# 基于对话结构的生成式对话摘要技术研究与应用

## 中期答辩

汇报人：刘俊鹏

时间：2022年7月22日

# 目录

# Part 01

# 选题背景与意义

# 任务背景

- 对话摘要任务定义
  - 理解对话形式的文本，并简明扼要地概述其内容。
- 对话摘要应用场景
  - 会议、客服对话、闲聊对话、医疗问诊对话、邮件、辩论……
- 对话摘要示例

| 会议内容 |
|---|
| 工业设计师：如果我们有电源支架呢？<br>界面设计师：你可以为支架和遥控器设计一些简洁的小设计。<br>项目经理：这会增加成本……<br>项目经理：我们需要改变最终的成本。 |
| **标准摘要** |
| 工业设计师建议在设备中加入一个电源支 架，但最终被决定这不是一个有用的功能。 |

| 闲聊对话 |
|---|
| 鲍勃：老兄，你可以来接我一下吗？<br>汤姆：你在哪里？<br>鲍勃：在家，我的车坏了，我现在急需去上班，我需要你的帮助……<br>汤姆：我现在出发，10分钟之内到。 |
| **标准摘要** |
| 鲍勃的车坏了，汤姆会在10分钟内让他搭 便车，送他去上班。 |

| 医疗对话 |
|---|
| 医生：你最近有肿胀吗？<br>患者：时有时无。<br>医生：我知道了，什么时候开始的？<br>患者：大约在三周之前。<br>…… |
| **标准摘要** |
| 肿胀：大约三周之前开始，症状时有时无。 |

# 任务难点分析

- 难点分析

1. 随着对话的进行，对话中的话题会发生转换。
2. 对话摘要的关键信息常常散落在不同之处。

# 对话摘要的研究意义

- 研究意义
  - 会议摘要：帮助 与会者回顾/未参会者跟进 会议的核心内容。
  - 闲聊对话：帮助总结对话历史，快速开启新对话。
  - 客服对话摘要：帮助其他客服快速理解用户问题，提出解决方案。
  - 医疗问诊对话：帮助医生快速了解病人病情信息。
  - 邮件摘要：快速处理大量办公邮件。
  - ……

# Part 02

# 研究现状

# 对话摘要的研究现状

- 研究现状（数据集）
  - 2019年，Gliwa等人提出SAMSum数据集。
    - 由16369个闲聊对话组成，每个对话带有人工书写的抽象式摘要。
  - 2021年，Zhu等人提出MediaSum数据集
    - 由463596个新闻媒体采访对话组成，将每个采访里出现在正文前面的总体内容描述作为抽象式摘要。

0 Mia: Could anybody help me to buy a flight ticket?
1 Rebecca: Sure, but what's the problem?
2 Mia: I don't have a credit card at the moment
3 Mia: I've always used Peter's card, but now you know... I'd prefer not to
4 Tom: you can use mine!
5 Mia: Should I send you the link?
6 Tom: Just send me the flight, company and your personal data that I may need
7 Mia: great, so nice of you, thanks!

Summary: Tom will help Mia buy a flight ticket as she doesn't have a credit card and doesn’t want to use Peter's now. Tom needs the flight, company and your personal data.

SAMSum对话示例

- 研究现状
  - 1）Li et.al. (2019) 利用指针网络学习话题分段信息，帮助对话摘要的生成。

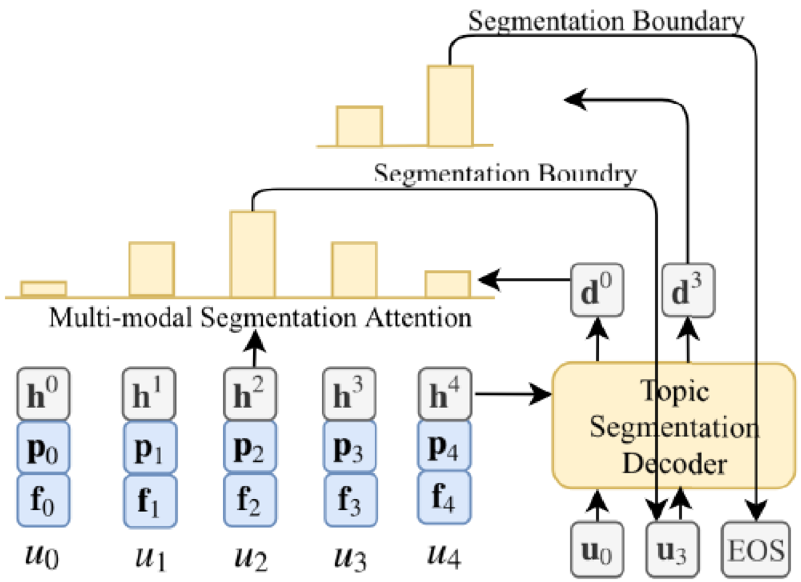



Figure 3: Topic Segmentation Decoder

Keep Meeting Summaries on Topic: Abstractive Multi-Modal Meeting Summarization

- 研究现状
  - 1）Li et.al. (2019) 利用指针网络学习话题分段信息，帮助对话摘要的生成。
  - 2）Zou et.al. (2021) 采用强化学习抽取重要句子，然后进行抽象式摘要。

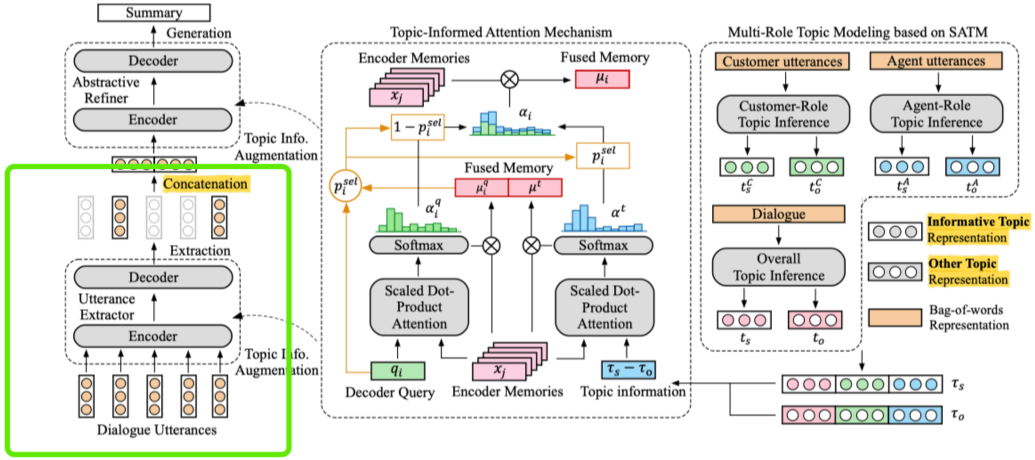



Figure 2: Overview of our proposed TDS with multi-role topic modeling based on SATM.

Topic-Oriented Spoken Dialogue Summarization for Customer Service with Saliency-Aware Topic Modeling

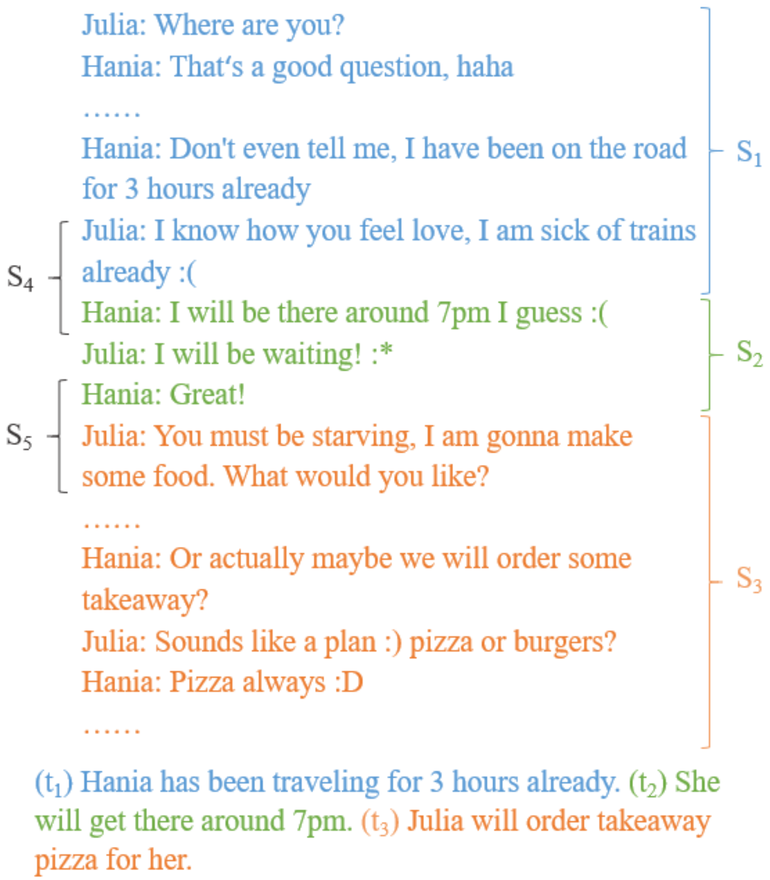
- 研究现状
  - 1）Li et.al. (2019) 利用指针网络学习话题分段信息，帮助对话摘要的生成。
  - 2）Zou et.al. (2021) 采用强化学习抽取重要句子，然后进行抽象式摘要。
  - 3）现有的工作在对对话结构（话题分段结构，话语结构等等）建模时，依赖于各式各样的人工标注，带有这类标注的数据集较少且标注成本高昂。

# Part 03

## 研究内容和目标

- 研究内容和目标
  - 文档具有天然的段落区分，且段落信息能够帮助文档摘要的生成，而对话中没有可以直接利用的段落信息。本课题从对话过程中话题转移的角度，研究对话中的话题分段结构，进而帮助对话摘要的生成。
  - 对话的话语通常存在重要性的区分，且对话摘要的目的是对其中的重要话语进行抽象式的总结。本课题研究对话中话语之间重要性的区别，即话语重要性结构，从而更好地进行对话摘要的生成。

- 研究内容一
  - 缺少人工标注信息的情况下，建模对话中的话题分段结构。
  - 如右图所示，对话中包含三个话题片段，每个话题分段表达的中心含义都在摘要中体现。



- 研究内容二
  - 话语作为对话的基本单元，建模话语之间重要性的区别，鼓励模型更多地关注重要信息，从而生成真正重要的信息。
  - 如右图所示，对话话语之间存在重要性大小的显著区别，只有其中重要话语表达的中心含义会出现在摘要中。

0 Mia: Could anybody help me to buy a flight ticket?
1 Rebecca: Sure, but what's the problem?
2 Mia: I don't have a credit card at the moment
3 Mia: I've always used Peter's card, but now you know... I'd prefer not to
4 Tom: you can use mine!
5 Mia: Should I send you the link?
6 Tom: Just send me the flight, company and your personal data that I may need
7 Mia: great, so nice of you, thanks!

Summary: Tom will help Mia buy a flight ticket as she doesn't have a credit card and doesn’t want to use Peter's now. Tom needs the flight, company and your personal data.

注：红-> 橙 -> 黄，重要性递减

# Part 04

# 研究方案设计

Beijing University of Posts and Telecommunications 北京邮电大学

- 研究内容一（话题分段结构建模）
  - 模型整体框架
    - 基础模型
      - BART
    - 连贯性分数预测模块
      - 预测片段的连贯性分数，区分话题分段边界
    - 子摘要生成模块
      - 对于话题分段，生成对应的子摘要

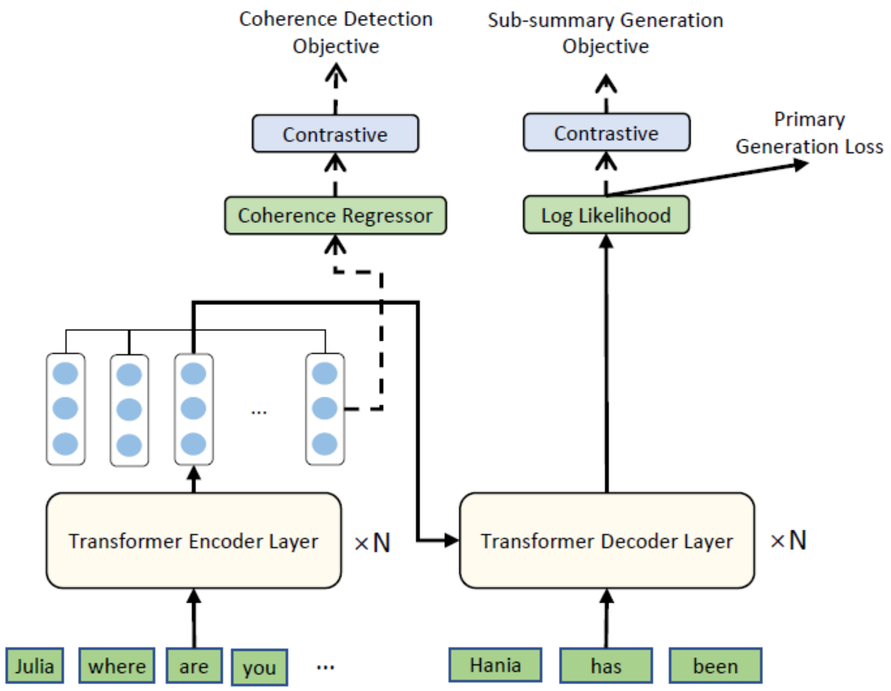



Figure 2: Model structure with contrastive objectives.

➢ 连贯性分数预测模块

- **正例**：原始对话中的连续片段
- **负例**：打乱顺序后的片段

**positive**

……
Hania: Don't even tell me, I have been on the road for 3 hours already
Julia: I know how you feel love, I am sick of trains already :(
— $S_1$

Hania: I will be there around 7pm I guess :(
Julia: I will be waiting! :*
Hania: Great!
— $S_2$

Julia: You must be starving, I am gonna make some food. What would you like?
……
Hania: Or actually maybe we will order some takeaway?
Julia: Sounds like a plan :) pizza or burgers?
Hania: Pizza always :D
……
— $S_3$

shuffle →

**negative**

Julia: I will be waiting! :*
Julia: I know how you feel love, I am sick of trains already :(
Hania: I will be there around 7pm I guess :(
Hania: Don't even tell me, I have been on the road for 3 hours already
Julia: You must be starving, I am gonna make some food. What would you like?
Hania: Great!

## ➢ 连贯性分数预测模块训练目标

- ➢ Encoder端最后一层隐层表示经过线性层预测连贯性分数
- ➢ 对比正负例的连贯性得分

$$y_{S_k^{\mathcal{D}}} = w_1 * E_{S_k^{\mathcal{D}}} + b_1; \quad y_{\widetilde{S_k^{\mathcal{D}}}} = w_1 * E_{\widetilde{S_k^{\mathcal{D}}}} + b_1$$

$y_{S_k^{\mathcal{D}}}, y_{\widetilde{S_k^{\mathcal{D}}}}$ : coherence scores

$$[co(S_k^{\mathcal{D}}), co(\widetilde{S_k^{\mathcal{D}}})] = softmax([y_{S_k^{\mathcal{D}}}, y_{\widetilde{S_k^{\mathcal{D}}}}])$$

$$\mathcal{L}_{CO}^{\mathcal{D}} = \frac{1}{N_{co}} \sum_{n=1}^{N_{co}} \max(0, \delta_{co} - (co(S_{k,n}^{\mathcal{D}}) - co(\widetilde{S_{k,n}^{\mathcal{D}}})))$$

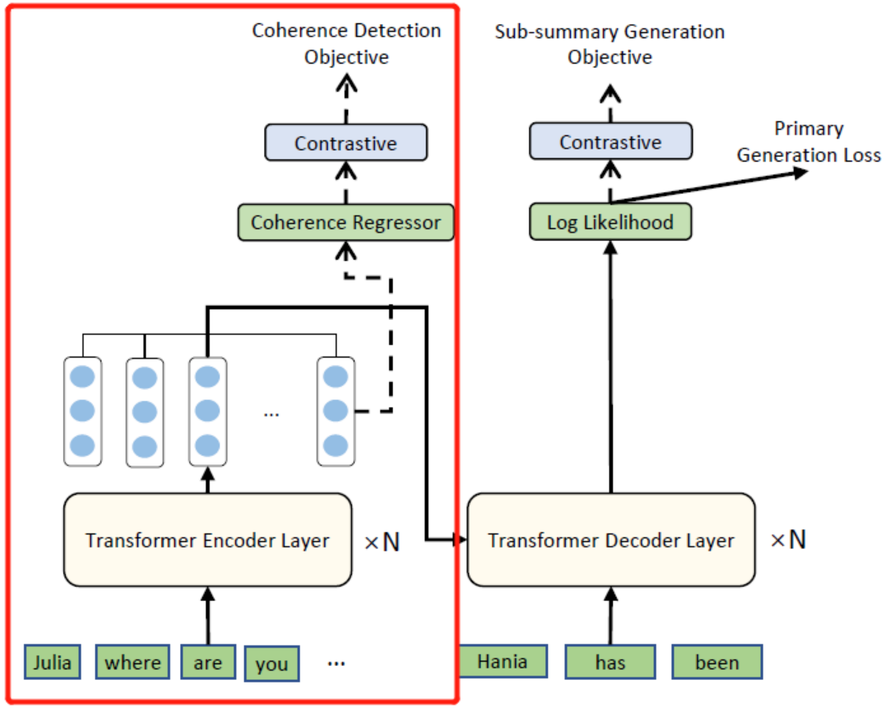



Figure 2: Model structure with contrastive objectives.

## ➢ 子摘要生成模块

- 正例: 利用和摘要的Rouge分数选择对话片段
- 负例: 随机采样对话片段作为负例

……
Hania: Don't even tell me, I have been on the road for 3 hours already
Julia: I know how you feel love, I am sick of trains already :(
— $S_1$

Positive
Hania: I will be there around 7pm I guess :(
Julia: I will be waiting! :*
Hania: Great!
— $S_2$

Julia: You must be starving, I am gonna make some food. What would you like?
……
Hania: Or actually maybe we will order some takeaway?
Julia: Sounds like a plan :) pizza or burgers?
Hania: Pizza always :D
……
— $S_3$

($t_1$) Hania has been traveling for 3 hours already.
**($t_2$) She will get there around 7pm.**
($t_3$) Julia will order takeaway pizza for her.

**Algorithm 1** Snippet selection for a sub-summary

**Input:** A sub-summary $t_i \in T$, a dialogue $\mathcal{D}$ containing $|\mathcal{D}|$ utterances, sliding window size interval $[a, b]$
**Output:** $(S^i_{\text{pos}}, S^i_{\text{neg}})$ for $t_i$

$\mathcal{W} = \emptyset$
**for** $w = a$ to $b$ **do**
  **for** $j = 1$ to $|\mathcal{D}| - w$ **do**
    cand $= \mathcal{D}_{j,j+w}$
    $r(j, w) \leftarrow \text{ROUGE}(\text{cand}, t_i)$
    $\mathcal{W} \leftarrow \mathcal{W} \cup \text{cand}$
    $j \leftarrow j + w/2$
  $w \leftarrow w + 1$
$j_{\text{best}}, w_{\text{best}} \leftarrow \arg\max_{j,w} r(j, w)$
$S^i_{\text{pos}} \leftarrow \mathcal{D}_{j_{\text{best}},(j_{\text{best}}+w_{\text{best}})}$
$S^i_{\text{neg}} \leftarrow \mathcal{W} \setminus S^i_{\text{pos}}$

以第二个子摘要($t_2$ ) 举例说明

- 子摘要生成模块

- 正例: 利用和摘要的Rouge分数选择对话片段
- 负例: 随机采样对话片段作为负例

……
Hania: Don't even tell me, I have been on the road for 3 hours already
Julia: I know how you feel love, I am sick of trains already :(

$S_1$

**Positive**

Hania: I will be there around 7pm I guess :(
Julia: I will be waiting! :*
Hania: Great!

$S_2$

Julia: You must be starving, I am gonna make some food. What would you like?
……
Hania: Or actually maybe we will order some takeaway?
Julia: Sounds like a plan :) pizza or burgers?
Hania: Pizza always :D
……

$S_3$

($t_1$) Hania has been traveling for 3 hours already.
**($t_2$) She will get there around 7pm.**
($t_3$) Julia will order takeaway pizza for her.

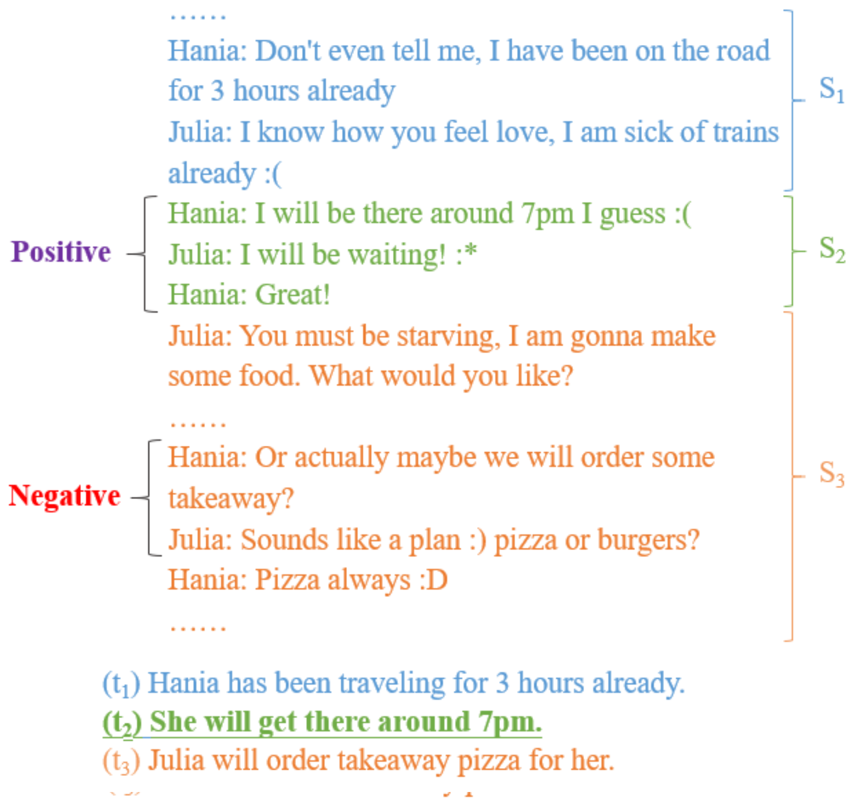
**Algorithm 1** Snippet selection for a sub-summary

**Input:** A sub-summary $t_i \in T$, a dialogue $\mathcal{D}$ containing $|\mathcal{D}|$ utterances, sliding window size interval $[a, b]$
**Output:** $(S^i_{\text{pos}}, S^i_{\text{neg}})$ for $t_i$
$\mathcal{W} = \emptyset$
**for** $w = a$ to $b$ **do**
  **for** $j = 1$ to $|\mathcal{D}| - w$ **do**
    $\text{cand} = \mathcal{D}_{j,j+w}$
    $r(j, w) \leftarrow \text{ROUGE}(\text{cand}, t_i)$
    $\mathcal{W} \leftarrow \mathcal{W} \cup \text{cand}$
    $j \leftarrow j + w/2$
  $w \leftarrow w + 1$
$j_{\text{best}}, w_{\text{best}} \leftarrow \arg\max_{j,w} r(j, w)$
$S^i_{\text{pos}} \leftarrow \mathcal{D}_{j_{\text{best}},(j_{\text{best}}+w_{\text{best}})}$
$S^i_{\text{neg}} \leftarrow \mathcal{W} \setminus S^i_{\text{pos}}$

以第二个子摘要($t_2$) 举例说明

➢ 子摘要生成模块

- 正例: 利用和摘要的Rouge分数选择对话片段
- 负例: 随机采样对话片段作为负例




**Algorithm 1** Snippet selection for a sub-summary

**Input:** A sub-summary $t_i \in T$, a dialogue $\mathcal{D}$ containing $|\mathcal{D}|$ utterances, sliding window size interval $[a, b]$

**Output:** $(S^i_{\text{pos}}, S^i_{\text{neg}})$ for $t_i$

$\mathcal{W} = \emptyset$
**for** $w = a$ to $b$ **do**
  **for** $j = 1$ to $|\mathcal{D}| - w$ **do**
    $\text{cand} = \mathcal{D}_{j,j+w}$
    $r(j, w) \leftarrow \text{ROUGE}(\text{cand}, t_i)$
    $\mathcal{W} \leftarrow \mathcal{W} \cup \text{cand}$
    $j \leftarrow j + w/2$
  $w \leftarrow w + 1$
$j_{\text{best}}, w_{\text{best}} \leftarrow \arg\max_{j,w} r(j, w)$
$S^i_{\text{pos}} \leftarrow \mathcal{D}_{j_{\text{best}},(j_{\text{best}}+w_{\text{best}})}$
$S^i_{\text{neg}} \leftarrow \mathcal{W} \setminus S^i_{\text{pos}}$

以第二个子摘要($t_2$) 举例说明

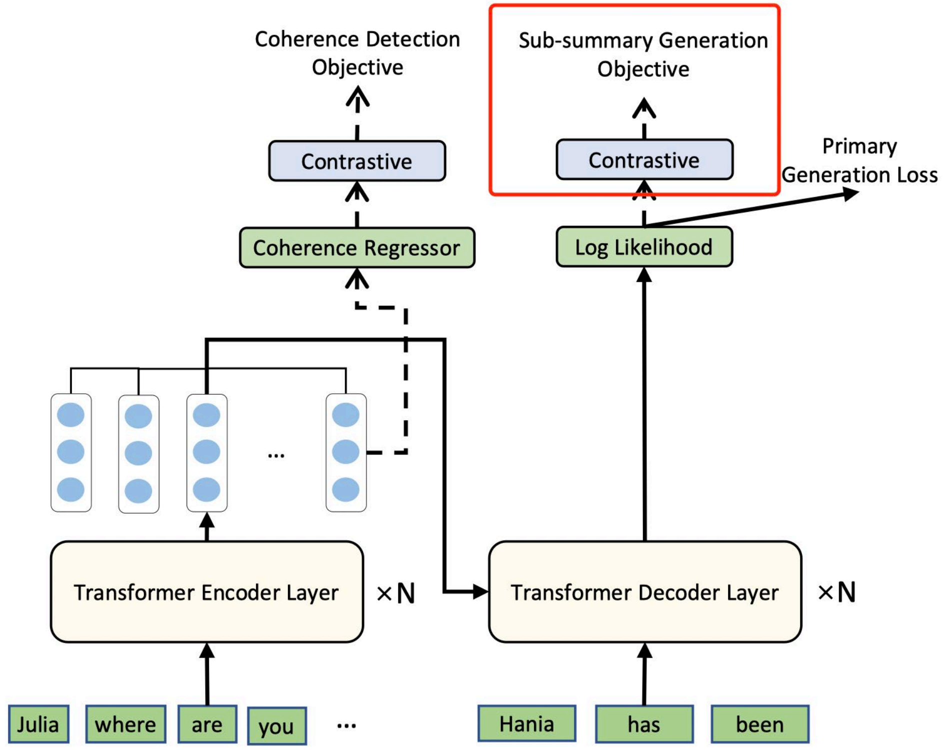
➢ 子摘要生成模块训练目标

$S^i_{\text{pos}}\ S^i_{\text{neg}}$ a contrastive pair $\quad t_i$ sub-summary

$$\mathcal{L}^{t_i}_{pos} = -\log(\prod_{j=1}^{|t_i|} p(t^i_j | t^i_{1:j-1}, \mathcal{S}^i_{\text{pos}}; \theta)) \qquad \mathcal{L}^{t_i}_{neg} = -\log(\prod_{j=1}^{|t_i|} p(t^i_j | t^i_{1:j-1}, \mathcal{S}^i_{\text{neg}}; \theta))$$

$$[su(S^i_{\text{pos}}), su(S^i_{\text{neg}})] = softmax([\mathcal{L}^{t_i}_{pos}, \mathcal{L}^{t_i}_{neg}])$$

$$\mathcal{L}^{\mathcal{D},T_{\mathcal{D}}}_{su} = \frac{1}{N_{su}} \sum_{n=1}^{N_{su}} \max(0, \delta_{su} - (su(S^n_{\text{neg}}) - su(S^n_{\text{pos}})))$$



- 实验结果分析（研究内容一）
  - 在SAMSum数据集上，融合话题分段信息后的模型相对于基线模型有显著提升
  - 和之前的SOTA模型相比，我们的方法不依赖于话题分段的标注信息，在ROUGE分数和BERTScore上都有更好的效果
  - 消融实验证明两个模块对于摘要质量都有提升

| Model | R-1 | R-2 | R-L | BERTS |
|---|---|---|---|---|
| *Lead3 | 31.4 | 8.7 | 29.4 | - |
| *PTGen | 40.1 | 15.3 | 36.6 | - |
| *DynamicConv + GPT-2 | 41.8 | 16.4 | 37.6 | - |
| *FastAbs-RL | 42.0 | 18.1 | 39.2 | - |
| *DynamicConv + News | 45.4 | 20.7 | 41.5 | - |
| Multiview BART | 53.9 | 28.4 | 44.4 | 53.6 |
| *$\text{BART}_{\text{BASE}}$ | 46.1 | 22.3 | 36.4 | 44.8 |
| *BART | 52.6 | 27.0 | 42.1 | 52.1 |
| *$\text{BART}_{\text{ORI}}$ | 52.6 | 27.2 | 42.7 | 52.3 |
| $\text{CONDIGSUM}_{\text{BASE}}$ | 48.1 | 24.0 | 39.2 | 48.0 |
| CONDIGSUM | **54.3** | **29.3** | **45.2** | **54.0** |
| w/o Sub-summary | 53.8 | 28.3 | 44.1 | 53.5 |
| w/o Coherence | 53.9 | 28.6 | 44.2 | 53.5 |

Table 1: Results on SAMSum test split. $*$ indicates that the results are significantly different from ours ($p < 0.05$).

- 实验结果分析（研究内容一）
  - 在MediaSum数据集上，我们的模型也超过了4个基线模型
  - 同时，消融实验证明了两个模块的有效性

| Model | R-1 | R-2 | R-L | BERTS |
|---|---|---|---|---|
| *Lead3 | 15.0 | 5.1 | 13.3 | - |
| *PTGen | 28.8 | 12.2 | 24.2 | - |
| *UniLM | 32.7 | 17.3 | 29.8 | - |
| *BART | 34.7 | 17.7 | 30.9 | 30.7 |
| *BART$_{ORI}$ | 35.0 | 17.9 | 31.1 | 31.2 |
| CONDIGSUM | **36.0** | **18.9** | **32.2** | **32.4** |
| w/o Sub-summary | 35.5 | 18.7 | 31.9 | 32.0 |
| w/o Coherence | 35.5 | 18.6 | 31.7 | 31.9 |

Table 2: Results on MediaSum test split. $*$ indicates that the results are significantly different from ours ($p < 0.05$).

- 实验结果分析（研究内容一）
  - 连贯性分数预测模块和子摘要生成模块的作用大小
  - 学习权重在0.1~10倍的范围内变化时，摘要质量对其变化不敏感
  - 过大的学习权重（>100倍）会导致开始模型弱化主摘要任务
  - 辅助任务的学习权重要适中

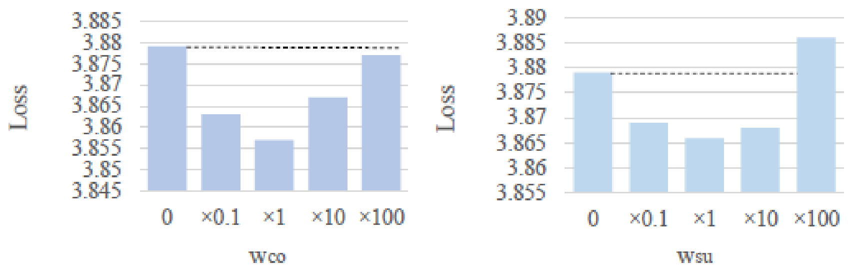
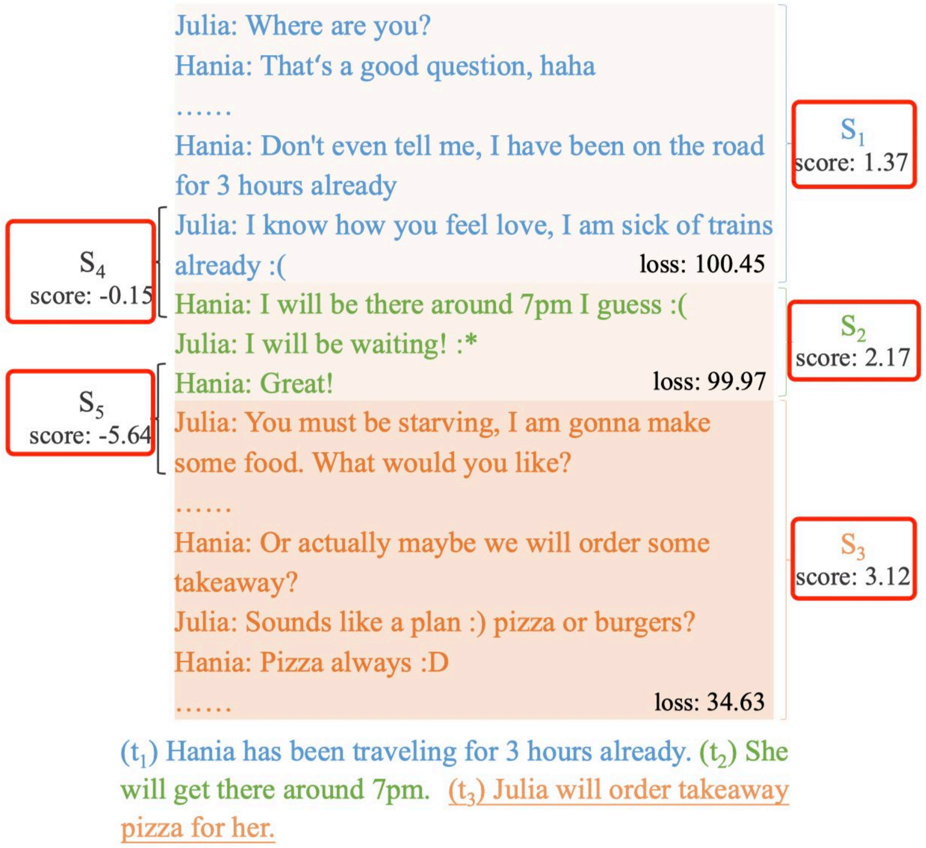



Figure 4: Impact of different values of task-coefficients of coherence detection (left) and sub-summary generation (right) objectives on the validation loss of the primary dialogue summarization task.

- 可视化结果（研究内容一）
  - 模型对于没有发生话题转移的片段预测出了高连贯性分数（S1, S2, S3），对于发生话题转移的片段预测出了低连贯性分数（S4, S5）



- 可视化结果（研究内容一）
  - 可视化子摘要生成目标，我们的模型能够更集中到关键信息，并生成更相关的摘要

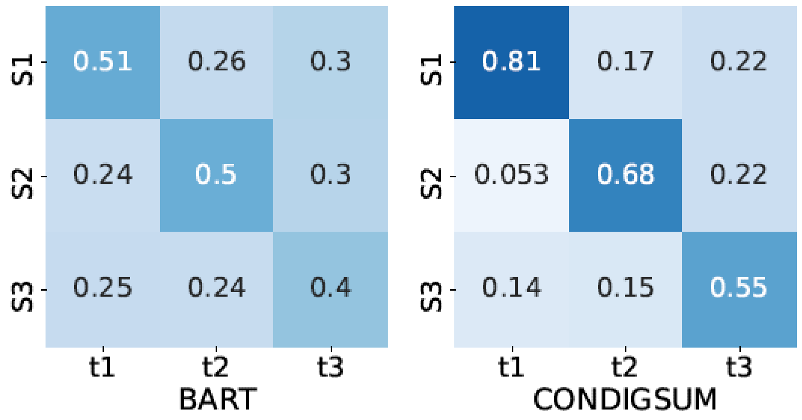



Figure 3: Visualization of how much a sub-summary is related to different snippets (the sum of every column is equal to 1). The result of CONDIGSUM is more concentrated on diagonal.

- 研究内容二（话语重要性结构建模）
  - **第一步，获取不同话语对于摘要文本的重要性**
    - 以预训练的文档摘要Bart-large-cnn模型的生成loss作为评测话语针对真实摘要的重要性得分的度量
    - 依次迭代删除对于真实摘要生成影响最小的话语，即相对不重要的话语
    - 最终可得到如下图所示的话语重要性得分

0 Mia: Could anybody help me to buy a flight ticket?
1 Rebecca: Sure, but what's the problem?
2 Mia: I don't have a credit card at the moment
3 Mia: I've always used Peter's card, but now you know... I'd prefer not to
4 Tom: you can use mine!
5 Mia: Should I send you the link?
6 Tom: Just send me the flight, company and your personal data that I may need
7 Mia: great, so nice of you, thanks!

0.00 ===== unimportant
0.14 =====
0.29 =====
0.43 =====
0.57 =====
0.71 =====
0.86 =====
1.00 ===== important

注：红-> 橙 -> 黄，重要性递减

Summary: Tom will help Mia buy a flight ticket as she doesn't have a credit card and doesn’t want to use Peter's now. Tom needs the flight, company and your personal data.

话语重要性评分结果示例

- 研究内容二（话语重要性结构建模）
  - 测试预训练模型进行对话语重要性排序的性能
    - 人工评测，50个对话，将每个对话的话语分成重要/不重要两组
    - 对于人工划分为重要的话语，预训练模型预测其重要性名次，计算平均名次
  - 发现，大模型预测出对话重要性排序性能还可以
    - 比random高2.32
    - 比human低1.00

| | 重要话语平均名次（越小越好） |
|---|---|
| random | 4.66 |
| bart-large-cnn | 2.34 |
| human | 1.34 |

- 研究内容二（话语重要性结构建模）
  - 第一步，获取不同话语对于摘要文本的重要性
  - **第二步，利用话语重要性结构帮助对话摘要的生成**
    - 第一种方法，区分重要话语和不重要话语，针对重要话语进行数据增强
    - 第二种方法，利用对比学习把话语重要性的排序融合到对话摘要模型中

- 区分重要话语和不重要话语，针对重要话语进行数据增强
  - 1）只保留重要话语，摘要不变
  - 2）混合两个对话的重要话语，新摘要为两个对话各自摘要的拼接
  - 3）保留对话的重要话语，用另一个对话的不重要话语对其中不重要话语进行替换，摘要不变

- 实验结果（数据增强）
  - 针对话语重要性结构进行的数据增强相对baseline提升了对话摘要的性能
  - 其中，混合两个对话的重要话语的增强方式（aug-2），是其中最好的增强方式

| | Rouge 1/2/L |
|---|---|
| bart-base | val: 44.90/20.71/37.12 test: 44.35/19.35/35.72 |
| bart-base + aug-1(single dialogue, sal ->summary) | val: 46.74/21.37/37.57 test: 45.93/19.99/36.36 |
| **bart-base + aug-2(two dialogues, D1_sal + D2_sal ->S1 + S2)** | val: 46.75/21.84/38.03 test: 46.29/20.62/37.30 |
| bart-base + aug-3(two dialogues, D1_sal + D2_unsal -> S1) | val: 46.49/21.05/37.35 test: 45.34/19.64/36.05 |

- 利用对比学习把话语重要性的排序融合到对话摘要模型中
  - 对比样例构建
    - 删除不重要话语
    - 用同batch其他样本的编码表示替换
      - 混入了其他对话的人名，命名实体

$X_1$

0 Mia: Could anybody help me to buy a flight ticket?
~~1 Rebecca: Sure, but what's the problem?~~
2 Mia: I don't have a credit card at the moment
3 Mia: I've always used Peter's card, but now you know... I'd prefer not to
4 Tom: you can use mine!
~~5 Mia: Should I send you the link?~~
6 Tom: Just send me the flight, company and your personal data that I may need
7 Mia: great, so nice of you, thanks!

$X_2$

0 Mia: Could anybody help me to buy a flight ticket?
~~1 Rebecca: Sure, but what's the problem?~~
2 Mia: I don't have a credit card at the moment
3 Mia: I've always used Peter's card, but now you know... I'd prefer not to
~~4 Tom: you can use mine!~~
~~5 Mia: Should I send you the link?~~
6 Tom: Just send me the flight, company and your personal data that I may need
7 Mia: great, so nice of you, thanks!

$X_3$

0 Mia: Could anybody help me to buy a flight ticket?
~~1 Rebecca: Sure, but what's the problem?~~
~~2 Mia: I don't have a credit card at the moment~~
~~3 Mia: I've always used Peter's card, but now you know... I'd prefer not to~~
~~4 Tom: you can use mine!~~
~~5 Mia: Should I send you the link?~~
6 Tom: Just send me the flight, company and your personal data that I may need
~~7 Mia: great, so nice of you, thanks!~~

$X_4, X_5, \cdots\cdots$

0 Mia: Could anybody help me to buy a flight ticket?
1 Rebecca: Sure, but what's the problem?
2 Mia: I don't have a credit card at the moment
3 Mia: I've always used Peter's card, but now you know... I'd prefer not to
4 Tom: you can use mine!
5 Mia: Should I send you the link?
6 Tom: Just send me the flight, company and your personal data that I may need
7 Mia: great, so nice of you, thanks!

$X_0$

- 利用对比学习把话语重要性的排序融合到对话摘要模型中
  - 话语重要性排序的训练目标
    - 输入中包含的高重要性得分话语越多，对应输出的似然概率也越大

$$\mathrm{Loss1} = \sum_{(x,y)\in D} \sum_{x_i} \sum_{j>i} \max(0, f(x_j) - f(x_i) + (j - i) * \lambda)$$

$$f(x) = \frac{\sum_{t=1}^{|y|} log p_{g\theta}(y_t|x, y_{<t}; \theta)}{|y|^a}$$

- 利用对比学习把话语重要性的排序融合到对话摘要模型中
    - 增强正例生成loss的训练目标

$$\text{Loss2} = \sum_{(x,y)\in D} \sum_{x_i} \sum_{j>i} \max((j - i) * \lambda * 2, f(x_j) - f(x_i))$$

$$f(x) = \frac{\sum_{t=1}^{|y|} log p_{g\theta}(y_t | x, y_{<t}; \theta)}{|y|^a}$$

- 利用对比学习把话语重要性的排序融合到对话摘要模型中
  - 融合话语重要性结构的模型超过了基线模型
  - 增强新构造的正例生成loss没有带来额外提升

| | validation rouge | test rouge |
|---|---|---|
| BART（baseline） | 52.34/28.00/43.49 | 51.08/26.32/42.24 |
| **BART + 排序话语对比** | **52.81/28.49/44.04** | **51.78/27.02/42.88** |
| BART + 排序话语对比 + 增强正例生成loss | 52.06/27.91/43.73 | 50.93/26.40/42.39 |

# Part 05

## 未来工作目标

- 未来工作目标
  - 研究内容二中，分析对话摘要模型预测重要话语排序能力和预训练模型的差距
  - 进一步改进话语重要性结构的建模方法
  - 搭建对话摘要系统，完成对话文本的自动摘要功能的展示

# 请老师批评指正