

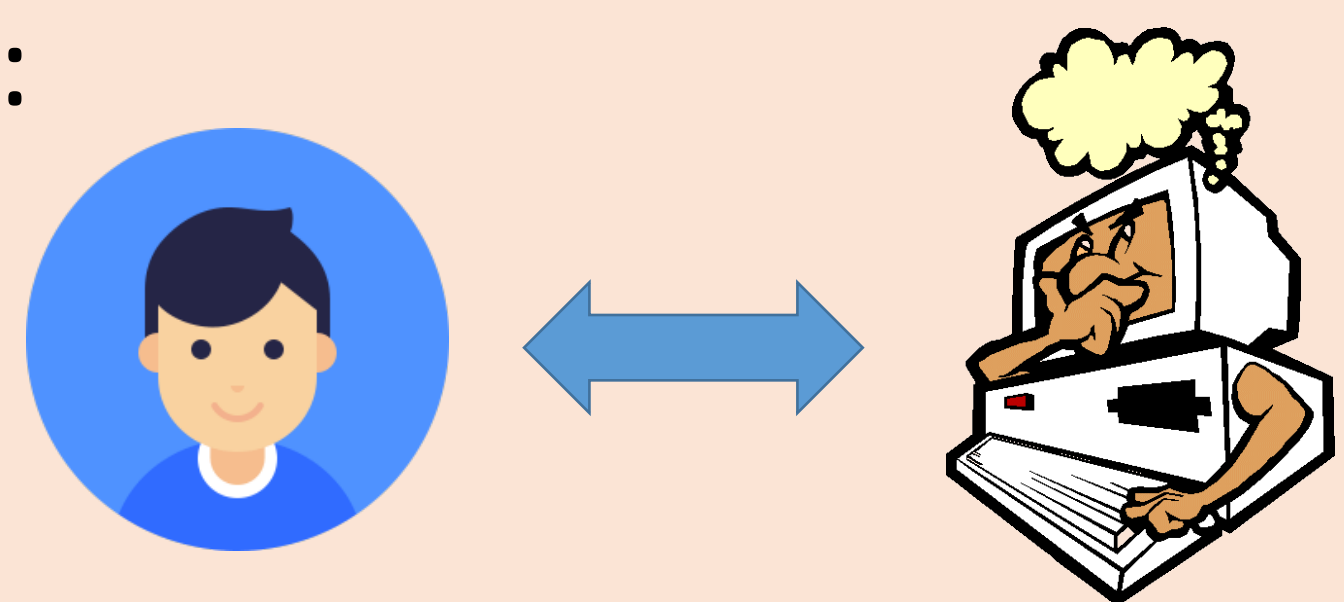
Joint Learning of Interactive Spoken Content Retrieval and Trainable User Simulator

Pei-Hung Chung*, Kuan Tung*, Ching-Lun Tai, Hung-Yi Lee

1. Introduction

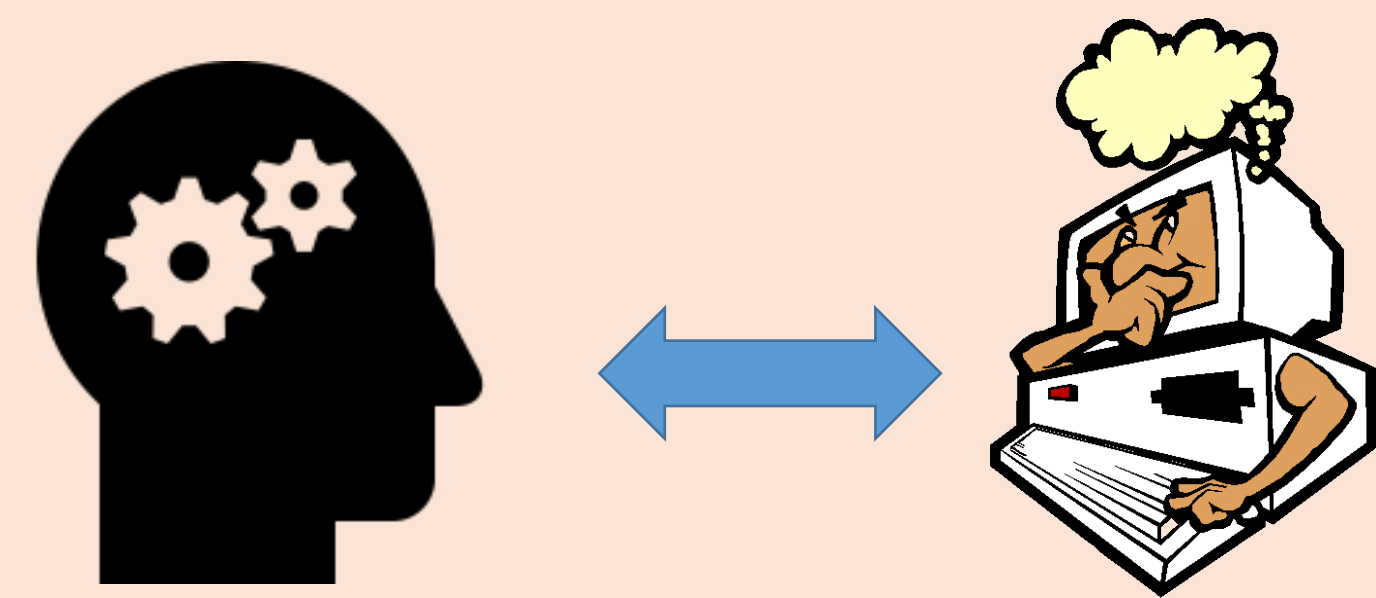
- User-machine interaction is **crucial** for spoken content retrieval.
- Motivation: build an interactive system with a **reliable** user simulator.

Ideal case:



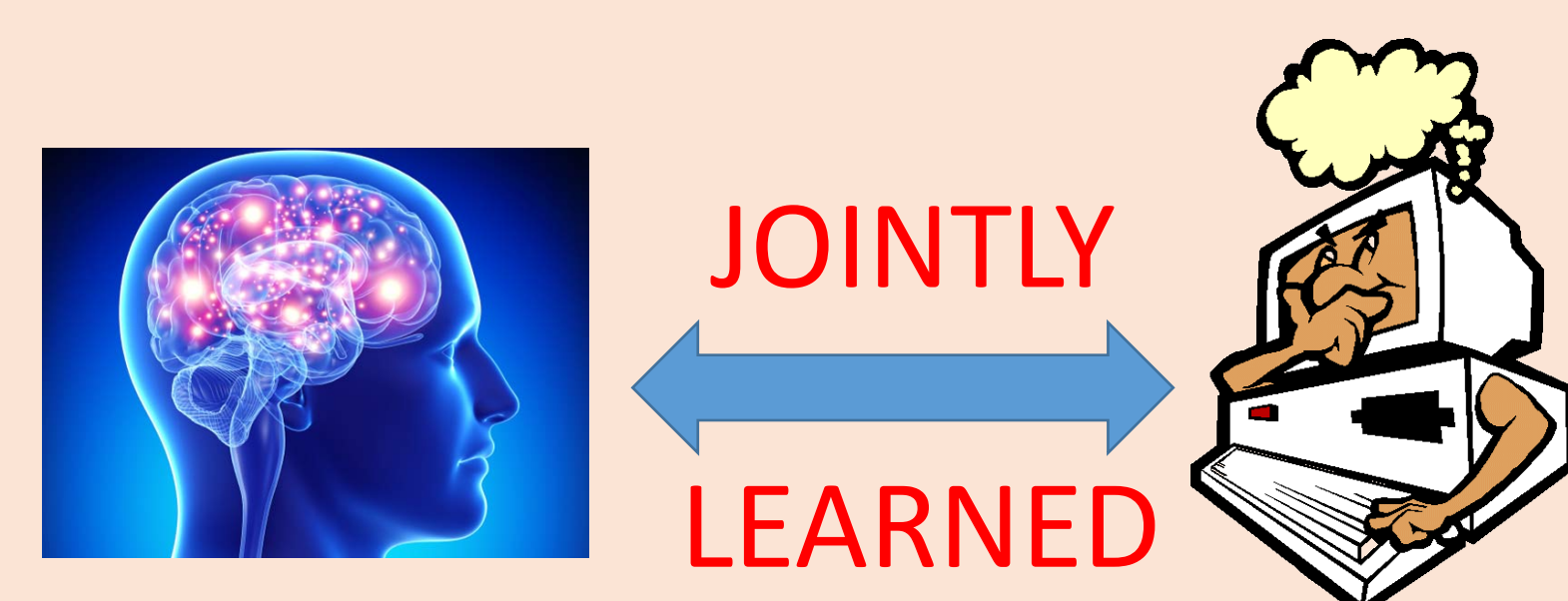
Previous method:

RULE-BASED USER SIMULATOR

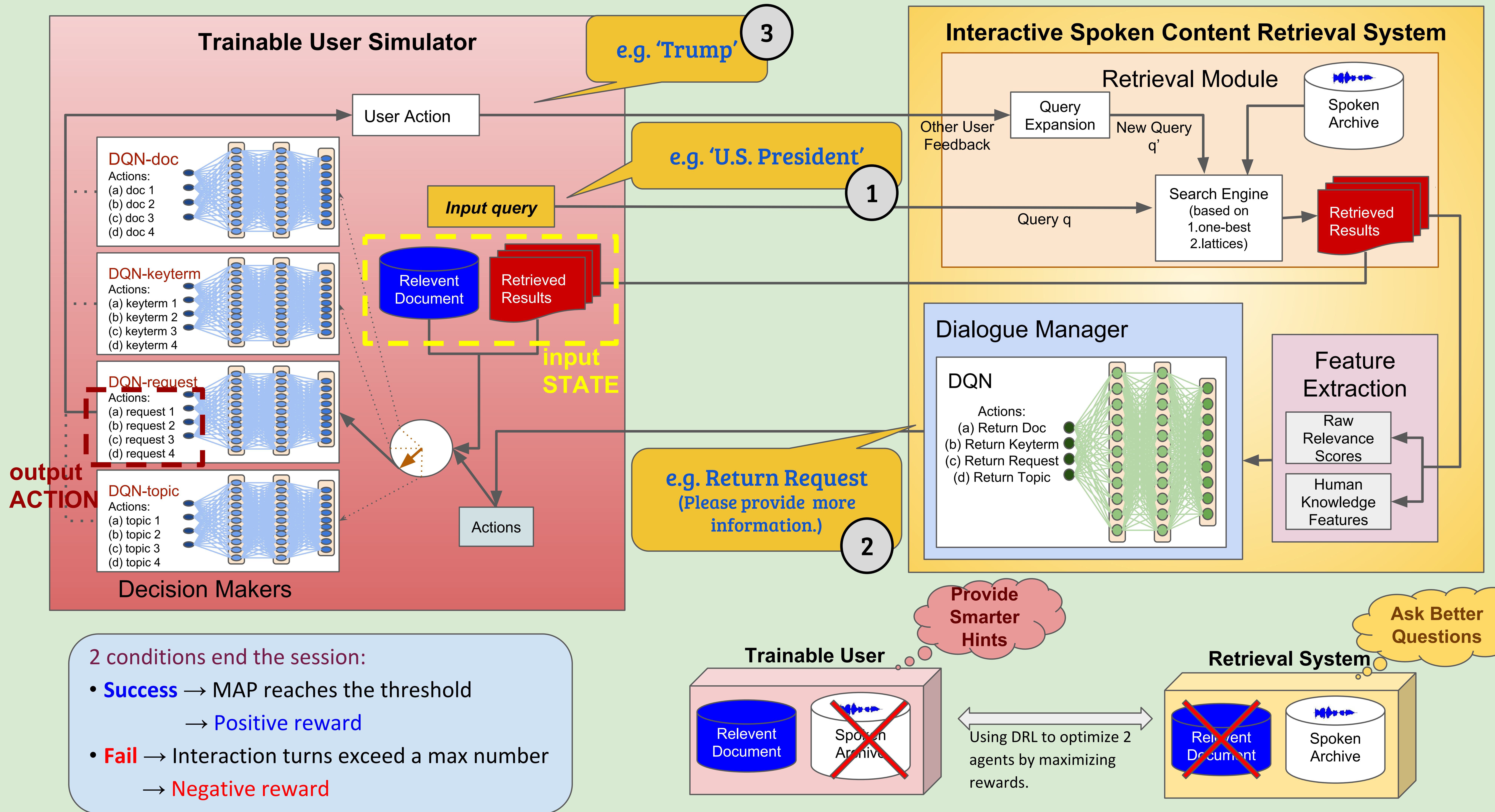


Proposed method:

TRAINABLE USER SIMULATOR



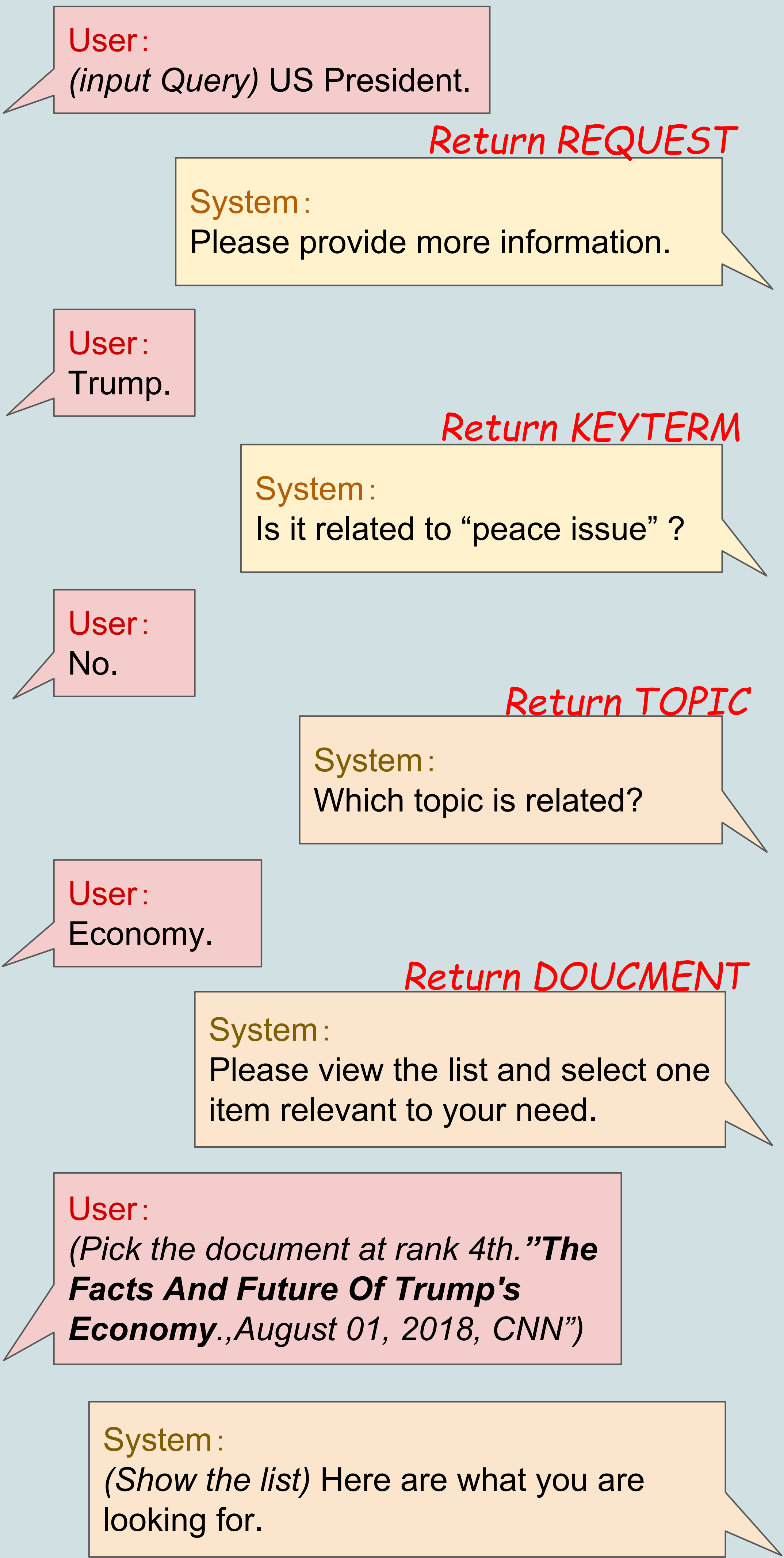
2. Proposed Approaches



2 conditions end the session:

- Success** → MAP reaches the threshold → Positive reward
- Fail** → Interaction turns exceed a max number → Negative reward

SCENARIO



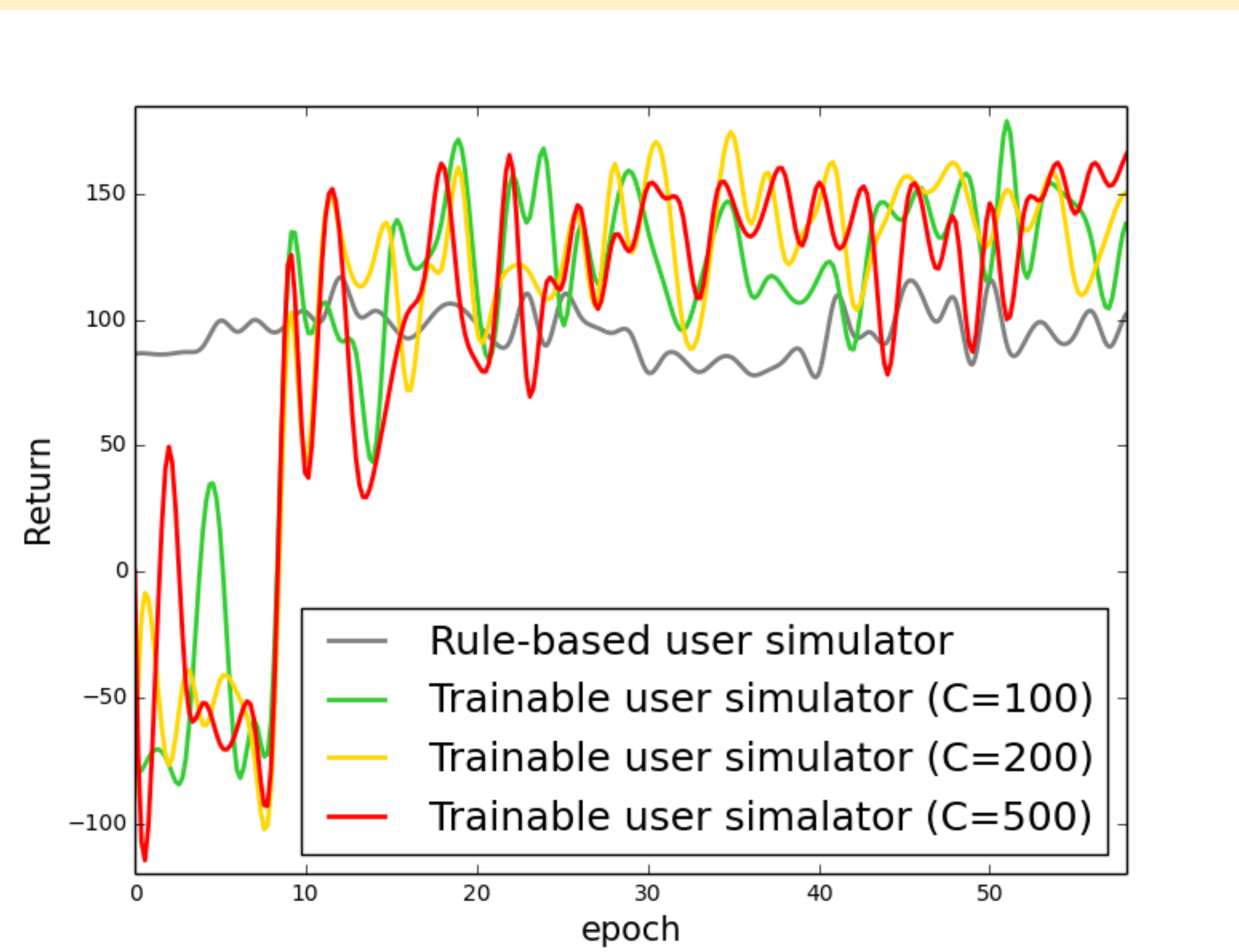
3. Experiments

Dataset:

- A Mandarin Chinese broadcast news corpus
- 5047 Spoken Content Documents for 198 hours

Training:

- Update 2 agents iteratively. Fix one for C steps, update another.



Results:

- Evaluation metrics: MAP (mean average precision) & Return (total reward)

Approaches			One-best		Lattices	
User simulator	Dialogue manager	Input feature	MAP	Return	MAP	Return
Rule-based	DQN	Raw	0.5641	99.31	0.5847	112.24
		Human +Raw	0.5655	107.02	0.5790	101.99
DQN	DQN	Raw	0.5758	113.02	0.6041	162.85
		Human +Raw	0.5820	130.83	0.6027	162.93
Double DQN	Double DQN	Raw	0.6063	190.39	0.6414	224.42
		Human +Raw	0.6066	199.17	0.6375	222.66
Dueling DQN	Dueling DQN	Raw	0.5780	108.72	0.5864	157.92
		Human +Raw	0.5757	127.49	0.5723	158.23
Double DQN	Dueling DQN	Raw	0.5733	134.20	0.5678	143.83
		Human +Raw	0.5573	111.84	0.5905	146.83
Dueling DQN	Double DQN	Raw	0.5899	164.00	0.6494	238.91
		Human +Raw	0.6139	208.25	0.6420	225.72

- Action distribution between Rule-based user/Trainable user/Human

KL-divergence	Rule-based v.s. Human	Trainable v.s. Human	Rule-based v.s. Trainable
	5.9899	2.1503	6.1859