# HETERMPC: A Heterogeneous Graph Neural Network for Response Generation in Multi-Party Conversations

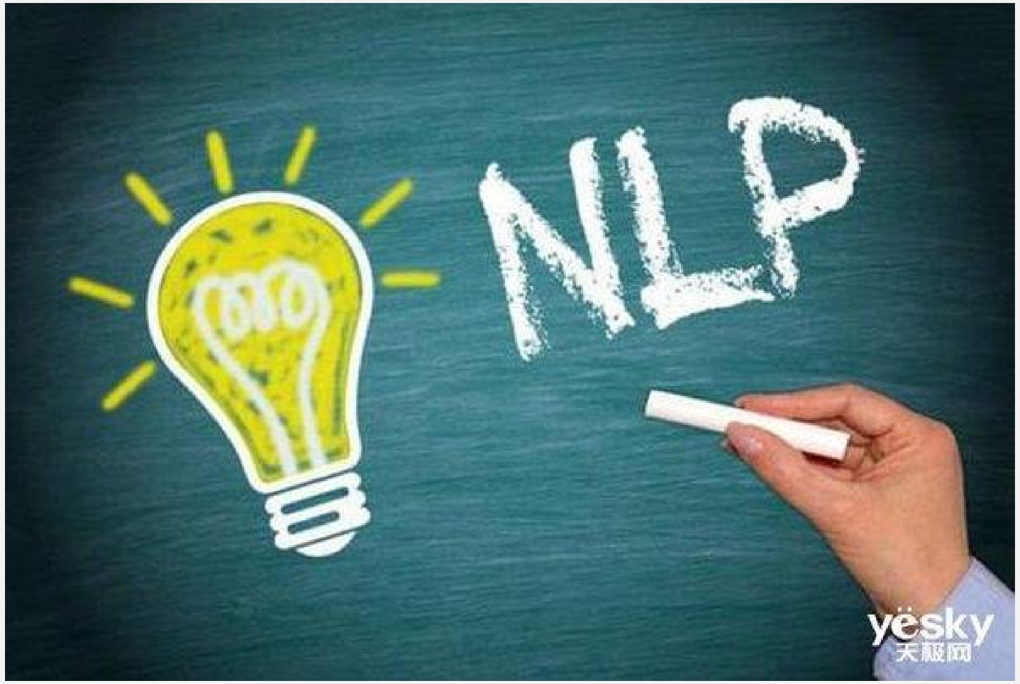
**Jia-Chen Gu[1*†], Chao-Hong Tan[1†], Chongyang Tao[2], Zhen-Hua Ling[1], Huang Hu[2], Xiubo Geng[2], Daxin Jiang[2‡]**

[1]National Engineering Research Center for Speech and Language Information Processing, University of Science and Technology of China, Hefei, China

[2]Microsoft, Beijing, China

{gujc,chtan}@mail.ustc.edu.cn, zhling@ustc.edu.cn,
{chotao,huahu,xigeng,djiang}@microsoft.com

**(ACL-2022)** **Reported by Jia Wang**




1. Introduction
2. Approach
3. Experiments

# Introduction

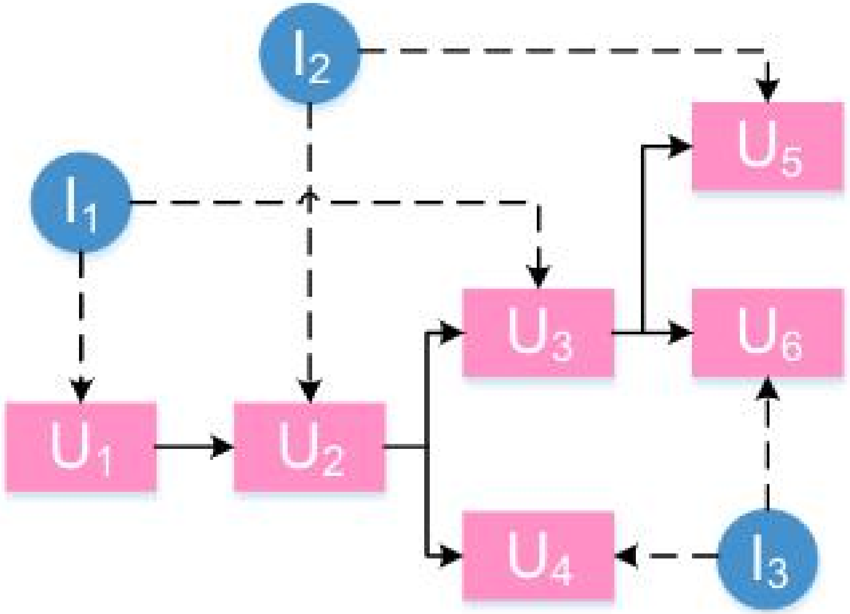



Figure 1: Illustration of a graphical information flow in an MPC. Pink rectangles denote utterances and blue circles denote interlocutors. Each solid line represents the "*replied-by*" relationship between two utterances. Each dashed line indicates the speaker of an utterance.

# Introduction

In summary, our contributions in this paper are three-fold:

• To the best of our knowledge, this paper is the first exploration of using heterogeneous graphs for modeling conversations;

• A Transformer-based heterogeneous graph architecture is introduced for response generation in MPCs, in which two types of nodes, six types of meta relations, and node-edge-type-dependent parametersare employed to characterize the heterogeneous properties of MPCs;

• Experimental results show that our proposed model achieves a new state-of-the-art performance of response generation in MPCs on the Ubuntu IRC benchmark.

# Approach

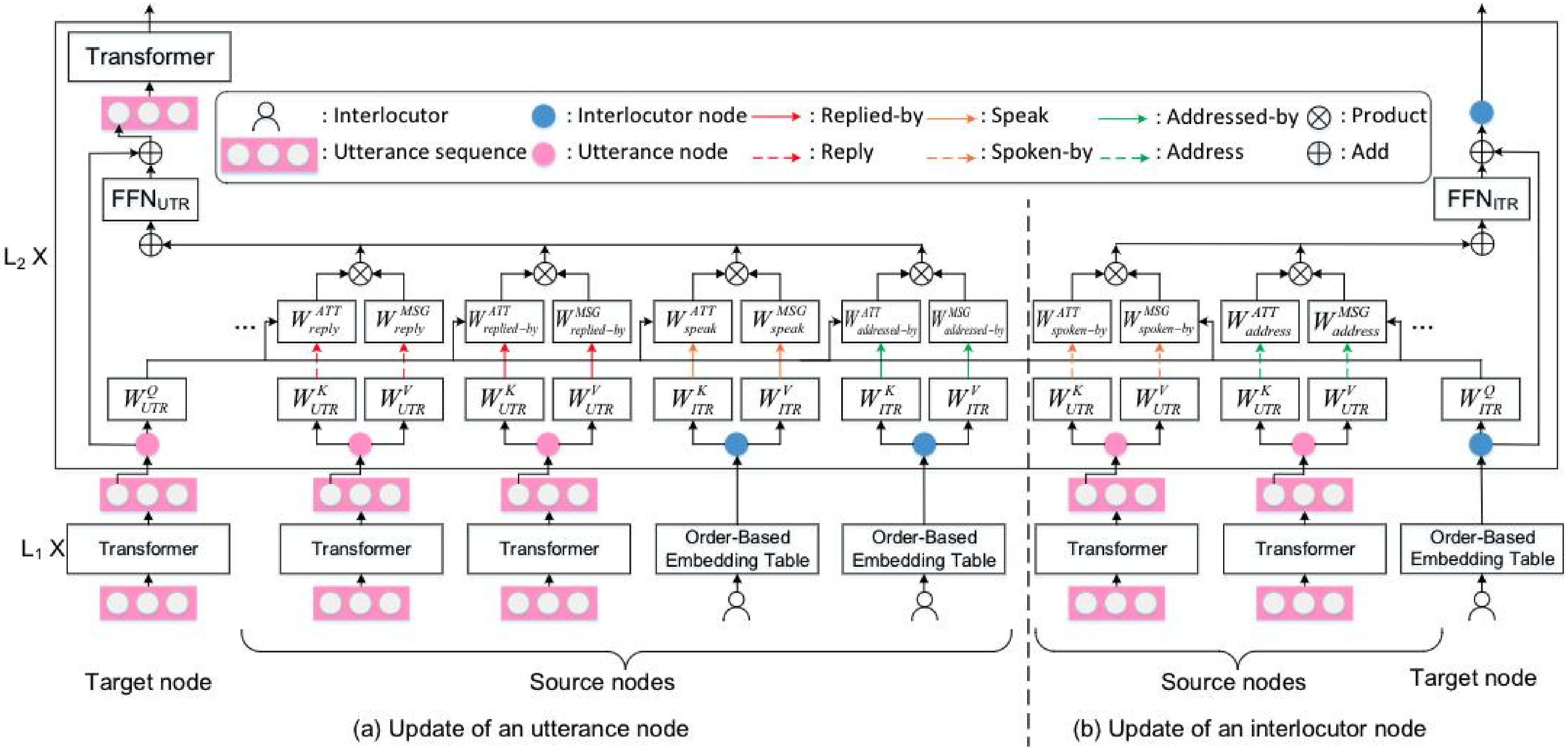



Figure 3: Model architecture of HeterMPC for (a) update of an utterance node and (b) update of an interlocutor node. "UTR" and "ITR" are abbreviations of "utterance" and "interlocutor" respectively. The set of $W_*^*$ denotes the node-edge-type-dependent parameters.

# Approach

## Graph Construction

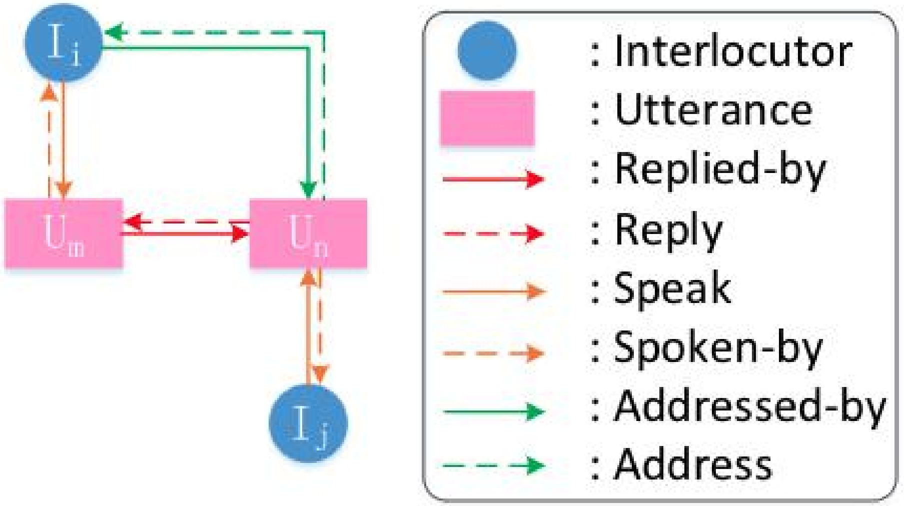
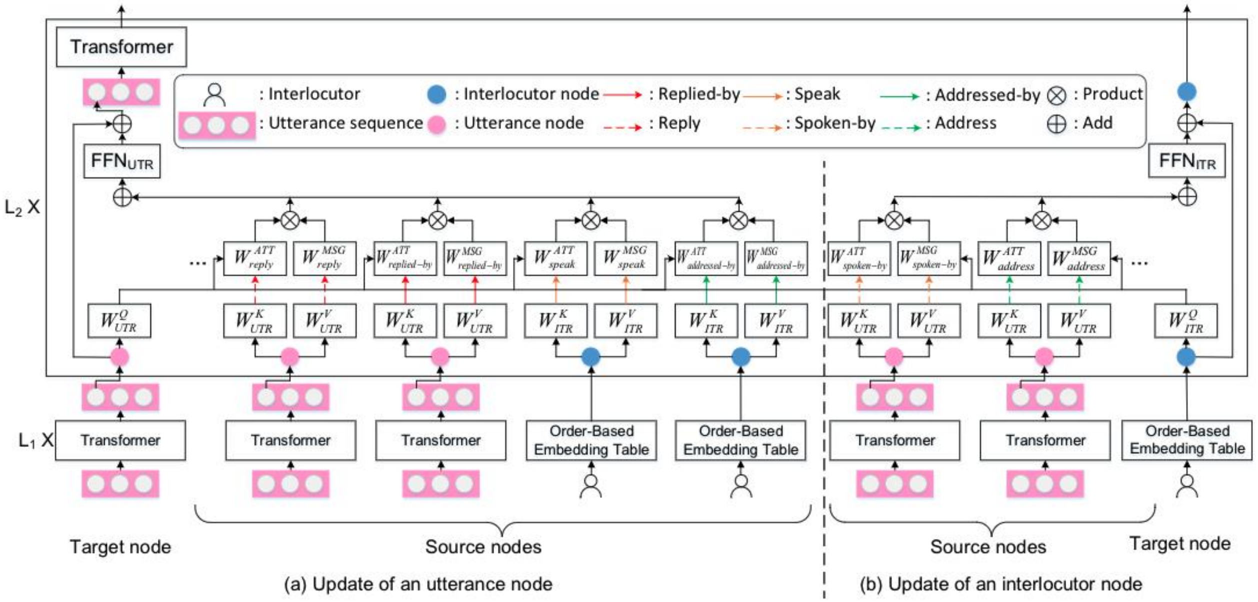


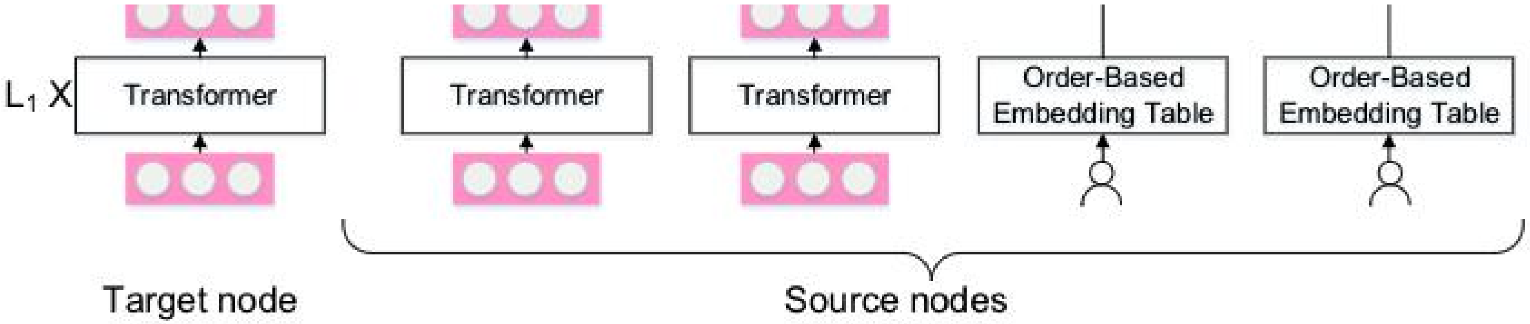

Figure 2: Illustration of the two types of nodes and six types of edges in a heterogeneous conversation graph. This example demonstrates that $\mathrm{I}_j$ speaks $\mathrm{U}_n$ replying $\mathrm{U}_m$ that is spoken-by $\mathrm{I}_i$.

Given an MPC instance composed of $M$ utterances and $I$ interlocutors, a heterogeneous graph $\mathbb{G}(\mathbb{V}, \mathbb{E})$ is constructed. Specifically, $\mathbb{V}$ is a set of $M + I$ nodes. Each node denotes either an utterance or an interlocutor. $\mathbb{E} = \{e_{p,q}\}_{p,q=1}^{M+I}$ is a set of directed edges. Each edge $e_{p,q}$ describes the connection from node $p$ to node $q$.

# Approach

## Node Initialization



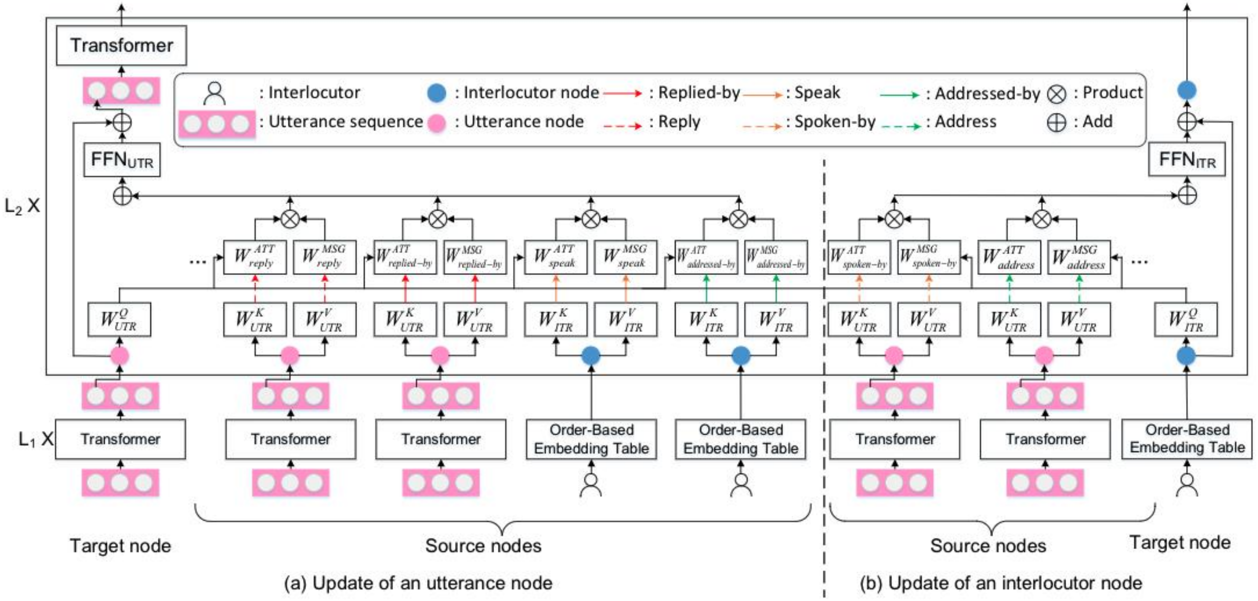

## Utterances:

$$\boldsymbol{H}_m^{l+1} = \text{TransformerEncoder}(\boldsymbol{H}_m^l), \qquad (2)$$

where $m \in \{1, ..., M\}$, $l \in \{0, ..., L_1 - 1\}$, $L_1$ denotes the number of Transformer layers for initialization, $\boldsymbol{H}_m^l \in \mathbb{R}^{k_m \times d}$, $k_m$ denotes the length of an utterance and $d$ denotes the dimension of embedding vectors.

# Approach

## Node Updating: Heterogeneous Attention




$$\boldsymbol{k}^l(s) = \boldsymbol{h}_s^l \boldsymbol{W}_{\tau(s)}^K + \boldsymbol{b}_{\tau(s)}^K, \tag{3}$$

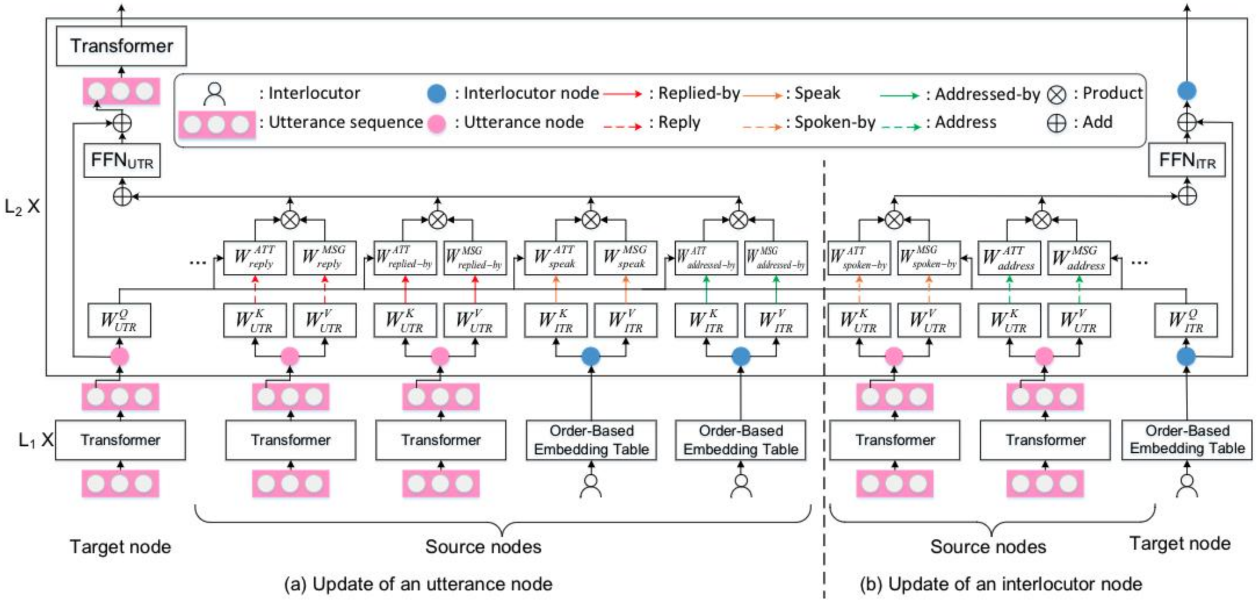
$$\boldsymbol{q}^l(t) = \boldsymbol{h}_t^l \boldsymbol{W}_{\tau(t)}^Q + \boldsymbol{b}_{\tau(t)}^Q, \tag{4}$$

$$w^l(s,e,t) = \boldsymbol{k}^l(s) \boldsymbol{W}_{e_{s,t}}^{ATT} \boldsymbol{q}^l(t)^T \frac{\mu_{e_{s,t}}}{\sqrt{d}}. \tag{5}$$

Here, $\tau(s), \tau(t) \in$ {*UTR*, *ITR*} distinguish utterance (*UTR*) and interlocutor (*ITR*) nodes. Eqs. (3) and (4) are node-type-dependent linear transformations. Eq. (5) contains an edge-type-dependent linear projection $\boldsymbol{W}_{e_{s,t}}^{ATT}$ where $\mu_{e_{s,t}}$ is an adaptive factor scaling to the attention. All $\boldsymbol{W}^* \in \mathbb{R}^{d \times d}$ and $\boldsymbol{b}^* \in \mathbb{R}^d$ are parameters to be learnt.

# Approach

## Node Updating: Heterogeneous Message Passing



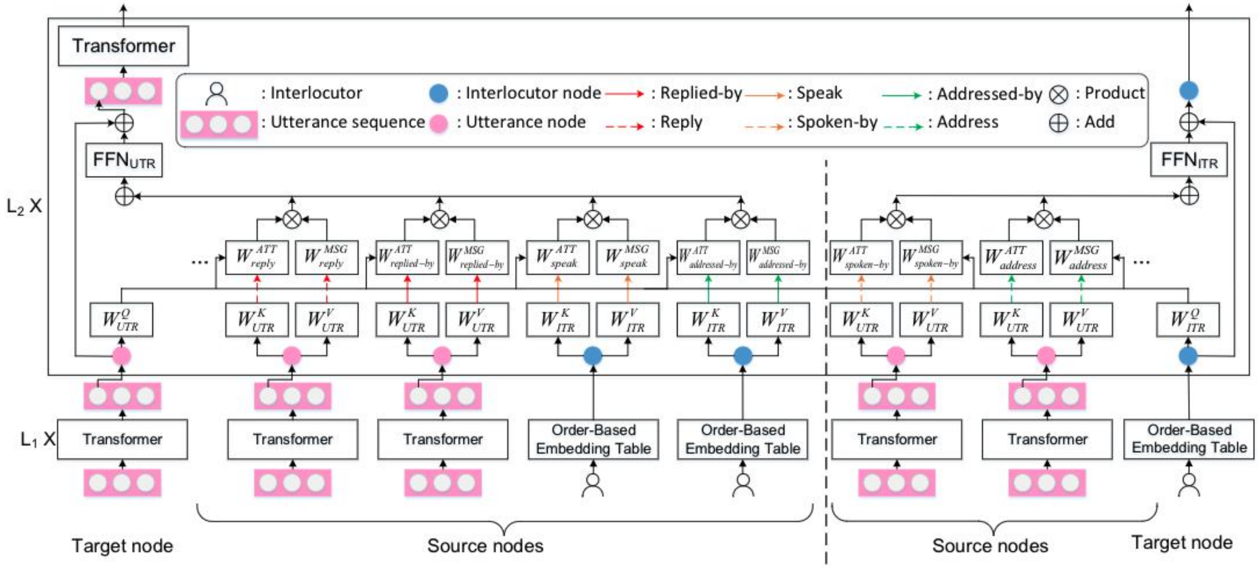

$$\boldsymbol{v}^l(s) = \boldsymbol{h}_s^l \boldsymbol{W}^V_{\tau(s)} + \boldsymbol{b}^V_{\tau(s)}, \qquad (6)$$

$$\bar{\boldsymbol{v}}^l(s) = \boldsymbol{v}^l(s) \boldsymbol{W}^{MSG}_{e_{s,t}}, \qquad (7)$$

where $\bar{\boldsymbol{v}}^l(s)$ is the passed message and all $\boldsymbol{W}^* \in \mathbb{R}^{d \times d}$ and $\boldsymbol{b}^* \in \mathbb{R}^d$ are parameters to be learnt.

# Approach

## Node Updating: Heterogeneous Aggregation




(a) Update of an utterance node (b) Update of an interlocutor node

$$\bar{\boldsymbol{h}}_t^l = \sum_{s \in S(t)} \text{softmax}(w^l(s,e,t))\bar{\boldsymbol{v}}^l(s), \quad (8)$$

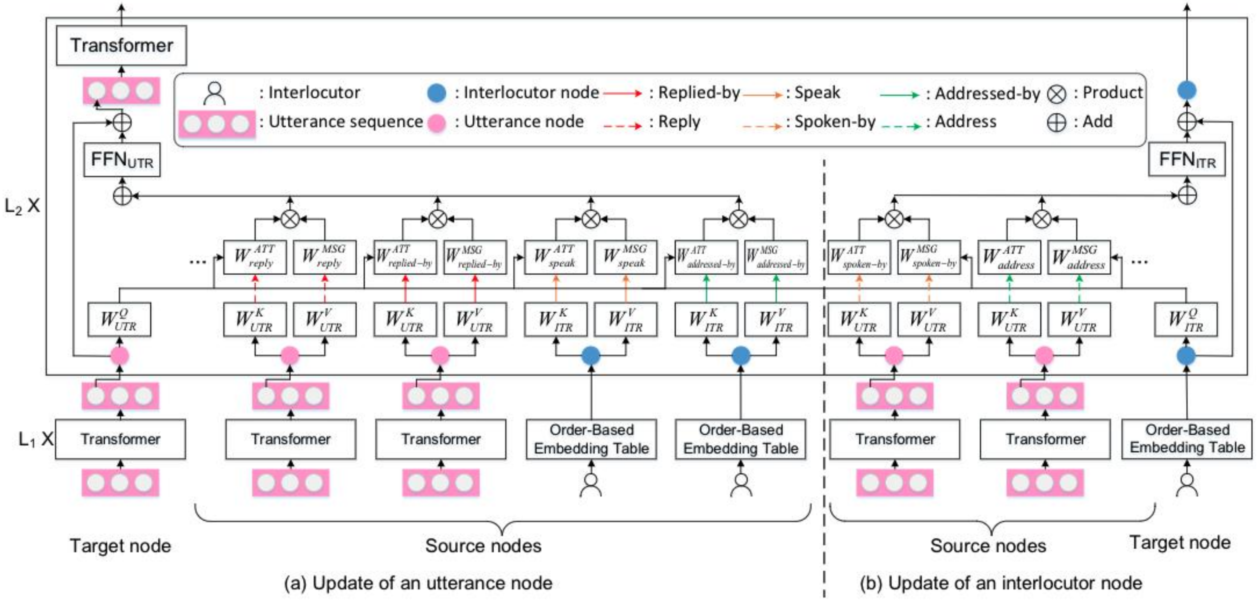
$$\boldsymbol{h}_t^{l+1} = FFN_{\tau(t)}(\bar{\boldsymbol{h}}_t^l) + \boldsymbol{h}_t^l, \quad (9)$$

$$\hat{\boldsymbol{h}}_t^{l+1} = [\boldsymbol{h}_t^l; \boldsymbol{h}_t^{l+1}]\boldsymbol{W}_{com} + \boldsymbol{b}_{com}, \quad (10)$$

where $\boldsymbol{W}_{com} \in \mathbb{R}^{2d \times d}$ and $\boldsymbol{b}_{com} \in \mathbb{R}^d$ are parameters. Then, $\hat{\boldsymbol{h}}_t^{l+1}$ replaces the representation of `[CLS]` (i.e., $\boldsymbol{h}_t^l$) in the sequence representations of the whole utterance.

# Approach

## Decoder



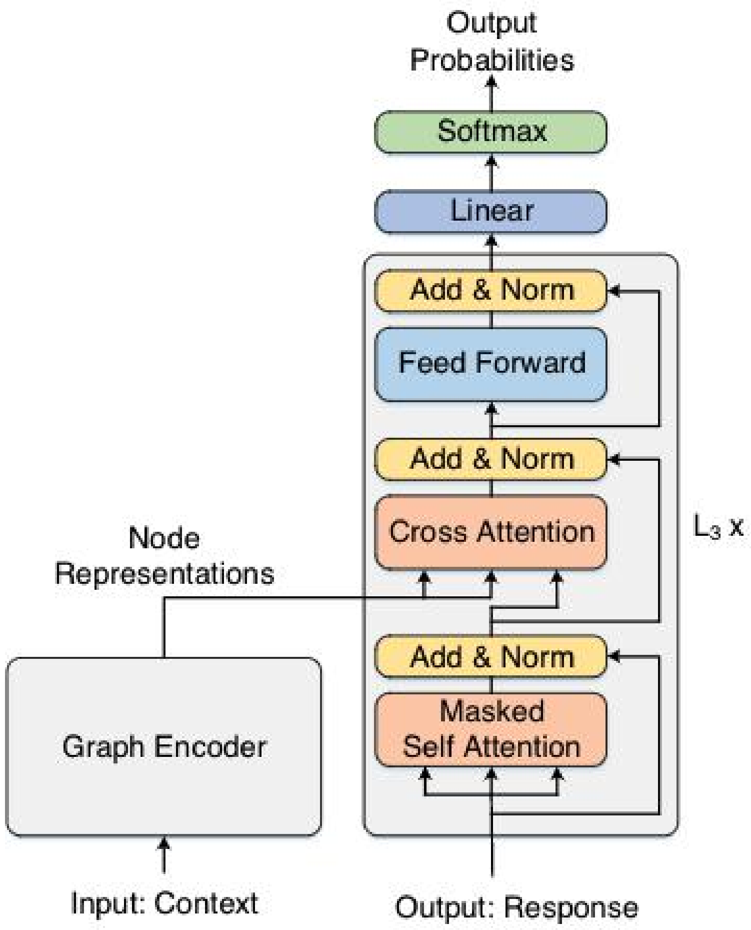




Figure 4: The decoder architecture of HeterMPC.

# Experiments

| Models \ Metrics | BLEU-1 | BLEU-2 | BLEU-3 | BLEU-4 | METEOR | ROUGE$_L$ |
|---|---|---|---|---|---|---|
| Seq2Seq (LSTM) (Sutskever et al., 2014) | 7.71 | 2.46 | 1.12 | 0.64 | 3.33 | 8.68 |
| Transformer (Vaswani et al., 2017) | 7.89 | 2.75 | 1.34 | 0.74 | 3.81 | 9.20 |
| GSN (Hu et al., 2019b) | 10.23 | 3.57 | 1.70 | 0.97 | 4.10 | 9.91 |
| GPT-2 (Radford et al., 2019) | 10.37 | 3.60 | 1.66 | 0.93 | 4.01 | 9.53 |
| BERT (Devlin et al., 2019) | 10.90 | 3.85 | 1.69 | 0.89 | 4.18 | 9.80 |
| HeterMPC$_{BERT}$ | **12.61** | **4.55** | **2.25** | **1.41** | **4.79** | **11.20** |
| HeterMPC$_{BERT}$ w/o. node types | 11.76 | 4.09 | 1.87 | 1.12 | 4.50 | 10.73 |
| HeterMPC$_{BERT}$ w/o. edge types | 12.02 | 4.27 | 2.10 | 1.30 | 4.74 | 10.92 |
| HeterMPC$_{BERT}$ w/o. node and edge types | 11.60 | 3.98 | 1.90 | 1.18 | 4.20 | 10.63 |
| HeterMPC$_{BERT}$ w/o. interlocutor nodes | 11.80 | 3.96 | 1.75 | 1.00 | 4.31 | 10.53 |
| BART (Lewis et al., 2020) | 11.25 | 4.02 | 1.78 | 0.95 | 4.46 | 9.90 |
| HeterMPC$_{BART}$ | **12.26** | **4.80** | **2.42** | **1.49** | **4.94** | **11.20** |
| HeterMPC$_{BART}$ w/o. node types | 11.22 | 4.06 | 1.87 | 1.04 | 4.57 | 10.63 |
| HeterMPC$_{BART}$ w/o. edge types | 11.52 | 4.27 | 2.05 | 1.24 | 4.78 | 10.90 |
| HeterMPC$_{BART}$ w/o. node and edge types | 10.90 | 3.90 | 1.79 | 1.01 | 4.52 | 10.79 |
| HeterMPC$_{BART}$ w/o. interlocutor nodes | 11.68 | 4.24 | 1.91 | 1.03 | 4.79 | 10.65 |

Table 1: Performance of HeterMPC and ablations on the test set in terms of automated evaluation. Numbers in bold denote that the improvement over the best performing baseline is statistically significant (t-test with $p$-value $< 0.05$).

# Experiments

| Metrics / Models | Score | Kappa |
|---|---|---|
| Human | 2.81 | 0.55 |
| GSN (Hu et al., 2019b) | 2.00 | 0.50 |
| BERT (Devlin et al., 2019) | 2.19 | 0.42 |
| BART (Lewis et al., 2020) | 2.24 | 0.44 |
| HeterMPC$_{BERT}$ | 2.39 | 0.50 |
| HeterMPC$_{BART}$ | 2.41 | 0.45 |

Table 2: Human evaluation results of HeterMPC and some selected systems on a randomly sampled test set.

# Experiments

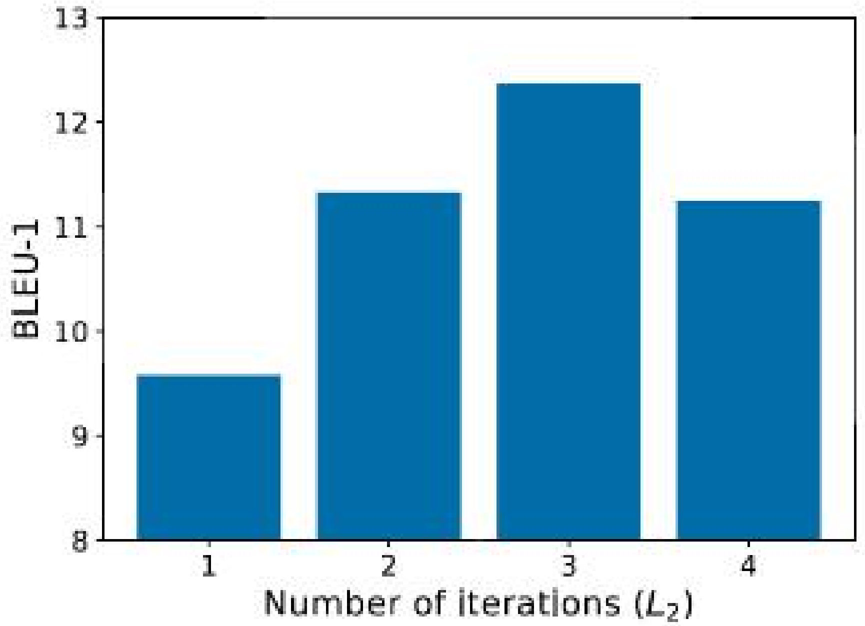



Figure 5: Performance of HeterMPC$_{BERT}$ under different numbers of iterations ($L_2$) on the test set.

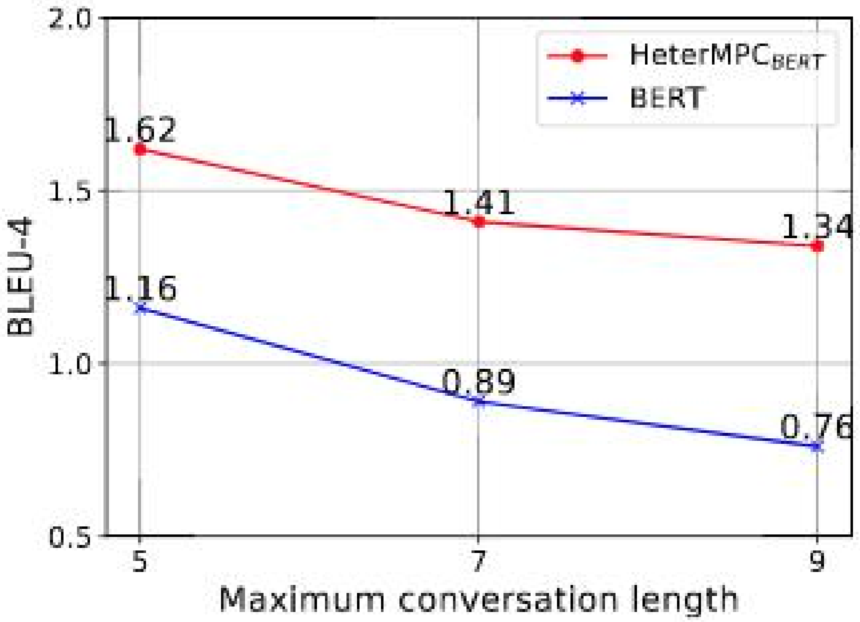



Figure 6: Performance of BERT and HeterMPC$_{BERT}$ on test samples with different session lengths.

# Experiments

| *Speaker* | *Utterance* | *Addressee* |
|---|---|---|
| | Case 1 | |
| I.1 | geev: in gparted now? | - |
| I.2 | there is no such command in my computer | I.1 |
| I.1 | open a terminal, type: sudo apt-get install gparted | I.2 |
| I.1 | after: gparted aksks for your pw, opens a gui | I.2 |
| I.2 | ok if found fail i will contact you | I.1 |
| I.3 | does it say "to install type sudo apt-get install gparted"? **(Human)** | I.2 |
| | do you have remote access to remote desktop? **(GSN)** | |
| | ok, i will try that. **(BERT)** | |
| | you can use the command "sudoapt-get install gparted" to install the gparted livecd. **(HeterMPC$_{BERT}$)** | |
| | i'm not sure what you mean by "in gparted" **(BART)** | |
| | you need to install gparted, it is a gui tool **(HeterMPC$_{BART}$)** | |

| | Case 2 | |
|---|---|---|
| I.1 | got a bit of a mount problem. fstab is still the same but the secondary harddrive shows as busy when doing mount | - |
| I.2 | you can see what is mounted with : mount | I.1 |
| I.1 | yes it is not mounted | I.2 |
| I.3 | how would i check what module the cam uses? | I.2 |
| I.2 | lsmod less will show you te loaded modules | I.3 |
| I.1 | lmount - t ext3 filepath fails cause already mounted or busy | I.2 |
| I.2 | then try to mount it at command line **(Human)** | I.1 |
| | i'm not sure how to do that **(GSN)** | |
| | i'm not sure what the problem is **(BERT)** | |
| | you need to mount it as a mount point **(HeterMPC$_{BERT}$)** | |
| | i'm not sure what the problem is **(BART)** | |
| | you need to check the filepath file **(HeterMPC$_{BART}$)** | |

# Thanks！