A Graph Reasoning Network for Multi-turn Response Selection via Customized

刘永康，冯时，王大玲，宋凯嵩，任飞亮，张一飞

报告人：刘永康

东北大学

目录

CONTENT

- Graph Reasoning
- Adaptive Pretraining
- Experiments

Reason Vs Match?

Context

u_1 : Good morning , two tickets to London , please .

u_2 : Express train or regular one ?

u_3 : Any difference ?

u_4 : The regular ticket is 80 dollars while price …

u_5 : I see , but how long does the express train take ?

u_6 : Comparatively , …the next one is leaving at 10:45

u_7 : Oh , we only have 15 minutes to go .

u_8 : But that 's the time to catch the train . …earlier...

Reason

Match

Response

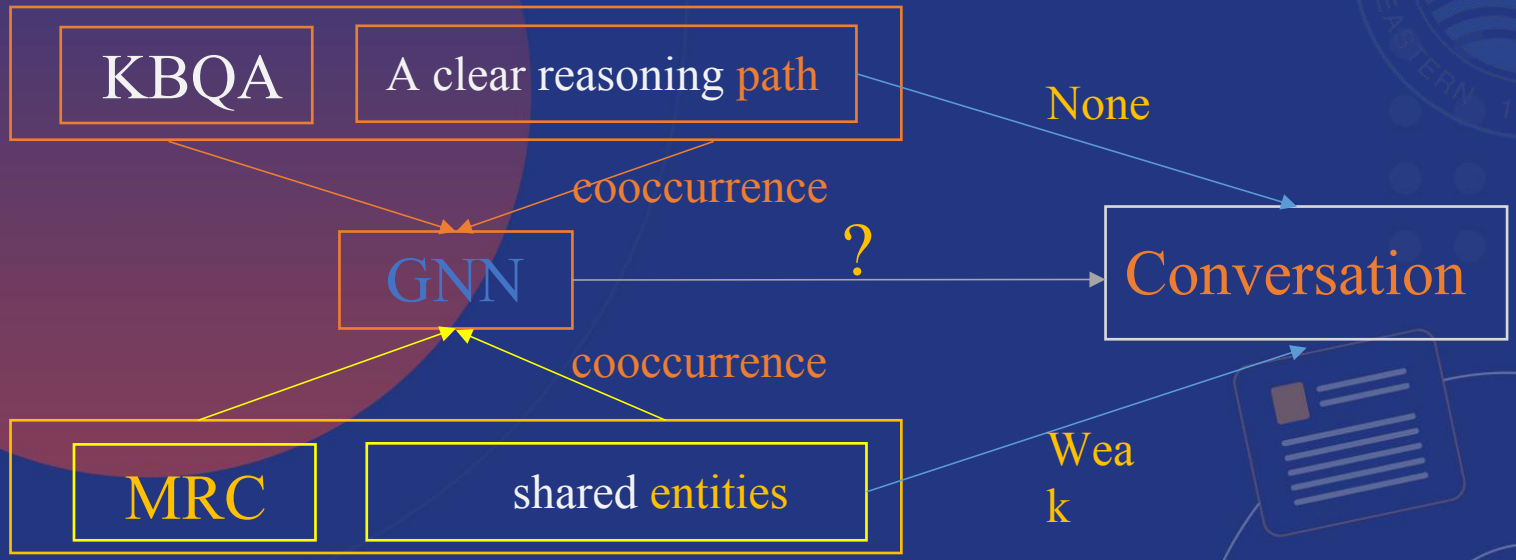
u_9 : …it 's already 10:30. you only have 15 minutes…

u_9 : …it 's already 10:45. you only have 15 minutes…

u_9 : …it 's already 10:15. you only have 30 minutes…

u_9 : ……

Distinction on Conversation



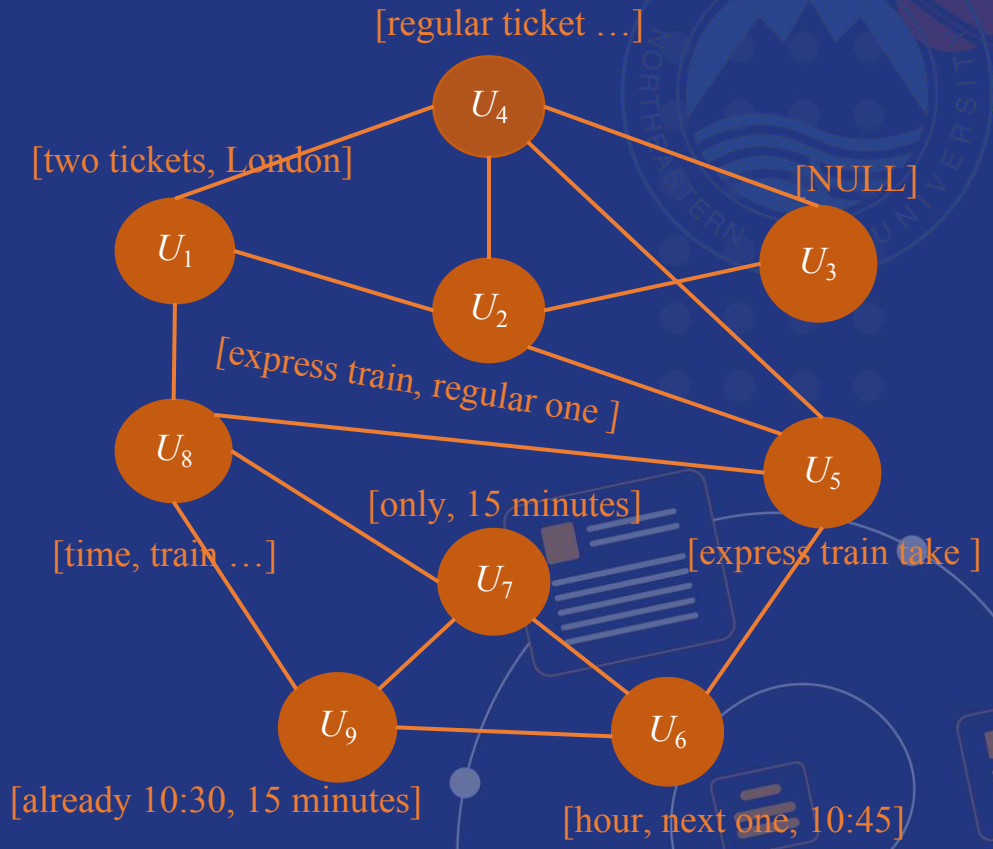
Graph Builder

- Token-level
  - Node : token
  - Sparse
- Utterance-level
  - Node : utterance
  - Simple
  - Sparse degree: moderation

u_1 : Good morning , two tickets to London , please .
u_2 : Express train or regular one ?
u_3 : Any difference ?
u_4 : The regular ticket is 80 dollars while price …
u_5 : I see , but how long does the express train take ?
u_6 : Comparatively , …the next one is leaving at 10:45 .
u_7 : Oh , we only have 15 minutes to go .
u_8 : But that 's the time to catch the train . …earlier...

Graph Builder

- Chronological Rule
  - Add an edge between adjacent utterances
- Topic Dependency
  - Topic: abstract
    - Each utterance has a central idea
    - Keyword set
  - Clustering/community detection



Better Representation

Different Response

u_1: Good morning , two tickets to London , please .
u_2 : Express train or regular one ?
u_3 : Any difference ?
u_4 : The regular ticket is 80 dollars while price …
u_5 : I see , but how long does the express train take ?

u_1: Good morning , two tickets to London , please .
u_2 : Express train or regular one ?
u_3 : Any difference ?
u_4 : The regular ticket is 80 dollars while price …

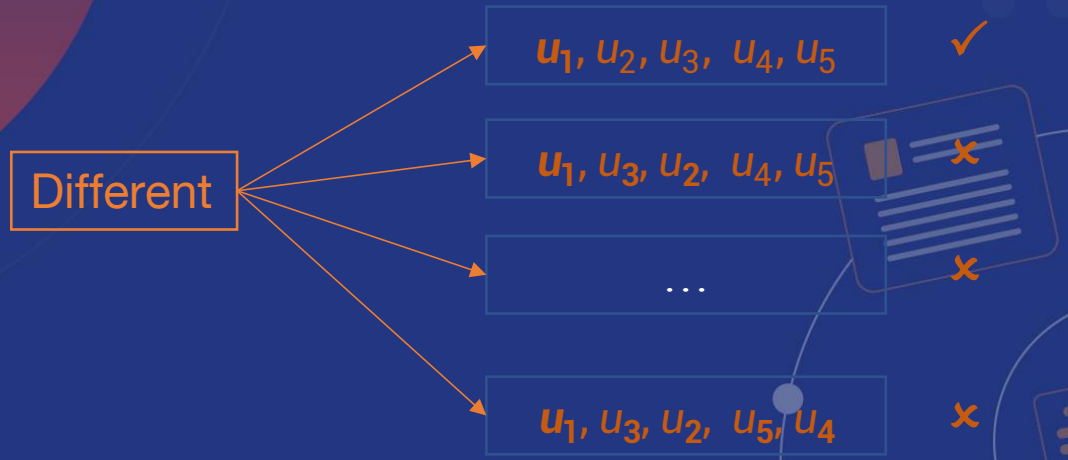
- How Model ?
  - NUP

Suppose utterance $u = \{u_1, u_2, \ldots, u_n\}$
$u_l = \{u_1, u_2, \ldots, u_{i-1}\}$; $u_r = \{u_{i+1}, u_{i+2}, \ldots, u_n\}$
$<u_l, u_i>$ $<u_r, u_i>$

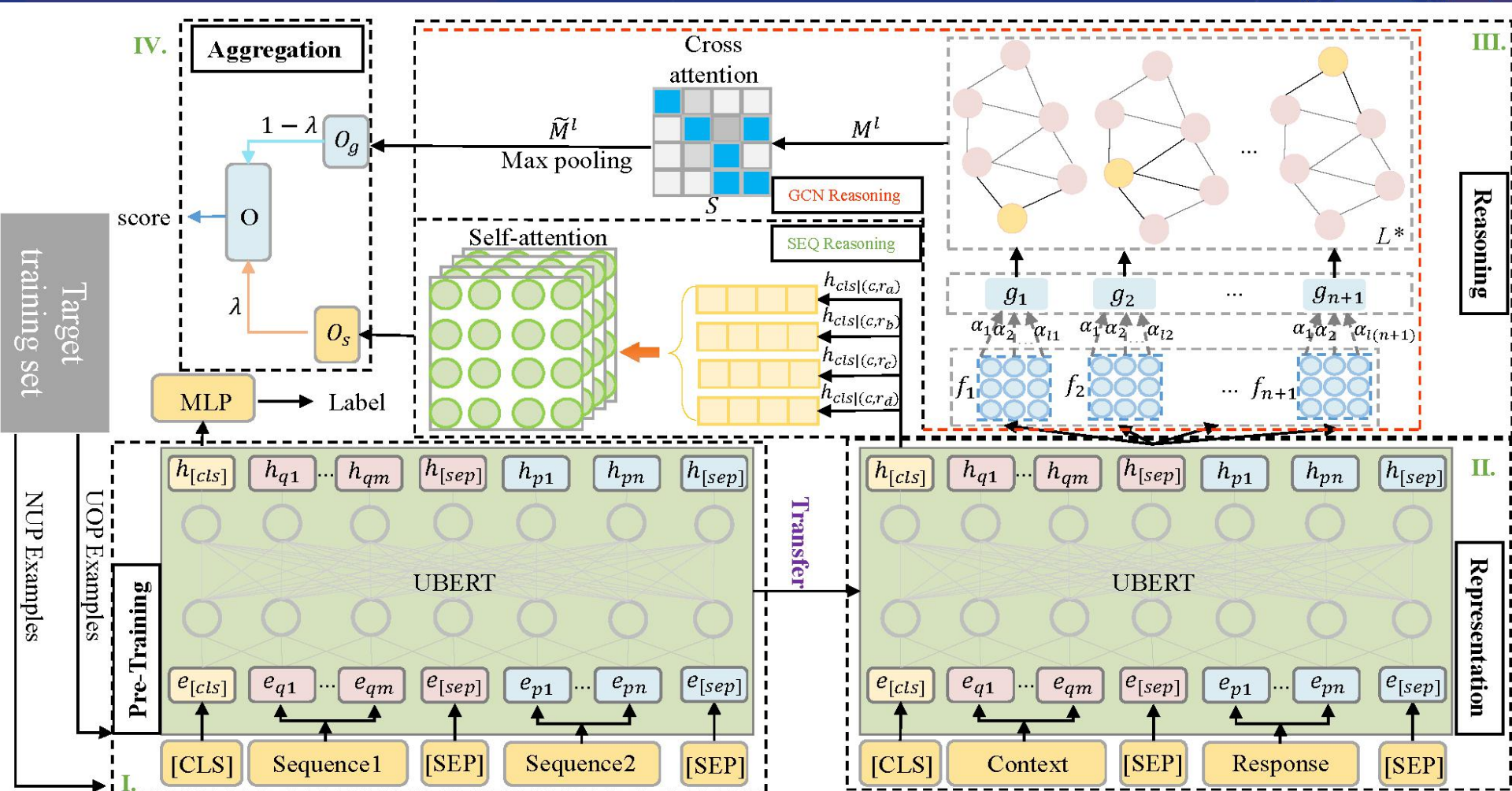
Better Representation

u_1 : Good morning , two tickets to London , please .
u_2 : Express train or regular one ?
u_3 : Any difference ?
u_4 : The regular ticket is 80 dollars while price …
u_5 : I see , but how long does the express train take ?

- How Model ?
  - NOP



Model Architecture



Experiments

- GRN VS Baselines(https://nealcly.github.io/MuTual-leaderboard/)

	MuTual			MuTual plus		
Method	R@1	R@2	MRR	R@1	R@2	MRR
Human	93.8	97.1	96.4	93.0	97.2	96.1
Random	25.0	50	60.4	25.0	50	60.4
TF-IDF	27.9	53.6	54.2	27.8	52.9	76.4
DuLSTM	26.0	49.1	74.3	25.1	47.9	51.5
SMN	29.9	58.5	59.5	26.5	51.6	62.7
DAM	24.1	46.5	51.8	27.2	52.3	69.5
BIDAF	35.7	58.9	58.9	33.4	49.2	56.2
R-NRT	27.0	43.5	51.3	26.1	50.6	53.2
QANET	24.7	51.7	52.2	25.1	49.5	51.9
BERT	64.8	84.7	79.5	51.4	78.7	71.5
RoBERTa	82.5	95.3	90.4	75.7	92.8	85.6
SpanBERT	80.6	94.8	89.3	70.3	88.4	83.0
GPT-2	33.2	60.2	58.4	31.6	57.4	56.8
GPT-2-FT	39.2	67.0	62.9	22.6	61.1	53.5
BERTMC	66.7	87.8	81.0	58.0	79.2	74.9
RoBERTMC	68.6	88.7	82.8	64.3	84.5	79.2
ALBERT	84.7	96.2	91.6	78.9	94.6	88.4
GRN	91.5	98.3	95.4	84.1	95.7	91.3

Experiments

- Ablation Study

Method	R@1	R@2	MRR
GRN	93.5	98.5	97.1
-pre-training	90.5	97.3	94.7
-GCN match	91.5	97.9	95.5
-sequence match	91.3	97.6	95.2
-cross attention	92.2	97.3	95.6
-selfAtt	92.7	98.2	96.8

- Pre-training Methods

Pretraining Method	R@1	R@2	MRR
ALBERT	84.8	96.0	91.6
BERT	78.2	90.6	84.9
Our Method	87.6	96.7	93.6

THANKS

报告人：刘永康

东北大学

2020.12.19