBJECT : a uh almost looks like a martini
             glass at the end of that lake
    ADVERBIAL CLAUSE :
    [ VERB : get
      SUBJECT : you
      PREPOSITIONALPHRASE : to that village ] ] ]
```

Syntactic information from a syntactic parser

```
[ PREDICATE : see (meaning : "view")
  TEMPORAL CLAUSE :
  [ PREDICATE : get (meaning : "move")
    [ THING IN MOTION : you
      DESTINATION : to that village ] ]
  VIEWER : you
  THING VIEWED :
  [ PREDICATE : look (meaning : "seem")
    ADVERBIAL MODIFIER : almost
    SEEMED LIKE : a martini glass ]
  LOCATION :
  [ PREDICATE : end
    VALUE : south
    THEME : of that lake ] ]
```

Semantic information from a semantic role labeler

Acknowledgments: This work was sponsored by the U.S. Army Research, Development, and Engineering Command (RDECOM). The content does not necessarily reflect the position or the policy of the U.S. Government, and no official endorsement should be inferred.