Developing a spoken dialogue system requires paying attention to several critical issues:
- Imperfect Automatic Speech Recognition module leads to task failures.
- Partial Sentence Tree to keep all keywords and some non-keywords.
-.Services used by Spoken Dialogue Understanding.
- Derivation Rules represented as a vector.
- Dialogue Management module require functionality, including capturing intentions and interactional patterns of the user.
- Data-driven approach to decide the Dialogue Act (DA) types.
- Dialogue History for dialogue controlling.

---

## Introduction

Semantic Information and Derivation Rules for Robust Dialogue Act Detection in a Spoken Dialogue System

Wei-Bin Liang, Chung-Hsien Wu, and Chia-Ping Chen*

Dept. of Computer Science and Information Engineering, National Cheng Kung University, Tainan, Taiwan

*Dep. of Computer Science and Engineering, National Sun Yat-sen University, Kaohsiung, Taiwan

---

### Data Collection

- **ASR**: Automatic Speech Recognition
- **PST**: Partial Sentence Tree
- **DR**: Derivation Rules
- **DM**: Dialogue Management

---

### Models for Dialogue Act Detection

**At dialogue turn $t$, given user’s utterance $U_t$ and dialogue historical information $H_t$, the most likely DA**

$$A_t' = \arg \max_{A_t} P(A_t | U_t, H_t)$$

**for all possible word sequence**

- $A_t$ best -1 ASR output
- $U_t \perp H_t$

---

### History Score

Based on a Markov model assumption for the chain of the DAs

$$h(A, H_t) = P(A_t = A | A_{t-1})$$

---

### ASR Score

HTK-based Mandarin speech recognizer

- 297 Lexical words
- 39 dim MFCC
- 86.1% Accuracy

$z$-score to detect the unreliable recognized words

$$z(w) = \frac{f(w) - \mu(w)}{\sigma(w)} < -2$$

**ASR Output**

- Where $\text{ether} \rightarrow \text{Anping Fort}$
- Detected and Substitution

---

### Lexical Score

- Broken into two terms

$$g(A, W) = \max_{A_t} \left( b_{(A,t)} \prod_{t} v(A_t, k_t) \right)$$

**Named Entity Score**

DA $\land$ Keywords (KW) $\alpha$

- Greet, Welcome, Hello
- Spot, Anping-Fort
- Time, Morning

---

### DR-DA Matrix Training

**Partial Sentence Tree Construction**

- **Ref** Where is the Anping Fort
- **NonRef**

**Each Partial Sentence $\sigma$ will be used to extract the Derivation Rules $h_{\sigma}$**

**Data-driven approach to decide DA types**

- Concept of Spectral Clustering Algorithm

**DR-DA Matrix Construction**

- All utterances and transcriptions are used to building this matrix
- Entropy-based normalization

---

### Evaluations

- **Number of DA Type**: (#DA Types) 37 38 39
- **Accuracy**: 82.7 84.3 77.2

**Detection accuracy (%) for the lexical score**

- 40%-SIM 60%-SIM 86.1%-ASR 100%-REF
  - DR-DA: 26.3 47.4 82.9 93.3

**Detection accuracy (%) for eighed history score**

- $\lambda_1 = 0.5$, $\lambda_2 = 0.6$
- Accuracy (%): 84.3 84.6
Where is the Anping Fort

- **Ref**: Stanford Parser
- **NonKW**: Where Filler Spot

Derivation Rules:
- **DR1**: WHADVP (WRB Where)
- **DR2**: SQ (VBP filler)
- **DR3**: WHADVP (WRB Where)
- **DR4**: NP (NNP Spot)

Named Entity Substitution

\[ \sigma = \psi \]