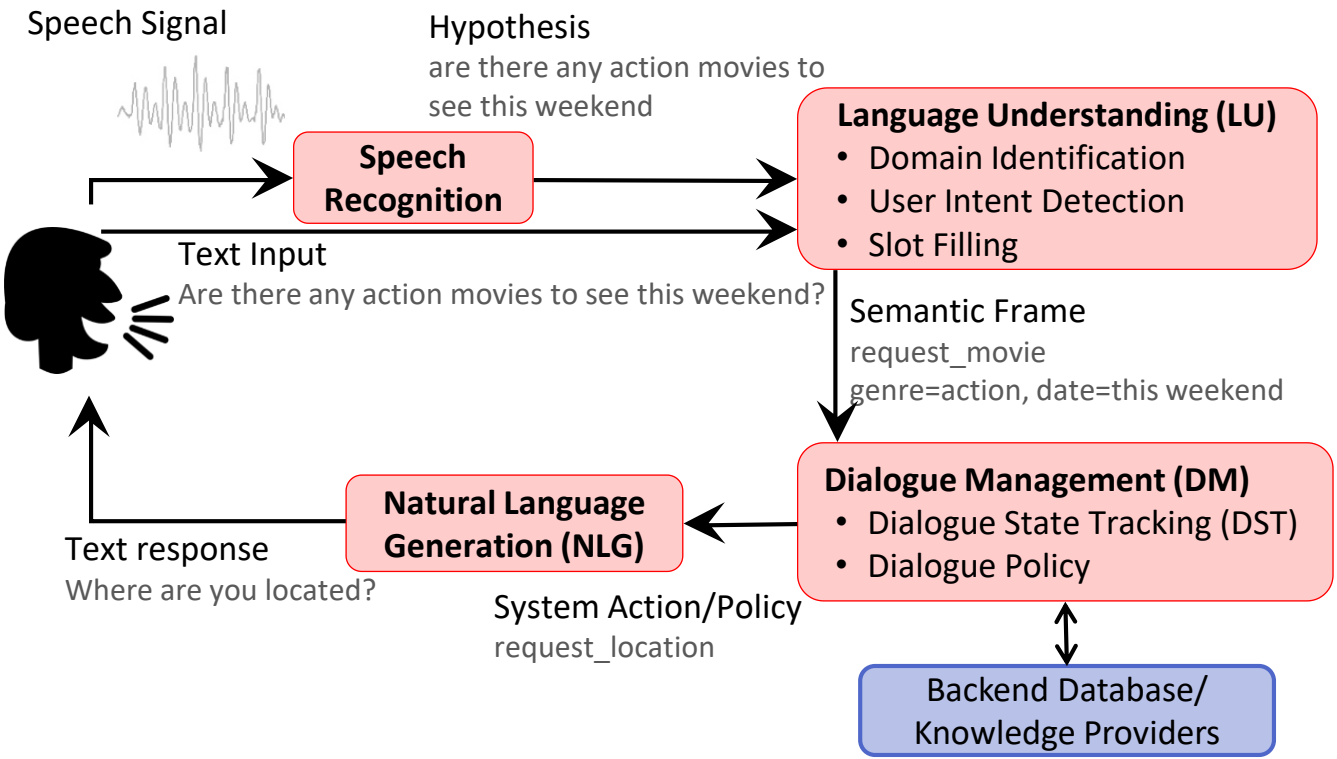
Intelligent Conversational Bot

Poster Exhibition Tips

YUN-NUNG (VIVIAN) CHEN WWW.CSIE.NTU.EDU.TW/~YVCHEN/S105-ICB

國立臺灣大學
National Taiwan University

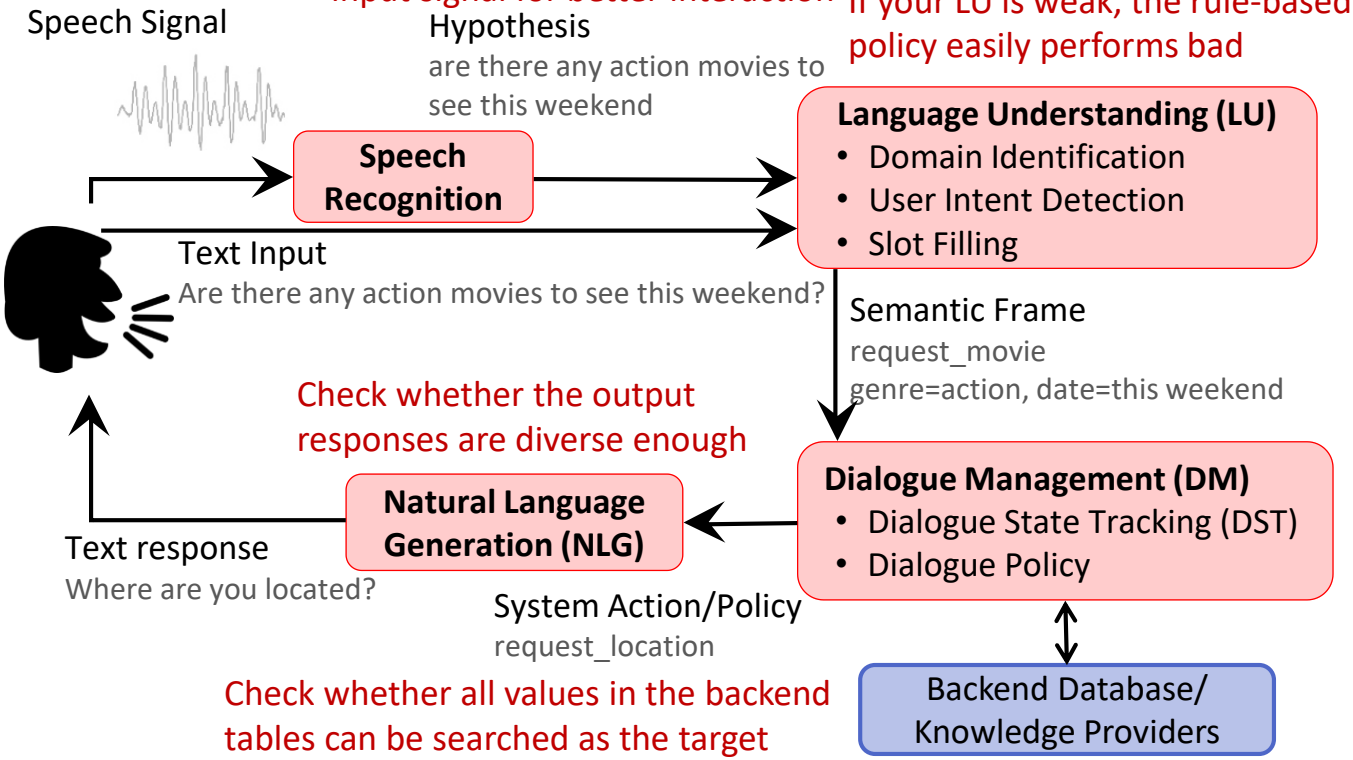
Framework


Speech Signal
Hypothesis
are there any action movies to see this weekend
Speech Recognition
Language Understanding (LU)
• Domain Identification
• User Intent Detection
• Slot Filling
Text Input
Are there any action movies to see this weekend?
Semantic Frame
request_movie
genre=action, date=this weekend
Dialogue Management (DM)
• Dialogue State Tracking (DST)
• Dialogue Policy
Natural Language Generation (NLG)
Text response
Where are you located?
System Action/Policy
request_location
Backend Database/ Knowledge Providers

Framework

If possible, try richer multimodal input signal for better interaction

If your LU is weak, the rule-based policy easily performs bad

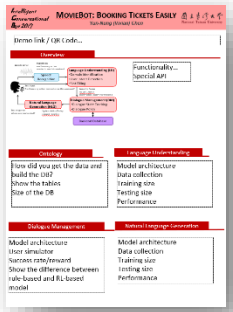


System Improvement

- **Ontology:** check whether all columns in the table can be searched as the target
- **LU:** evaluate the LU to see the *coverage* of the understanding module
  - Testing data should come from real human
    - Provide the system link to collect more dialogues and then annotate them for evaluation
- **DM:** add multi-turn interactions into the simulator for training the RL agent
  - The RL agent should handle misunderstanding better than the rule-based agent
    - Check whether the agent can handle *misrecognized* texts or *misunderstanding*
  - If the RL agent performs worse than the rule agent, increase your system complexity
    - More functionality/backend databases, more complex simulated interactions
    - Please check the strategies this agent applied to make sure your RL agent has increasing performance trend
- **NLG:** improve diverse and interesting responses
- **Multimodality:** try richer multimodality for interesting interactions
  - Emotion recognition, speaker recognition, etc for better greeting

Final Score

- System functionality
  - #tables, #slots, #intents
- System success performance
  - Human testing performance evaluated by TAs
    - ~30 dialogues
    - If the failed dialogues are fixed, we use the refined performance
- Evaluation
  - Correctness and reasonability
    - Testing data should be from real human instead of generated patterns
- Creativity
  - Multimodality usage (e.g. emotion)
  - Diverse/interesting responses
- The poster template can be revised freely [link]
  - Due: 6/17 23:59:59



Milestone 3 / Peer Demo Log

- Improve your system based on the feedback
  - Milestone 3 [link]
  - Peer demo feedback [link]
- Team peer review form
  - Due 6/15 23:59:59

Poster Content (1)

- Demo link / QR code for app
- Input
  - Interaction example
  - Supported APIs (speech, vision, emotion, etc)
  - Functionality your system supports
- Ontology
  - DB tables (size of the DB, #column, #slot, #intent)
  - How did you get the DB data

3 numbers should be close

- LU
  - Model architecture
  - Training data size
  - Testing data size (should come from real human)
  - Performance on testing data (frame accuracy, etc)

Poster Content (2)

- DM
  - Model architecture
  - User simulation summary
  - Trend of the learning curves for rule-based and RL agents (success rate, reward, etc)
  - Show the example with the difference between two agents
- NLG
  - Model architecture
  - Training data size
  - Testing data size (should come from real human)
  - Performance on testing data (BLEU score, naturalness)

Poster Presentation

- 2 minute presentation
  - Supported functions
  - Special features
  - Whole system performance
- 3 minute demonstration
  - Allow the user to test the system

Agenda

- 9 am – 10 am
  - Preparation (poster, system, etc.)
- 10 am – 11:50 am
  - Presentation
- 12 pm – 12:20 pm
  - Lunch break & judge discussion
- 12:20 pm – 1 pm
  - Company sharing
  - Award announcement

Final Report / Code

- Due: **6/25 (Sun) 23:59:59**
- Code
  - README, Requirements
- Report
  - GitHub page [link]
  - Put the poster contents / figures into the page as the report (can be more detailed)