Brief Introduction to Conversation Systems

Yiping Song

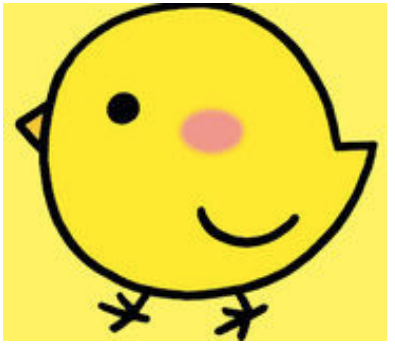
18.4.2

ILLINOIS

Overview

- Problem definition
- Retrieval-based methods
- Generation-based methods
- Combination

Conversation Systems



amazon

Microsoft

人人网 renren.com

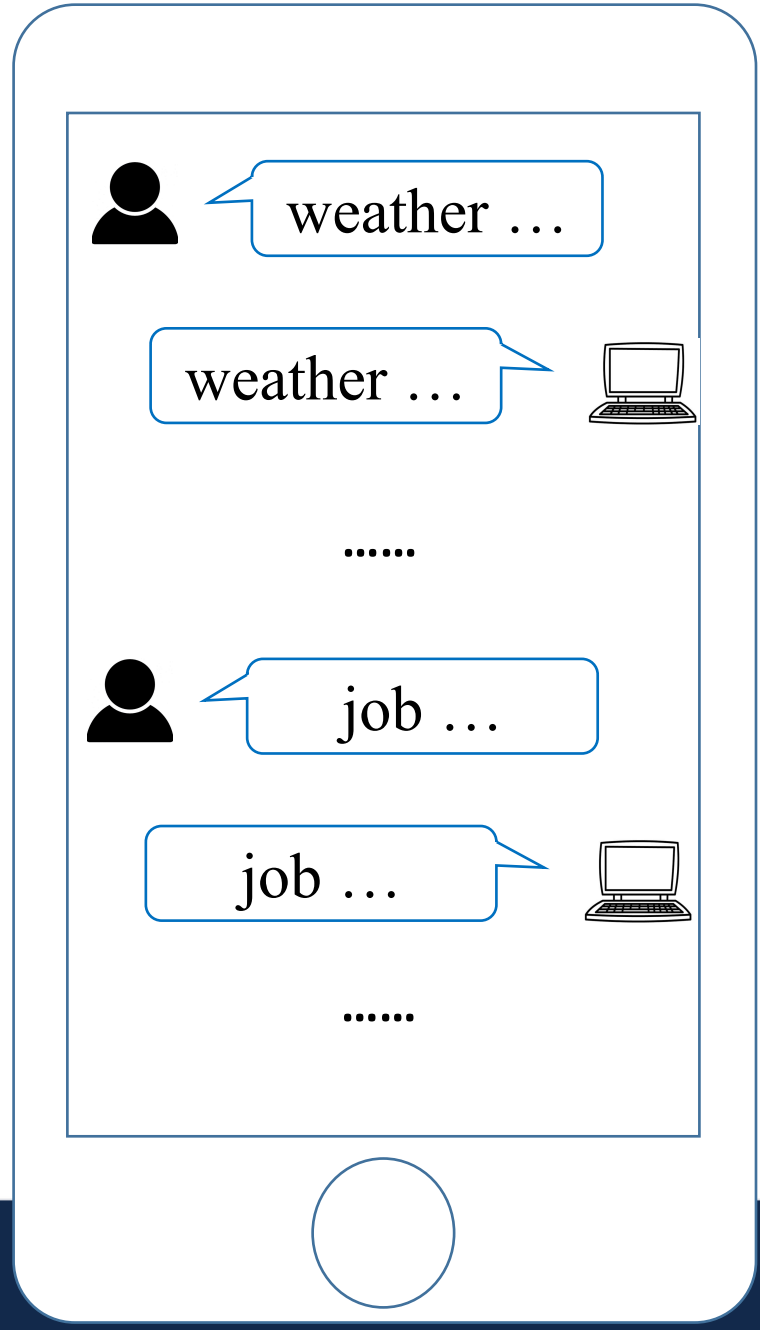
Baidu 百度

Conversation Systems

- open-domain VS specific-domain

- open-domain
  - free talk about anything

- specific-domain
  - book a flight
  - complete a task

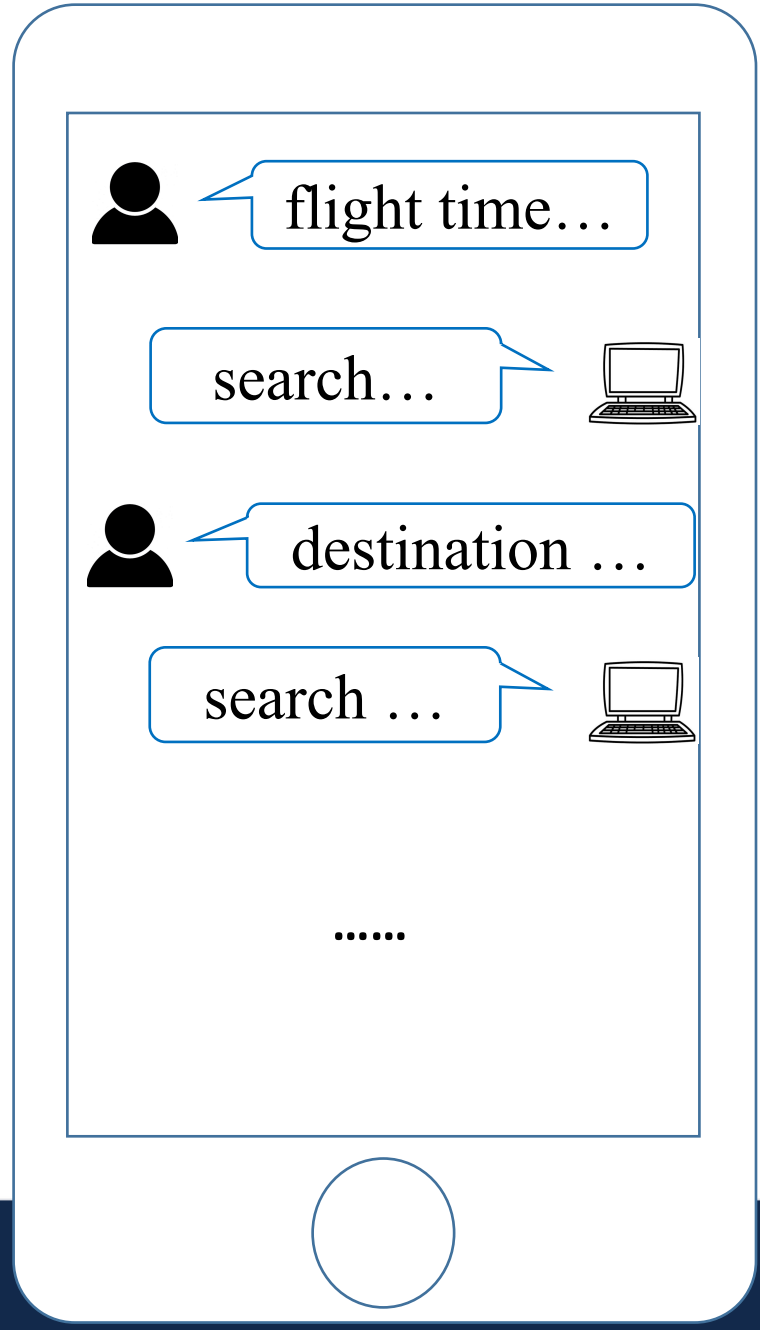


Conversation Systems

- open-domain VS specific-domain

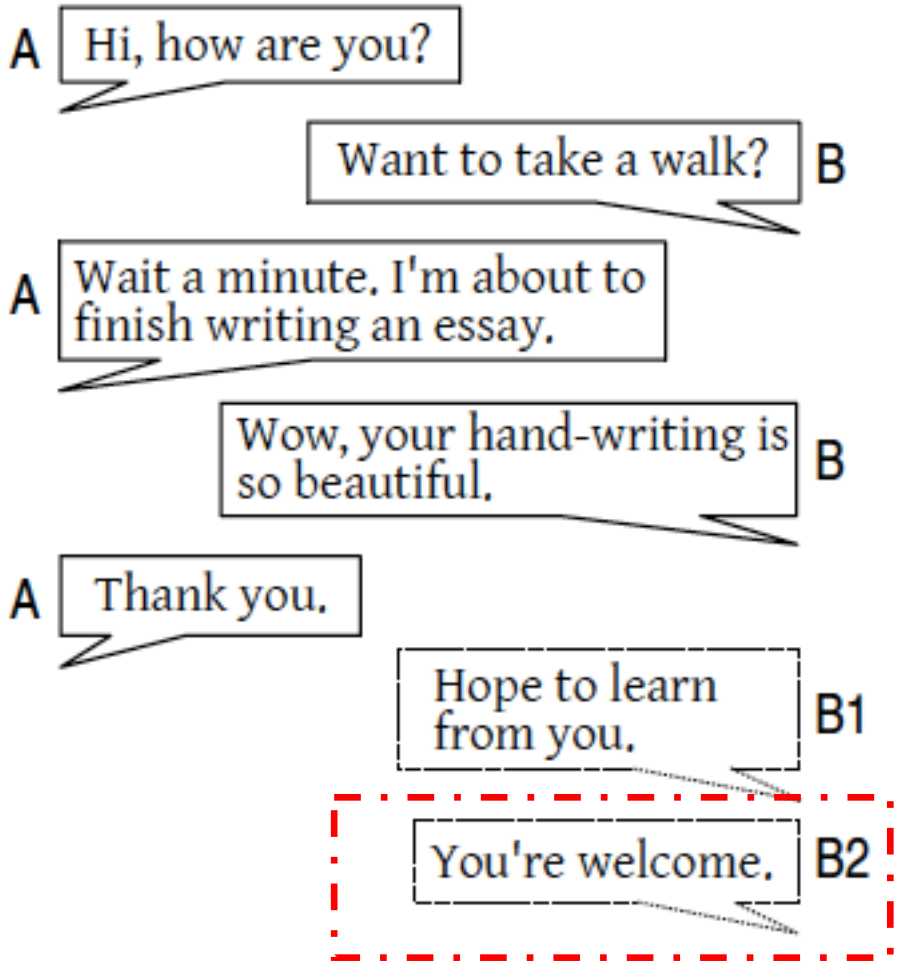
- open-domain
  - free talk about anything

- specific-domain
  - book a flight
  - complete a task



Conversation Systems

- single turn VS multi turn
- single turn
- multi turn



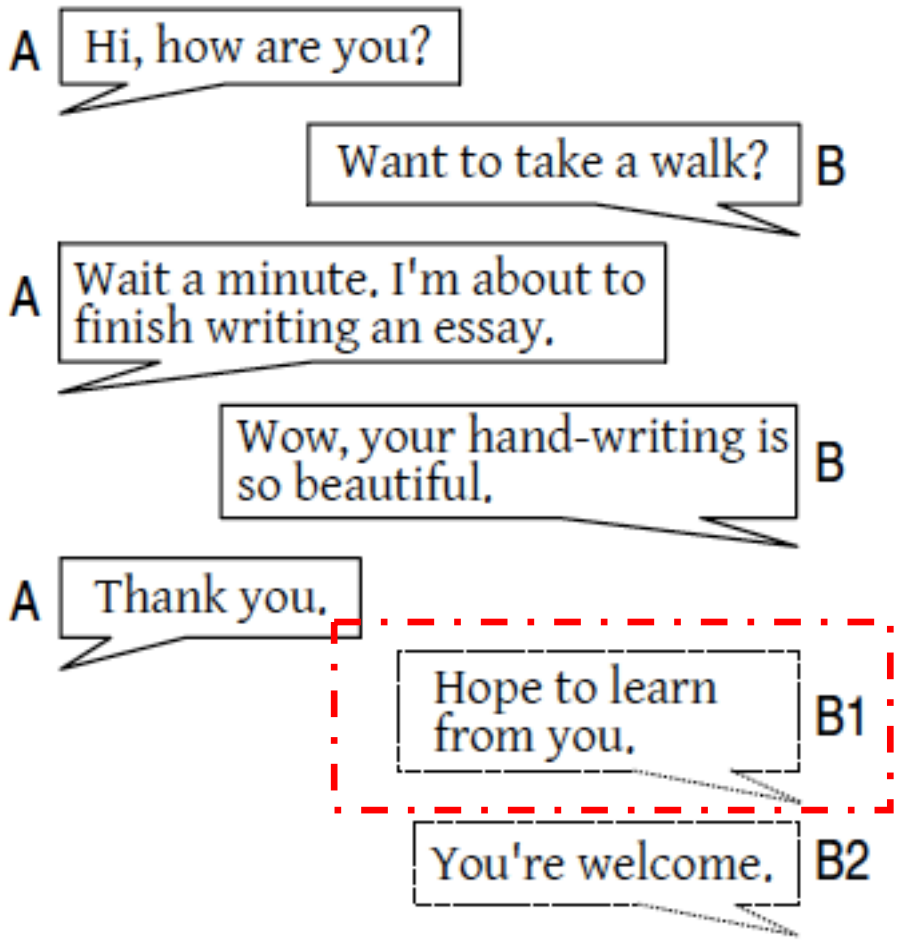
Song et al, INTERSPEECH, 2016

Conversation Systems

- single turn VS multi turn

- single turn

- multi turn



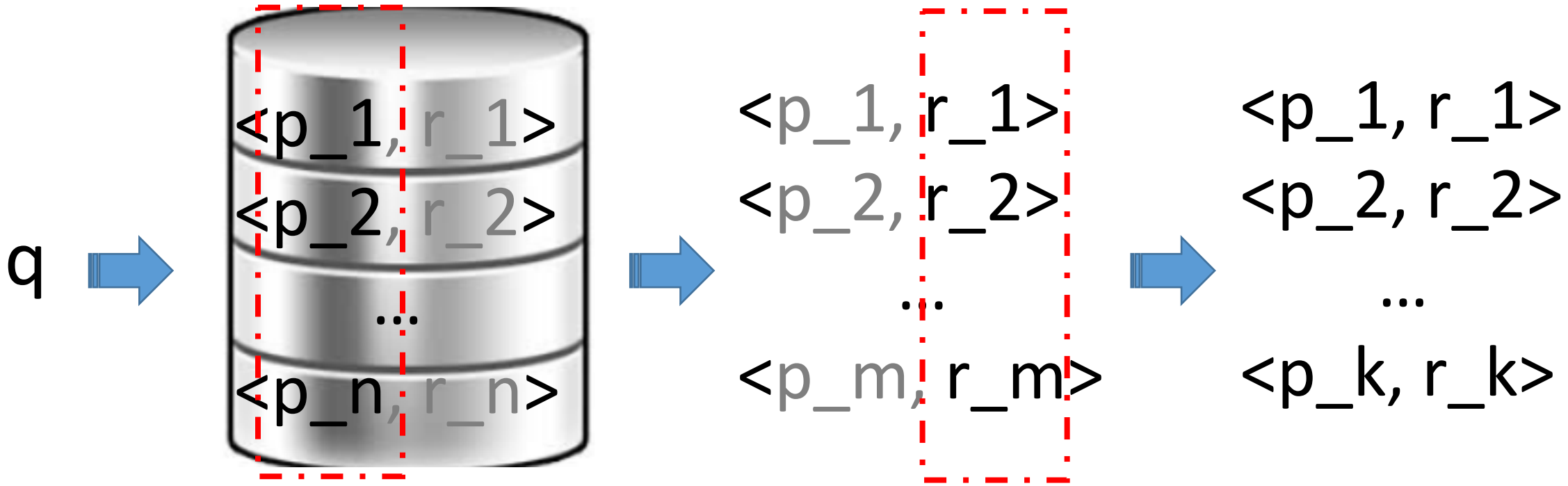
Conversation Systems

- personality

18-year old girl

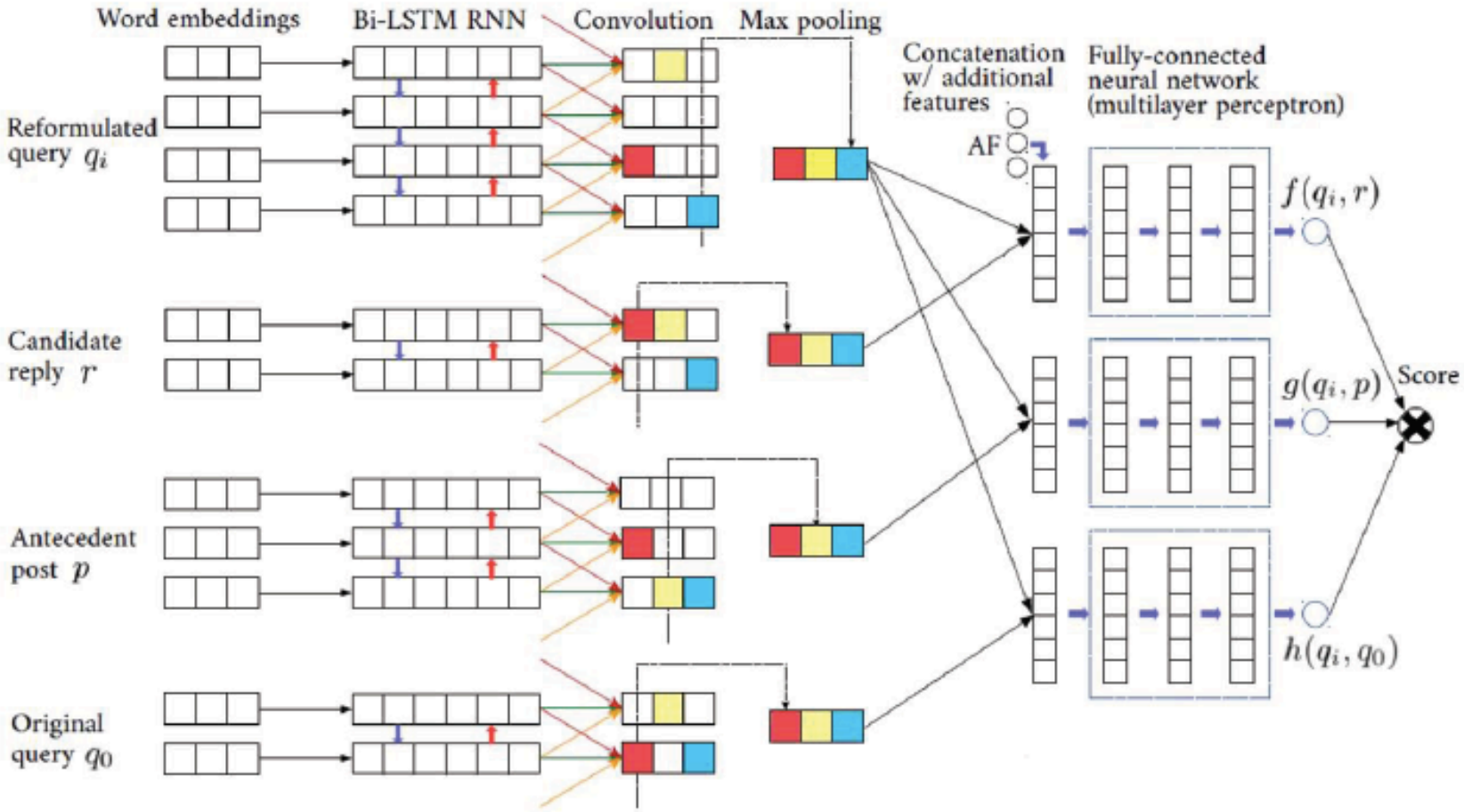
Dongbei(dialect) style

Retrieval-based methods



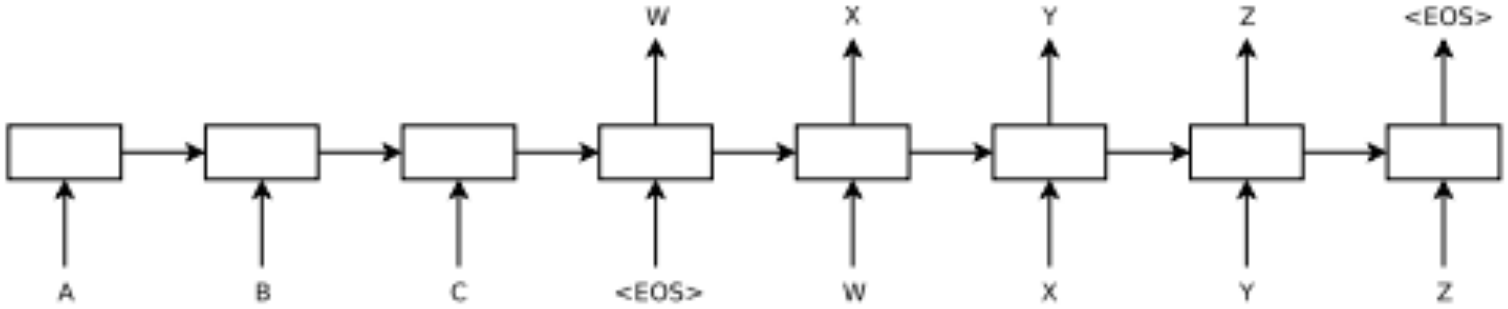
n (the bigger, the better) >> m (1000) > k (5)

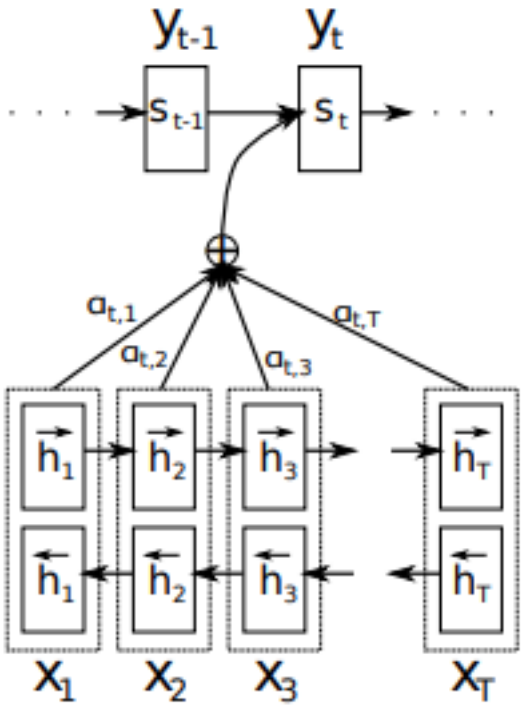
Retrieval-based methods



Yan et al, SIGIR, 2016

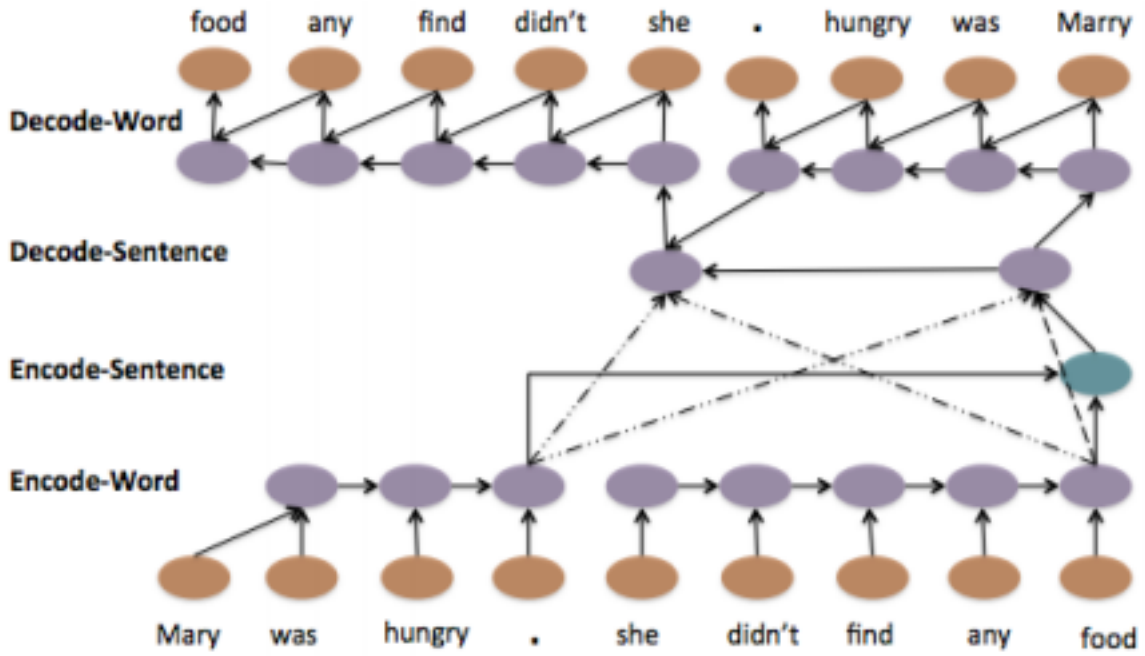
Generation-based methods





Generation-based methods

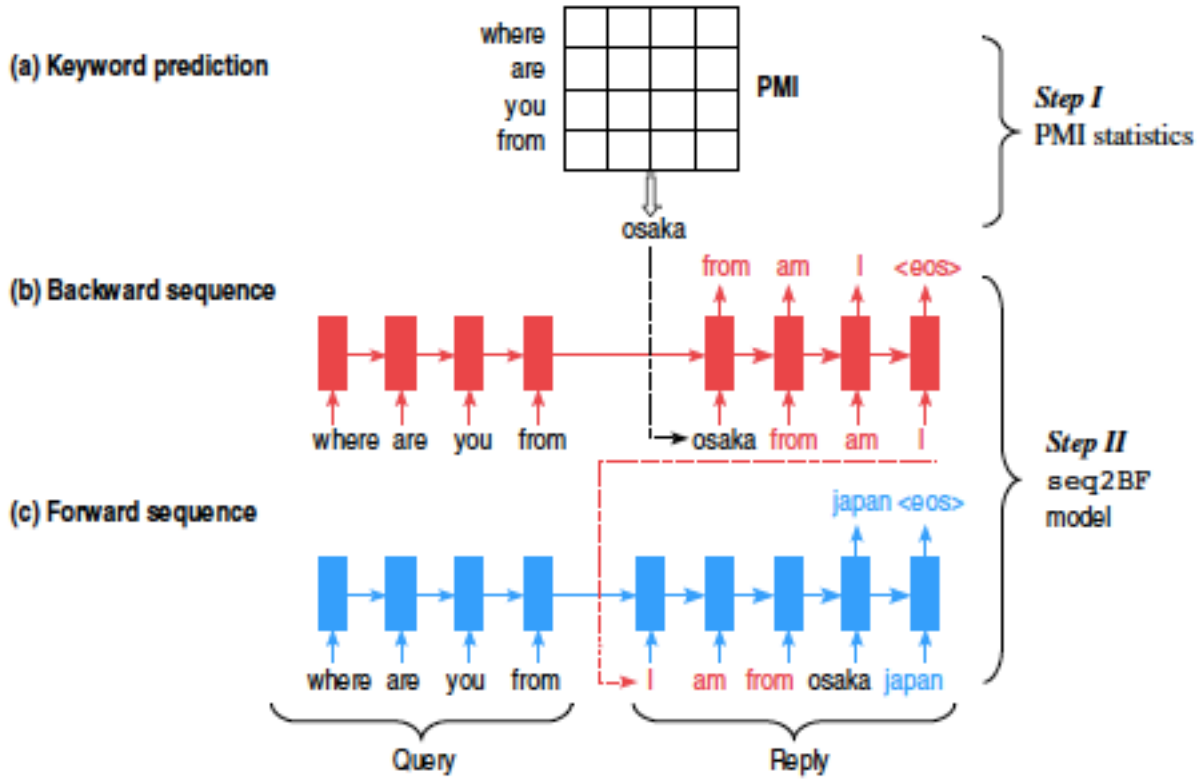
- context-aware



li et al, arxiv, 2015

Generation-based methods

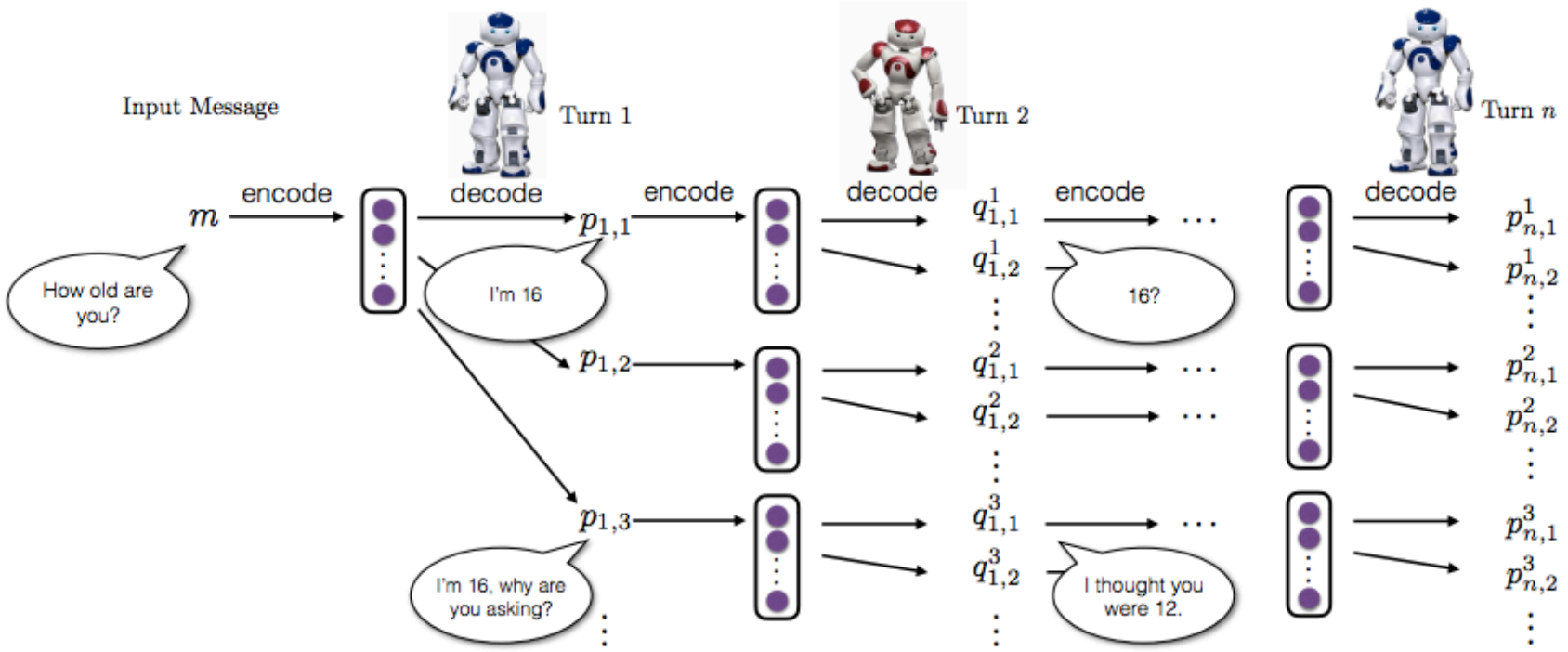
- controlled generation



Mou et al, COLING, 2016

Generation-based methods

- two agents



Li et al, arxiv, 2016

Generation-based methods

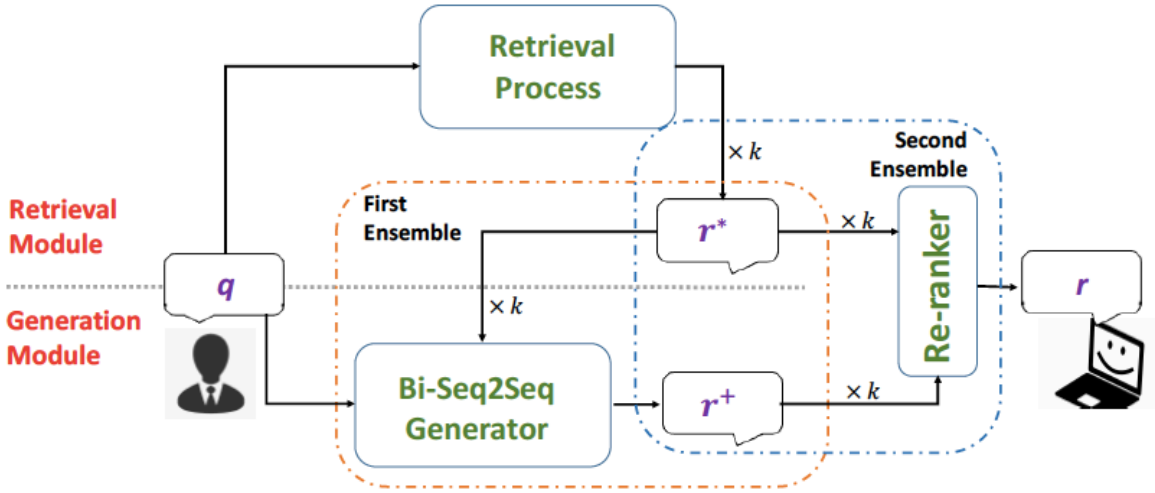
- adversarial training

For number of training iterations **do**
. **For** i=1,D-steps **do**
. Sample (X,Y) from real data
. Sample $\hat{Y} \sim G(\cdot|X)$
. Update D using (X, Y) as positive examples and $(X, \hat{Y})$ as negative examples.
. **End**
.
. **For** i=1,G-steps **do**
. Sample (X,Y) from real data
. Sample $\hat{Y} \sim G(\cdot|X)$
. Compute Reward r for $(X, \hat{Y})$ using D.
. Update G on $(X, \hat{Y})$ using reward r
. Teacher-Forcing: Update G on (X, Y)
. **End**
End

Li et al, arxiv, 2017

Combination

Category	Pros	Cons
Retrieval	literal human utterances; various expressions with great diversity; long in length	not tailored to queries; bottleneck: the size of repository
Generation	tailored for queries; highly coherent	insufficient information; short in length



Song et al, arxiv, 2016

Ultimate goal

Thank you !