**任务：面向通信领域的过程类知识抽取**

# 任务描述

通信领域存在多种的过程类知识，如硬件安装（基站主设备安装操作步骤）、参数配置（配置网元开通与对接相关的参数）、集成调测（网元开通调试和功能验证）、故障处理（修复网元开通或正常运行中出现的故障）等， 其中故障处理过程类知识尤为重要。通信运维过程中，通过“事件”及“事件关系”对故障过程知识进行梳理，给用户呈现故障发生的逻辑，提供故障排查和故障恢复方案，指导一线处理现网故障。在故障知识整理过程中，“事件抽取”与“事件共指消解”是实现故障脉络、排查步骤和恢复步骤梳理的重要手段。 通信领域“事件抽取”的挑战在于通信业务的复杂和任务本身的复杂，其中通信业务的在于复杂的领域长词、歧义事件、元素共用等，而任务本身的复杂在于多任务识别（包括触发词和角色识别）、元素间的依赖关系等。“事件共指消解”的难点在于事件元素表述多样化和事件元素缺损（漏抽取、文本描述缺损）。

本次评测任务的语料来源主要是华为公司的公开故障处理案例。本次评测任务的事件类型包括：指标恶化类、软硬件异常、采集数据、核查类、配置类故障、外部事件、调整机器、操作机器等。

本次评测包括两个子任务：事件抽取和事件共指消解。具体内容如下：

**子任务一：事件抽取**

**该任务目标是从自由文本中抽取事件触发词及事件角色**：即给定自由文本T，抽取文本T中所有的事件集合E，对于E中的每个事件e，从文本T中抽取e的触发词（包含词、位置、分类）和各个角色（包含词、位置、分类）。

**输入**：一段文本T

**输出**：事件触发词和事件角色

**示例**：

***样例1***

*输入：*“将直放站小区半径参数由2000m改为10000m”

*输出：{"trigger": ["SetMachine", 14, "改为"], "argument": [["Object", 1, "直放站小区半径"], ["InitialState", 9, "2000m"], ["FinalState", 16, "10000m"]]}*

***样例2***

*输入：“XX小区晚上8点之后苹果终端接入5G网络失败*”

*输出： {"trigger": ["SoftHardwareFault", 14, "接入"], "argument": [["Subject", 10, "苹果终端"], ["Object", 16, "5G网络"], ["State", 20, "失败"]]}]*

其中各事件元素及其说明如下表：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 指标恶化类 | 事件要素 | IndexFault | Index | Owner | State |  |
| 说明 | 触发词，如下降、低、偏大等 | 技术指标，如掉话率、接通率等 | 指标主体，如核心网、站点等 | 网络制式，如严重、长时间等 |  |
| 软硬件异常 | 事件要素 | SoftHardwareFault | Subject | Object | State | Owner |
| 说明 | 触发词，如脱网、掉线等 | 故障发生的位置，如手机、基站等 | 故障相关的宾语，如信息等 | 故障状态，如失败、频繁 | 故障发生位置所属的实体，如小区、核心网等 |
| 采集数据 | 事件要素 | CollectData | Object | Source |  |  |
| 说明 | 触发词，如采集、导出、下载 | 采集对象，如配置文件、告警日志等 | 数据来源，如PRS系统、NodeB等 |  |  |
| 核查类 | 事件要素 | Check | Object | Standard |  |  |
| 说明 | 触发词，如核查、对比、查看等 | 核查对象，如指标状态、参数配置等 | 核查标准，如特性手册、设计图纸等 |  |  |
| 配置类故障 | 事件要素 | SettingFault | Setting | Owner | Reference | State |
| 说明 | 触发词，如配置、漏配、不一致等 | 配置，如邻区参数、QOS参数等 | 配置所属实体，如LTE小区、基站等 | 参考对象，如正常参数值、参考手册等 | 状态，如错误、失败等 |
| 外部事件 | 事件要素 | ExternalFault | State |  |  |  |
| 事件要素类型 | 触发词，如疫情、下雪、雷暴等 | 状态，如严重、长时间等 |  |  |  |
| 调整机器 | 事件要素 | SetMachine | Object | Network | InitialState | FinalState |
| 事件要素类型 | 触发词，如设置、替换、升级等 | 设置对象，如参数配置、版本、RRU | 网络，如GSM、LTE、高铁专网等 | 初始状态，如具体的版本或参数配置等 | 末态，如具体的版本或参数配置等 |
| 操作机器 | 事件要素 | Operate | Object | Owner |  |  |
| 事件要素类型 | 触发词，如复位、打开、开通等 | 操作对象，如单板、RRU、路由器等 | 操作对象主体，如基站、核心网等 |  |  |

**子任务二：事件共指消解任务**

**该任务的目的是给定两个自由文本及其中的事件判定两个事件是否为同一个事件。**即给定文本T1、T2、T1中事件e1和T2中事件e2，判定e1和e2是否相等（其中事件类型及元素表如子任务一介绍）

**样例1**

输入：

e1: {"text": "PTN链上端站交叉板隐性问题导致部分TDS站点下PS掉线率指标恶化及交互类业务异常问题", "trigger": ["IndexFault", 31, "恶化"], "argument": [["Index", 18, "TDS站点下PS掉线率指标"]]},

e2: {"text": "PS掉线率升高的问题", "trigger": ["IndexFault", 5, "升高"], "argument": [["Index", 0, "PS掉线率"]]} }

输出：True

**样例2：**

 e1: {"text": "特定用户协同功能开关未开启导致5GCPE无法接入19B版本的NSA组网", "trigger": ["SoftHardwareFault", 22, "接入"], "argument": [["Subject", 15, "5GCPE"], ["State", 20, "无法"], ["Object", 24, "19B版本的NSA组网"]]},

 e2: {"text": "5G 终端无法入网。", "trigger": ["SoftHardwareFault", 7, "入网"], "argument": [["Subject", 0, "5G 终端"], ["State", 5, "无法"]]}}

输出：False

# 数据描述

本次数据主要来源于华为外部公开的故障处理案例，其中包括事件抽取数据和事件共指消解数据，训练集、验证集及测试集的说明如下：

**子任务一：**

**训练集&验证集：**

在训练及验证数据发布阶段，我们会发布15k条左右的文本及其所标注事件元素和2k条左右的验证文本。训练集每行为一个json文档, “doc\_id”问题文本ID， “text”为文本，“event\_list”为文本的事件列表，如下：

{"doc\_id": "0579", "text": "基带板闭塞导致小区不可用", "event\_list": [{"trigger": ["SoftHardwareFault", 3, "闭塞"], "argument": [["Subject", 0, "基带板"]]}, {"trigger": ["SoftHardwareFault", 9, "不可用"], "argument": [["Subject", 7, "小区"]]}]}

验证集每行为一个json：

{"doc\_id": "0579", "text": "基带板闭塞导致小区不可用",}

提交格式：

{"doc\_id": "0579", "event\_list": [{"trigger": ["SoftHardwareFault", 3, "闭塞"], "argument": [["Subject", 0, "基带板"]]}, {"trigger": ["SoftHardwareFault", 9, "不可用"], "argument": [["Subject", 7, "小区"]]}] }

**测试集：**

在测试数据发布阶段，我们将会再发布2k千条左右的文本数据集（另外会有一定数量的干扰数据不参与结果评测），不含标注结果，作为测试。每行为一个json，

{"doc\_id": "0579"，"text": "基带板闭塞导致小区不可用"}

提交格式：

{"doc\_id": "0579", "event\_list": [{"trigger": ["SoftHardwareFault", 3, "闭塞"], "argument": [["Subject", 0, "基带板"]]}, {"trigger": ["SoftHardwareFault", 9, "不可用"], "argument": [["Subject", 7, "小区"]]}]}

*说明:一个文本可能有多个事件类，请预测出所有的事件及其关键角色，所以事件存在event\_list字段，每个事件为一个字典，“trigger”为触发词信息，“argument”为所有角色列表，每个文本的预测结果保存在一个json中，写入文本的一行*

**任务二：**

**训练集&验证集：**

在训练及验证数据发布阶段，我们会发布15k条左右的事件对和2k条验证集。训练集每行为一个json格式，“pair\_id”为事件对ID，“eventA”和“eventB”为事件的文本及其元素(包含text，trigger，argument三个字段)，label为标签（true共指，false非共指），具体格式示例如下：

{ "pair\_id": "XXX102", "eventA": {"text": "特定用户协同功能开关未开启导致5GCPE无法接入19B版本的NSA组网", "trigger": ["SoftHardwareFault", 22, "接入"], "argument": [["Subject", 15, "5GCPE"], ["State", 20, "无法"], ["Object", 24, "19B版本的NSA组网"]]}, "eventB": {"text": "5G终端无法入网。", "trigger": ["SoftHardwareFault", 6, "入网"], "argument": [["Subject", 0, "5终端"], ["State", 4, "无法"]]}, "label": false }

验证集格式：

{"pair\_id": "XXX103" , "eventA": {"text": "特定用户协同功能开关未开启导致5GCPE无法接入19B版本的NSA组网", "trigger": ["SoftHardwareFault", 22, "接入"], "argument": [["Subject", 15, "5GCPE"], ["State", 20, "无法"], ["Object", 24, "19B版本的NSA组网"]]}, "eventB": {"text": "5G终端无法入网。", "trigger": ["SoftHardwareFault", 6, "入网"], "argument": [["Subject", 0, "5终端"], ["State", 4, "无法"]]}}

提交格式：

{ "pair\_id": "XXX103", "label": false }

**测试集：**

在测试数据发布阶段，我们将会再发布2k条左右的文本数据集（另外会有一定数量的干扰数据不参与结果评测），其中只有的事件标注，不含共指标注结果，作为测试。格式为：

{"pair\_id": "XXX103", "eventA": {"text": "特定用户协同功能开关未开启导致5GCPE无法接入19B版本的NSA组网", "trigger": ["SoftHardwareFault", 22, "接入"], "argument": [["Subject", 15, "5GCPE"], ["State", 20, "无法"], ["Object", 24, "19B版本的NSA组网"]]}, "eventB": {"text": "5G终端无法入网。", "trigger": ["SoftHardwareFault", 6, "入网"], "argument": [["Subject", 0, "5终端"], ["State", 4, "无法"]]} }

提交结果格式：

{"pair\_id": "XXX103", "label": false}

# 评价指标

**任务1评价指标：**

本次任务采用精确率（Precision, P）、召回率（Recall, R）、F1值（F1-measure, F1）来评估事件要素的识别效果。采用微平均计算F值即所有样本的元素（触发词和角色）一起计算P和R。事件的触发词与每一个角色都是统计项。

给定一个文本，假设答案包含两个事件结果:

E1: 触发词：value0， role1：value1，role2：value2，role3：value3，role4：value4

E2: 触发词：value5，role1：value6，role2：value7，role3：value8

假设选手给出两个预测结果：

P1: 触发词：value0，role1：value9，role2：value2，role3：value3

P2: 触发词：value10，role1：value6，role2：value7

评测时首先匹配触发词，如果触发词相同则统计角色准确率否则角色全部算错，例如上例中P1和E1触发词相同且2个角色相同(+3)，P2和E2触发词不同(+0)，可以计算Precision和Recall，如下

Precision = (3+0)/(4+3)， Recall = (3+0)/(5+4)

注：在给每条预测结果找其对应的标注结果时，首先找到触发词的相同（词、位置和类型都相同）的标注， 然后计算相同的事件元素（包括触发词和角色）。

$$事件要素精确率（Precission）=\frac{预测事件元素与标注事件元素重合数量}{预测事件元素总量}$$

$$事件要素召回率（Recall）=\frac{预测事件元素与标注事件元素重合数量}{标注事件元素总量}$$

$$事件要素F1值=\frac{2\*事件要素精确率\*事件要素召回率}{事件要素精确率+事件要素召回率}$$

**任务2评价指标：**

本次任务采用准确率（Accuracy, A）来评估事件共指预测的效果。

给定一组事件对（4组）：

Event\_pair1:{event1, event2, label:true}

Event\_pair2:{event3, event4, label:true}

Event\_pair3:{event5, event6, label:false}

Event\_pair4:{event7, event8, label:false}

假设选手预测的结果如下：

Event\_pair1:{event1, event2, label:true}

Event\_pair2:{event3, event4, label:false}

Event\_pair3:{event5, event6, label:false}

Event\_pair4:{event7, event8, label:true}

则TP=1， FN=1， TN=1， FP=1

$$共指准确率Accuracy=\frac{TP+TN}{TP+TN+FP+FN}$$

# 任务提交

本次任务将采取刷榜的方式，验证集发布后，允许参赛队伍多次向平台提交结果，文件命名为“参赛队名称\_valid\_result.txt”，格式与结果文件result.txt相同，排名随时更新。参赛队伍可在评测集发布之前随时上传验证集的计算结果（每日最多可上传3次），管理系统会及时更新各队伍的最新排名情况；

测试集发布后，允许参赛队伍多次提交结果文件，结果文件提交格式如下。

**结果文件格式说明：**

选手将结果保存为result.txt，以utf-8编码格式保存。

对于子任务一：

{”doc\_id”:”xxx”,”event\_list”:[{“trigger”:[xxx],”argument”:[xxx]}]}

对于任务二：

{”pair\_id”:”xxx”,”label”:true/false }

最终提交文件要求：

每一个参赛队需提交的材料如下：

1. 抽取结果运行结果文件
2. 代码及说明
3. 方法描述文档（非评测论文，评测论文撰写要求见CCKS 2021官网）

以上三个文件需在任务提交截止日期前发送至邮箱：

**子任务1：**发送至ccks2021taskcpe\_t1@126.com。邮件的标题为：“CCKS-评测任务CPE子任务1-最终提交文件-参赛队名称”，例如“CCKS-评测任务CPE-子任务1-最终提交文件-火箭队”。

**子任务2：**发送至ccks2021taskcpe\_t2@126.com。邮件的标题为：“CCKS-评测任务FEE-子任务2-最终提交文件-参赛队名称”，例如“CCKS-评测任务CPE-子任务2-最终提交文件-火箭队”。

结果文件用result.txt命名，采用utf-8格式存储，文件格式需要与训练数据中的标注结果文件完全一致。

代码及其文档需打包成一个文件（tar，zip，gzip，rar等均可），用code.xxx命名，要求提交所有的程序代码及相关的配置说明，确保程序能够正确运行，且所得结果与result.txt相符。

方法描述文档用SM.pdf命名，包含算法描述及参数设置，需用pdf格式存储，页数不超过10页。

评测平台：本次评测将依托biendata平台（https://biendata.com/）展开，请有意向的参赛队伍关注平台上的竞赛列表。

# 时间安排

• 评测任务发布：4月1日

• 报名时间：4月1日—7月15日

• 训练及验证数据发布：4月20日

•   验证集提交开放：4月20日23:59 - 7月20日23:59 （北京时间）

•     测试数据发布：7月15日23:59（北京时间）

•     测试集提交开放： 7月20日23:59 - 8月5日23:59（北京时间）

•     评测论文提交：8月5日

•     CCKS会议日期：8月18日—21日

# 规则

1. 所有参赛选手都必须在biendata平台管理系统中注册，本次比赛的参赛对象仅限全日制在校大学生（本科、硕士、博士均可）和企业员工；

2. 参赛选手需确保注册时提交信息准确有效，所有的比赛资格及奖金支付均以提交信息为准；

3. 参赛选手在管理系统中组队，参赛队伍成员数量不得超过5个，报名截止日期之后不允许更改队员名单；

4. 每支队伍需指定一名队长，队伍名称不超过15个字符，队伍名的设定不得违反中国法律法规或公序良俗词汇，否则组织者有可能会解散队伍；

5. 每名选手只能参加一支队伍，一旦发现某选手以注册多个账号的方式参加多支队伍，将取消相关队伍的参赛资格；

6. 允许使用开源代码或工具，但不允许使用任何未公开发布或需要授权的代码或工具；

7. 除主办方提供的数据集外，参赛选手允许使用预训练数据（如词向量、字向量等），但是不能用额外的标注数据。

8. 参赛队伍可在参赛期间随时上传验证集的预测结果，一天不能超过3次 ，管理系统会实时更新各队伍的最新排名情况。

# 奖励设置

第一名：人民币10000元

第二名：人民币3000元

第三名：人民币2000元

技术创新奖：人民币5000元

备注：每个子任务单独设立奖项，每个子任务最多有一个队伍获得技术创新奖，技术创新奖和前三名可以兼得，优胜者有机会进入2021年华为校园招聘的优招通道。

#  任务组织者和联系人

**任务组织者：**

谢远涛 （华为）

孙伟健 （华为）

陈玉博 （中科院自动化所）

赵军 （中科院自动化所）

**任务联系人：**

谢远涛：xieyuantao2@huawei.com

陈玉博：yubo.chen@nlpr.ia.ac.cn

# 任务交流平台

任务交流平台采用钉钉群的方式。所有报名参赛的参赛队要求有一名成员加入到该讨论群中，后续所有有关本任务评测及数据的相关说明和通知将只在讨论群中发布和交流，不再另行发邮件或更新官网通知，请所有参赛队务必在报名成功之后加入该讨论群。钉钉群加入二维码如下，在钉钉上扫一扫添加：



子任务1：事件抽取钉钉群二维码



子任务2：事件共指钉钉群二维码

加入讨论组时请务必注明所属参赛队队名及所属单位，加入讨论组有问题的参赛队请联系：**ccks2021taskcpe\_t1@126.com 或ccks2021taskcpe\_t2@126.com**。