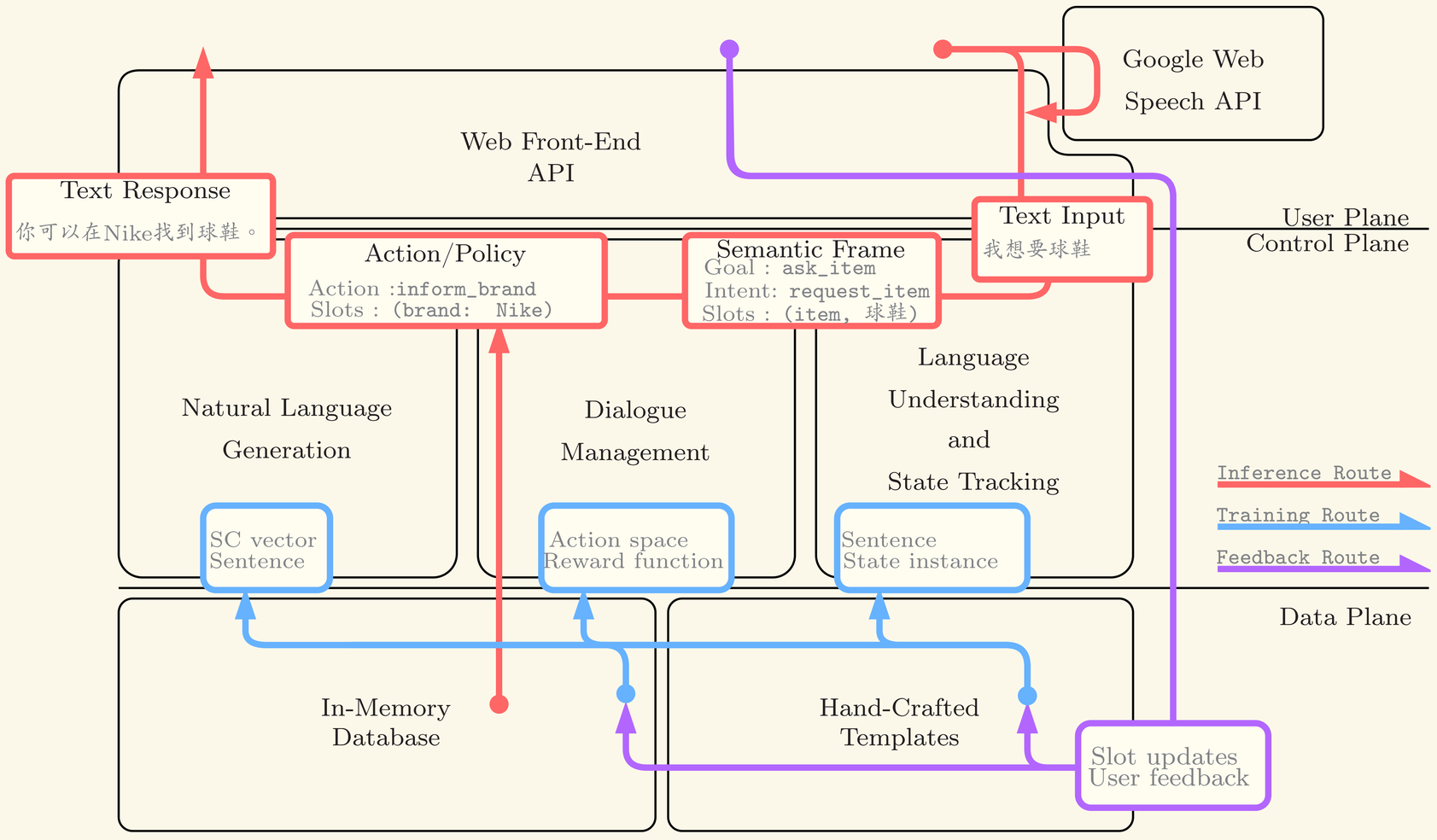
Department Store Bot

Che-Hsien Lin[1], Ian Chen[1], Yen-Chun Fang[2], En-Yu Fan[2], Wei Fang[3]

[1] *Graduate Institute of Electrical Engineering, National Taiwan University*
[2] *Graduate Institute of Electronics Engineering, National Taiwan University*
[3] *Department of Electrical Engineering, National Taiwan University*

National Taiwan University

Overview



Supported Functionalities

- Find a store(s)/restaurant(s)
- List all stores/restaurants that sell a particular item/food
- Show sale/phone/introduction of a store
- Get direction to a store/service

Technical Highlights

- Web Application Interface
- Supports speech input (Google Speech API)
- RL-based Chatbot
- RNN-based LU/State-tracker & LG
- Template-based data generator
- Handle irrelevant sentences in dialogue
- Rapid model update from user feedback

Ontology

- Metadata
  - size of DB: 1058 brands, 2 locations
  - DB source: 新光三越Website
- Brand

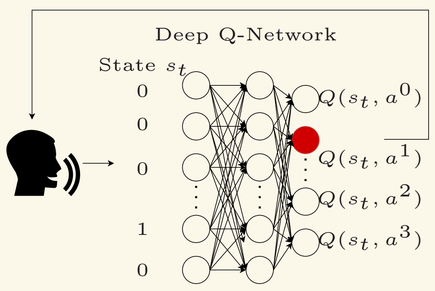
Brand	Category	Introduction	locations	suggestions	activities
Nike	運動用品	現今Nike已成爲...	[location]	球鞋	週年慶8折
Starbucks	cafe	starbucks...	[location]	none	none
⋮	⋮	⋮	⋮	⋮	⋮

- Location

branch	floor	phone
A9	5F	(02)2345-6789
⋮	⋮	⋮

Dialogue Management

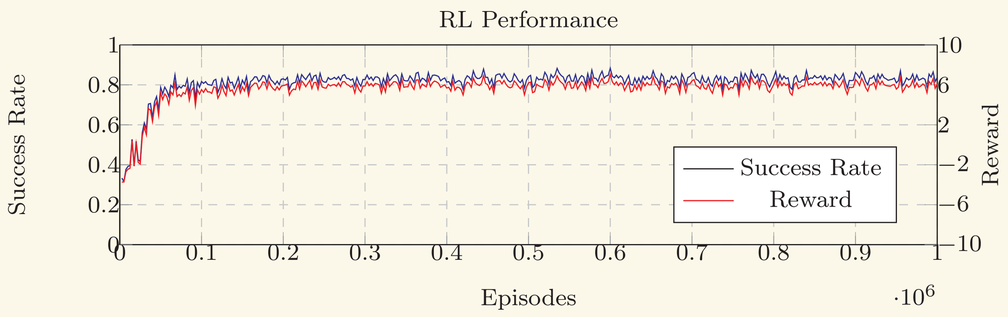
- RL Model: Deep Q-Network (DQN)
  - Fully-connected network
  - Inputs:
    - Tracked state (from DST)
    - 'n-hot' vector representation
  - 40 action-values as outputs



- User Simulator
  - Lets agent learn to minimize the dialogue turns
  - State remains the same unless agent responds reasonably
  - Reward:

Condition	Value	Meaning
Dialogue Success	10	User goal achieved
Dialogue Fail	-10	Fails to terminate dialogue
Others	-2	Responds correctly or incorrectly

- Performance



Rule-based Model Showcase

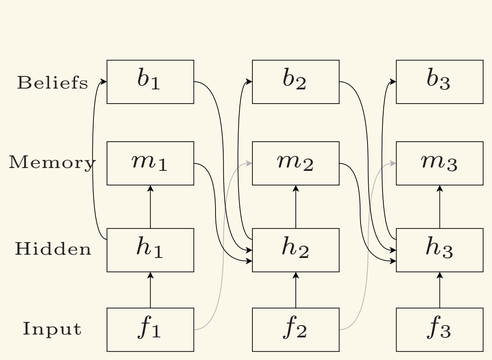
User: 我想買球鞋
Model: 請問你在哪一棟
User: 我在A8
Model: A8可以在Nike買到球鞋

RL-based Model Showcase

User: 你可以告訴我買D'URBAN要往哪邊去
Model: 你在哪一棟的哪一層呢
User: 我在A9的7F
Model: D'URBAN的位置在A8三樓

Language Understanding/State Tracking

- Model: Jordan-type RNN Tracker[1]
- Inputs:
  - Natural language (Bag-of-Words)
  - Previous belief state
  - Previous bot action
- Training Data:
  - 18 user goals
  - 13 intents (ex. `request_item`,...)
  - 7 slot types (ex. item, brand,...)
  - 1.26M sentence templates
  - 10k dialogues sampled from sentences
  - Samples error during training so tracker can learn to correct itself

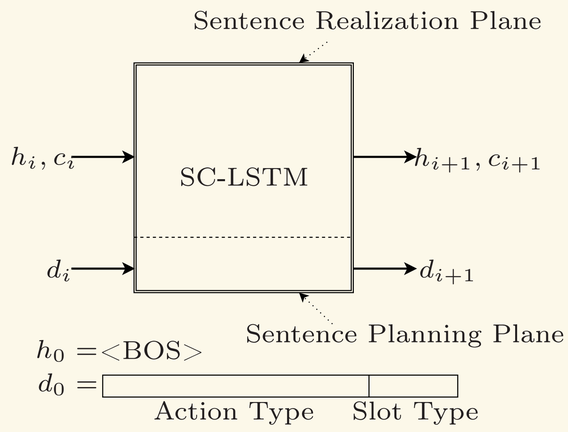


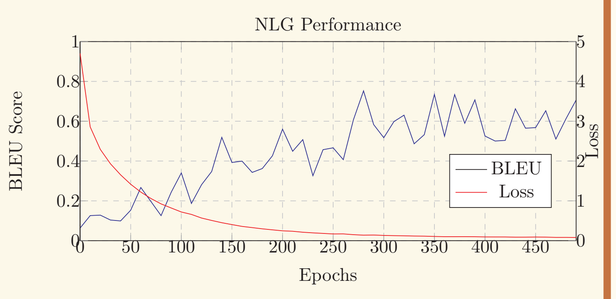
- Testing data:
  - 20 dialogues generated by human

Metric	Intent	Branch	Floor	Food	Item	Restaurant	Brand	Service
L2 distance	0.001	0.0	0.0	0.0	0.0	0.004	0.019	0.0
Precision	99.9	100	100	100	100	99.1	96.6	100
Recall	99.9	100	100	100	99.9	99.1	95.1	100
F1 score	99.9	100	100	100	100	99.1	95.8	100

Natural Language Generation

- Model: modified from Semantic-Conditioned LSTM[2]
- Inputs:
  - Semantic vector representing action types and slot types
- Training Data:
  - Generated by template
  - We can achieve response diversity by mapping an action to many sentences with the same meaning.





References

[1] Mrksic et al. Multi-domain Dialog State Tracking using Recurrent Neural Networks. In *ACL 2015*.
[2] Wen et al. Semantically Conditioned LSTM-based Natural Language Generation for Spoken Dialogue Systems. In *EMNLP 2015*.