

第10章 篇章分析

宗成庆

中国科学院自动化研究所

cqzong@nlpr.ia.ac.cn

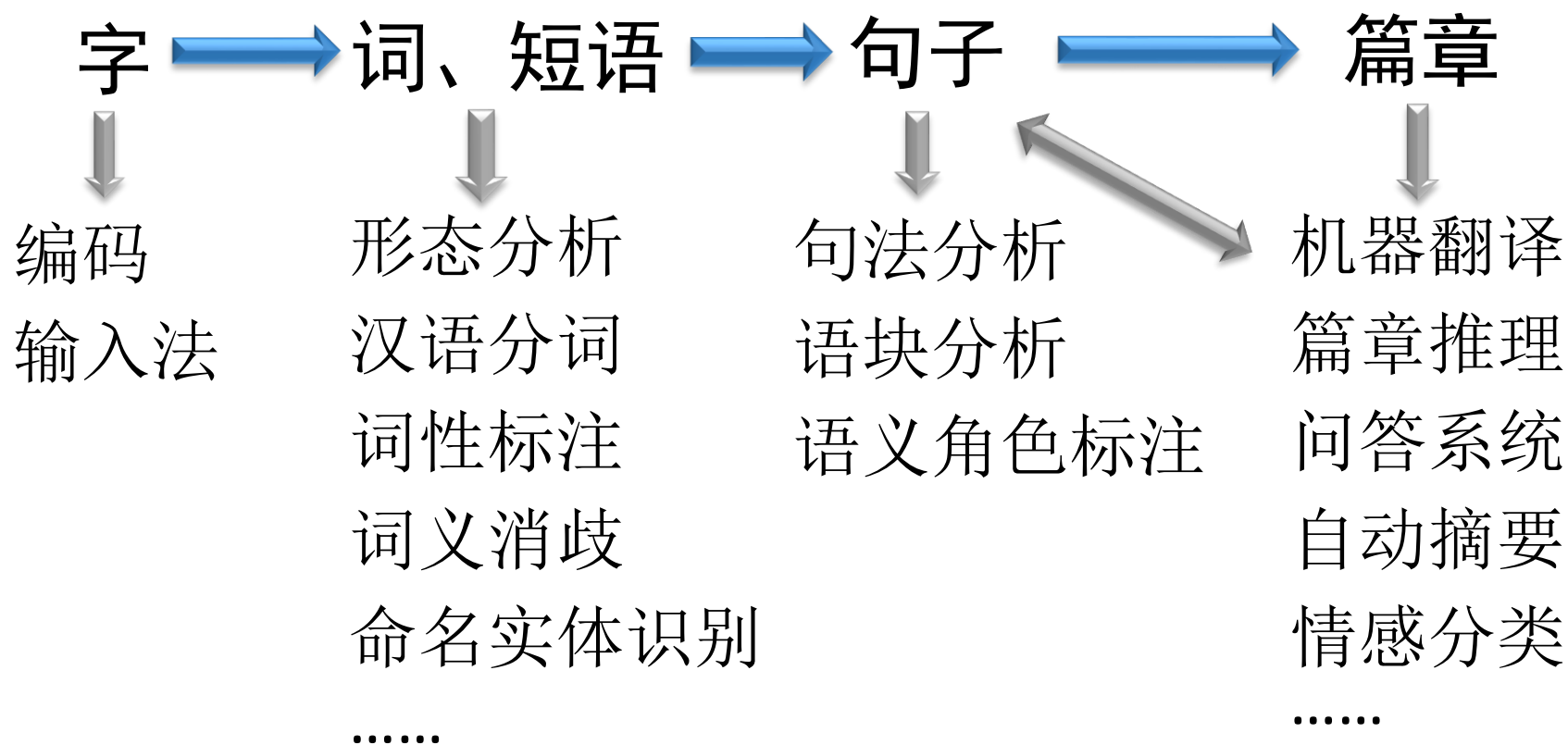
本章内容



1. 概述
2. 篇章表示理论
3. 篇章关系分析
4. 篇章关系应用
5. 习题

1. 概述

◆ 语言处理单位



1. 概述

◆以翻译为例

她穿着一件灰色的风衣，带着一副茶色的眼镜，左肩挎着一个白色的高档包，右手领着一一条白色的小狗，**据说是局长的女儿。**

据说是只名犬 / 据说是个模特

She was wearing a gray windbreaker, a pair of brown sunglasses, a white high-end bag on her left shoulder, and a white puppy in her right hand, **which is said to be the daughter of the director.**

She is wearing a gray windbreaker, a pair of brown sunglasses, a white high-end bag on her left shoulder, and a white puppy in her right hand, **which is said to be a famous dog.**

She is wearing a gray windbreaker, a pair of brown sunglasses, a white high-end bag on her left shoulder, and a white puppy in her right hand. **She is said to be a model.**

1. 概述

◆以对话系统为例

A



公司最近怎么样啊？

不怎么样



不是新款手机上市了吗？

市场不好



又在开发新产品？

不开发怎么办？



好好调研一下，别盲目上，要不然老白干

西北风都喝不上，还老白干呢



.....

B



海底捞两人多少钱？

咋的了，这两人？掉海里了？



我说海底捞吃饭多少钱？

吃饭怎么还得从海底捞？



算了，干脆铁板烧吧，两人多少钱？

铁板烧两人，那更不行了，那不是犯法嘛？



.....

本章内容

1. 概述

 2. 篇章表示理论

3. 篇章关系分析

4. 篇章关系应用

5. 习题

2. 篇章表示理论

◆ 词汇链

(S1)数年前，**北海**还是北部湾一个默默无闻的小渔村，然而三五年时间**北海**已建成了一个现代化都市的框架，街上客流如潮，楼房拔地而起。

(S2)**北海**已成为中国**对外开放**中升起的一颗明星。

(S3)**北海市**的崛起，是近年来**广西**壮族自治区**对外开放**取得卓著成就的重要标志之一。

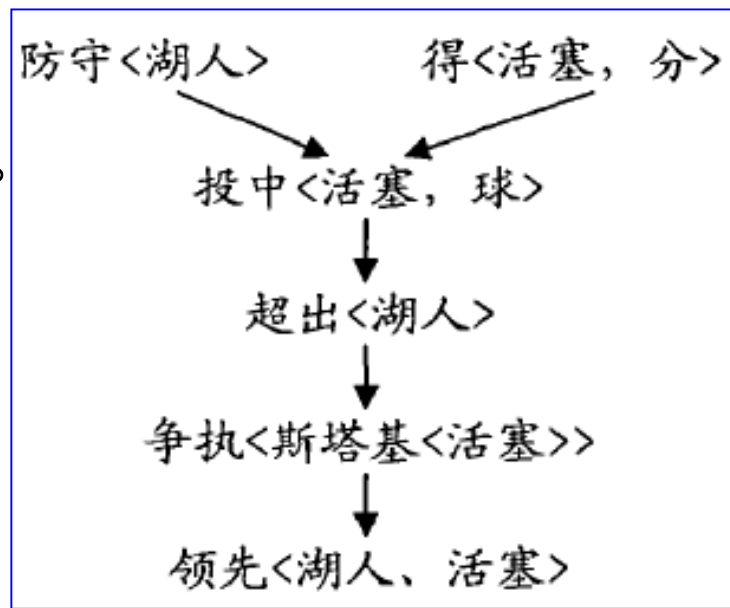
(S4)现在**广西**已初步形成了沿海开放城市、沿海经济开放区、**边境开放城镇**相结合，由沿海、沿边、沿江向腹地推进的多领域、多层次的**对外开放**总体格局。

(S5)统计资料显示，过去五年**广西****对外贸易**和利用**外资**规模迅速扩大，**进出口**贸易额累计达到一百亿美元，其中出口六十八点七亿美元，分别比“七五”时期（一九八六至一九九〇年）增长一点七八倍和一点四三倍；实际利用**外资**累计达到三十三点二四亿美元，……

2. 篇章表示理论

◆ 事件链

- a. 第三节湖人加强了防守，拼得活塞只得9分。
- b. 本节打了8分钟后，活塞只投中一球。
- c. 湖人以16:3开始本节，一举以57:48超出。
- d. 其间活塞斯塔基和裁判还发生了争执。
- e. 三节过后，湖人以 61:54 领先活塞。



● 构建过程:

- 抽取实体词汇链
- 抽取词汇链上每个词最近的谓词-论元，构成事件链
- 判断相邻事件之间关系

2. 篇章表示理论

◆ 话题链

一组以名词回指(noun anaphora, NA)、代词回指(pronoun anaphora, PA)和零形回指(zero anaphora, ZA)形式的话题连接起来的小句或句子

- **回指(anaphora)**: 是指一个词或短语在语篇中用于(回)指代同一语篇中的另一个词或短语的概念(Quirk et. Al, 1985)。

例1:

中国科学院大学简称国科大, 其前身是中科院研究生院, 拥有浓郁的学术氛围和雄厚的科研实力, 国科大招收的第一届本科生于2018年夏季毕业。

Diagram illustrating anaphora types in the example sentence:

- PA (pronoun anaphora)**: Indicated by a red arrow from "其" (its) to "中国科学院大学" (Chinese Academy of Sciences University).
- ZA (zero anaphora)**: Indicated by a red arrow from "国科大" (Guo Ke Da) to "中国科学院大学" (Chinese Academy of Sciences University).
- NA (noun anaphora)**: Indicated by a red arrow from "国科大" (Guo Ke Da) to "中国科学院大学" (Chinese Academy of Sciences University).

2. 篇章表示理论

例2：你们年纪还小，(S1)还要成家立业，(S2)不要虚度年华，
(S3)更不要成为社会讨厌的人。

话题链：你们—ZA—ZA—ZA

关联词：“还…更…”、“不要…不要…”

例3：我无意中碰到了身边的一个什么东西，(S1)伸手一摸(O1)，
(S2)是他给我开的饭，两个干硬的馒头。

主话题链：我—ZA； 次话题链：Φ—ZA

实体链：东西—Φ—饭—馒头

2. 篇章表示理论

◆ 修辞结构理论(Rhetorical Structure Theory, RST)

是美国学者Willian C. Mann 和 Sandra A. Thompson 1983年提出来的，代表作有：

- ① Rhetorical Structure Theory: A Theory of Text Organization, 1987
- ② Rhetorical Structure Theory: Toward a functional theory of text organization, 1988

2. 篇章表示理论

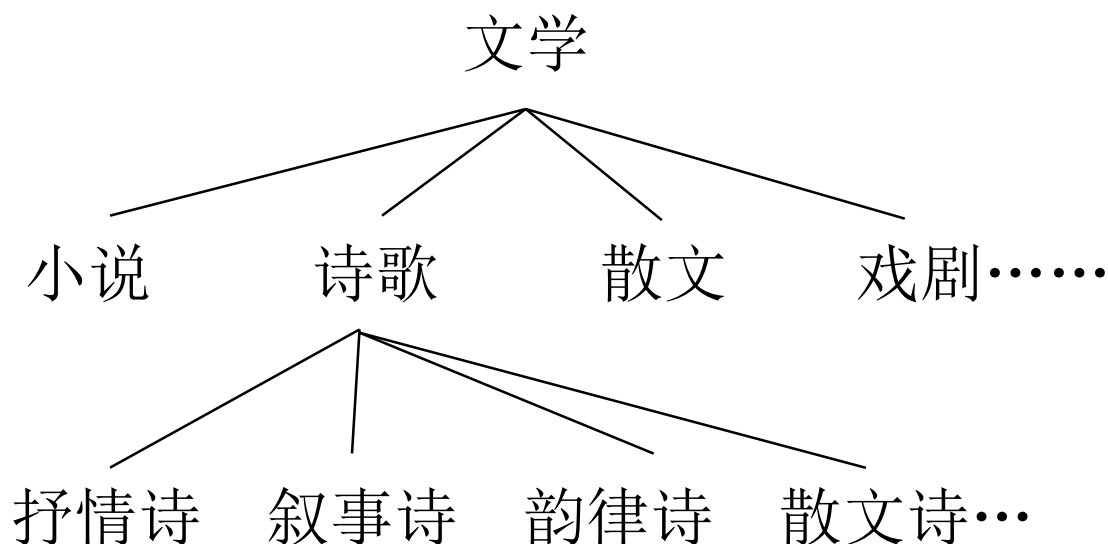
● RST理论的主要内容

- 语篇由各个具有重要功能的部分构成，较小的部分按照一定的关系模式组成更大的部分，直至生成语篇；
- 一段话语或文字要被确认为语篇，各个部分必须有机地结合，形成整体性和连贯性；
- 整体性和连贯性来自语篇的内在功能，一个语篇之所以产生整体性和连贯性的效果，是因为它的每一个部分都直接或间接地服务于语篇作者的同一个中心目的；
- 语篇的构成方式如下：两个基本部分组成一个较大的部分，较大部分再与另一部分组织更大的部分，直至语篇，因此语篇是层次化的结构；

2. 篇章表示理论

- RST区分三种语篇结构：类型结构、句法结构和关系结构。其中，类型结构由语篇的题材或类型特征决定，相对固定。语篇类型包括狭义的语篇类型和广义的语篇类型。狭义的语篇指人们处于特定的交际目的而选择的语篇体现形式，如小说、诗歌、散文、戏剧、书信、广告、标签、票据、留言、遗嘱、考卷、支票、菜单、日记等。

层次化结构：



2. 篇章表示理论

- 关系结构内部具有同一性：它不因结构层级的改变而不同；关系结构是多语句语篇的主要结构；
- 不对称性在关系结构中占主导地位。英语中最常用的一类结构关系是“核心-辅助关系”(Nucleus-Satellite)；
- 语篇的结构关系是功能关系，不是形式关系；语篇关系是语篇的深层结构关系；语篇关系的种类和数量原则上是无限的，以前未出现的关系类型也可能在新的语篇关系中出现。

2. 篇章表示理论

● 英语语篇常见的结构关系

- 环境关系(Circumstance)
- 目的关系(Purpose)
- 解答关系(Solution hood)
- 对照关系(Antithesis)
- 阐述关系(Elaboration)
- 让步关系(Concession)
- 背景关系(Background)
- 条件关系(Condition)
- 使能关系(Enablement)
- 析取关系(Otherwise)
- 动机关系(Motivation)
- 解释关系(Interpretation)
- 证据关系(Evidence)
-

2. 篇章表示理论

例句：

泽巴里说，以美国为首的联军是应伊拉克要求在伊存在并对伊提供保护的，因为目前的伊临时政府不能控制安全局势，无力追剿恐怖分子和破坏分子。

① 泽巴里说，

② 以美国为首的联军是应伊拉克要求在伊存在并
提供保护的，

③ 因为目前的伊临时政府不能控制安全局势，

④ 无力追剿恐怖分子和破坏分子。

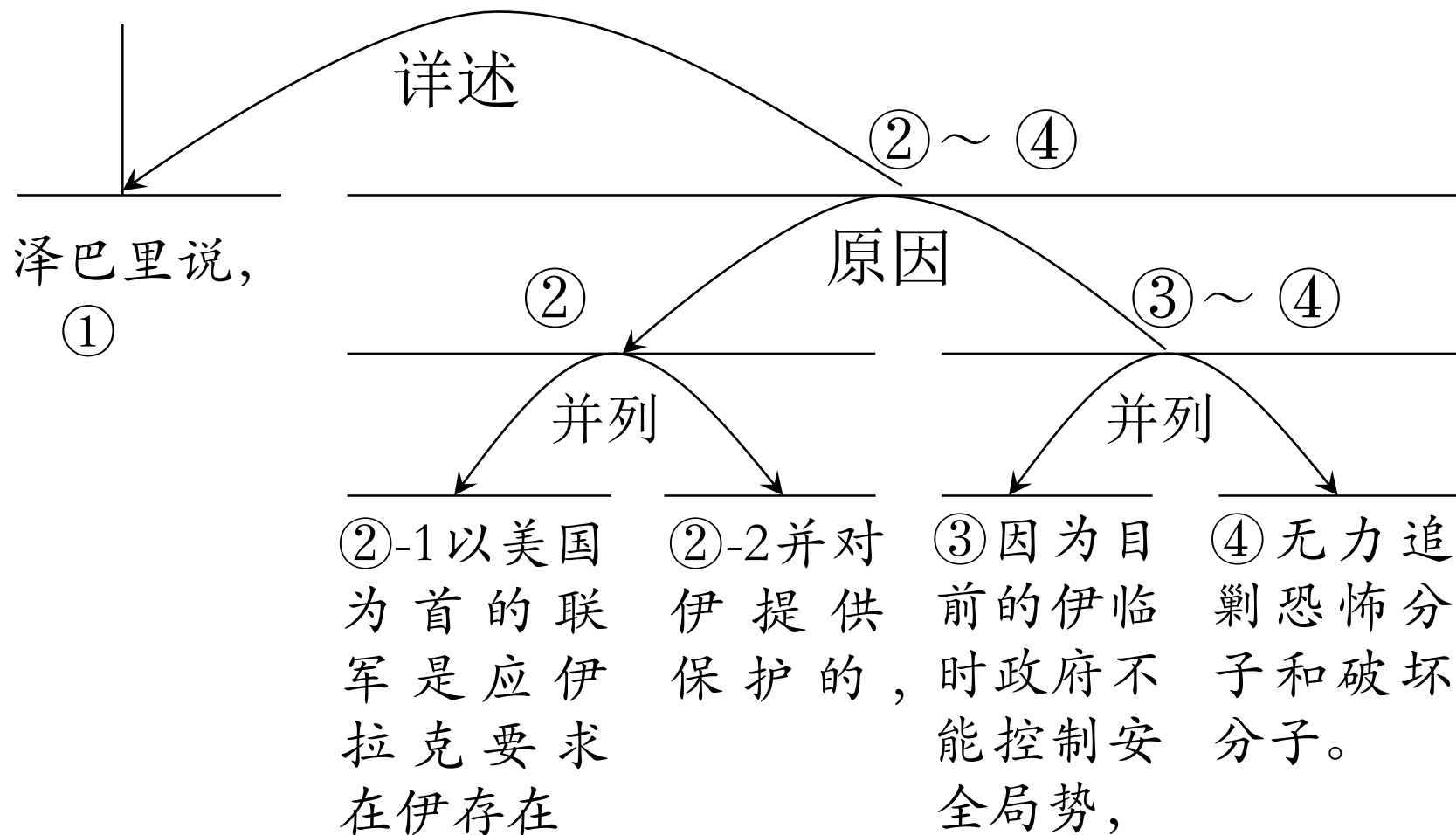
②-1 | ②-2

← | →

伊临时政府

并对伊

2. 篇章表示理论



2. 篇章表示理论

● 汉英篇章树库对比 (CDTB vs. PDTB)

| | 汉语 (CDTB) | 英语 (PDTB) |
|------|--|---|
| 隐式关系 | 78.27% | 54.53% |
| 连接词 | flexible (385 in corpus) many parallel connectives (“因为”、“所以”、 “因为...所以...” ...) | A fixed vocabulary table with 100 connectives |
| 论元定义 | by semantics (The span of an argument ranges from several words to sentences even to paragraphs) | by physical location of the connective (The span must belong to the pair of adjacent sentences) |
| 语义类型 | 11 types, flat structure | 4classes, 16types, hierarchical structure |
| 标点符号 | indicate discourse relations | no discourse function |

2. 篇章表示理论

● 汉语篇章举例

中国 建筑业 对 外 开放 呈现 新 格局

新华社 北京 二月 十三日 电

中国 建筑 市场 近年 来 对 外 开放 步伐 进一步 加快。

据 初步 统计，目前 在 中国 境内 承包 工程 的 国外 承包商 已 有 一百三十七 家，承包 的 工程 达 一百四十一 项，其中 最大 规模 的 项目 达 二十七点七亿 元；中外 合资 合作 的 建筑 企业 近 二千 家。

中国 建筑业 对 外 开放 始于 八十年代。

十 几 年 来，已 有 美国、日本、法国、英国、德国、芬兰、意大利、新加坡、香港、台湾等十几个国家和地区 的 境外 企业 进入 中国 进行 工程 总 承包 或 工程 分包。

世界 上 最大 的 二百二十五 家 国际 承包商 中，有 十几 家 已 进入 中国，其中 不少 公司 与 中国 公司 合 资 合 作 进行 建设。

根据 建设部 的 规定，凡 属于 国际 金融 组织 贷款 并 由 国际 公开 招标 的 工程 全部 由 外国 投资 或 赠款 建设 的 工程，以及 国内 企业 在 技术 上 难以 单独 承包 的 中外 合 资 建设 工程，境外 建筑 企业 在 取得 中国 审批 的 外国 企业 承包 工程 资质 证 后，皆 可 进入 中国 境内 承包 建设 项目。

一九九五年 九月 建设部 和 外经贸部 联合 发布 的 《关于 设立 外商 投资 建筑业 企业 的 若干 规定》，使 中国 的 建筑 市场 从 允许 境外 企业 到 中国 承包 工程 进入 到 允许 境外 企业 到 中国 办 合 资 建筑 企业。

(完)

1

Type: Implicit

Arg1: 中国 建筑 市场 近年 来 对 外 开放 步伐 进一步 加快

Arg2: 目前 在 中国 境内 承包 工程 的 国外 承包商 已 有 一百三十七 家，承包 的 工程 达 一百四十一 项，其中 最大 规模 的 项目 达 二十七点七亿 元；中外 合 资 合 作 的 建筑 企业 近 二千 家

Sense: Expansion

2

Type: Implicit

Arg1: 目前 在 中国 境内 承包 工程 的 国外 承包商 已 有 一百三十七 家

Arg2: 承包 的 工程 达 一百四十一 项

Sense: Conjunction

3

Type: Explicit

Connective: 其中

Arg1: 承包 的 工程 达 一百四十一 项

Arg2: 最大 规模 的 项目 达 二十七点七亿 元

Sense: Expansion

4

Type: Implicit

Arg1: 目前 在 中国 境内 承包 工程 的 国外 承包商 已 有 一百三十七 家，承包 的 工程 达 一百四十一 项，其中 最大 规模 的 项目 达 二十七点七亿 元

Arg2: 中外 合 资 合 作 的 建筑 企业 近 二千 家

Sense: Conjunction

2. 篇章表示理论

◆ 汉语广义话题结构理论 (Generalized Topic Structure)

由北京语言大学宋柔教授提出。宋柔教授于1968年本科毕业于北京大学数学力学系，1981年研究生毕业于北京大学计算机系，取得硕士学位。退休前为北京语言大学信息科学学院教授，博士生导师，主要从事智能软件工具、语言信息处理、人工智能应用方面的研究、开发和教学工作，在汉语篇章结构表示和分析方向做出重要贡献。



1946年9月出生于苏州

2. 篇章表示理论

● 小句复合体的语法结构：

- 关注书面语篇章(text)
- 采用如下层次(由低到高):

词/短语 → 小句 → 小句复合体 → 篇章

- 在语义上，小句表现为关于事物（事物与属性的关系，事物与状态的关系，事物与事物的关系，事物与行为的关系等）的命题，小句复合体是表现关系密切的若干命题的组合。
- 小句又常被称为分句(clause)，多个小句构成汉语的句子，称为复句。
- 写者/说者叙述一个事物时选定的各层次的话语开端称为话头。
- 标点句当作小句的主体、小句复合体的构件；标点句之间可成分共享；共享成分与接受它的标点句之间的关系为话头话身关系。

2. 篇章表示理论

● 换行缩进图式表示

北京冬季盛行西北风，经常出现大风、降温、寒冷、干燥天气；春季气温回升迅速，云量稀少，多大风；夏季当东南季风边缘摆动到北京附近时，南来的暖湿空气与北京冷空气相遇，形成7-8月高温多雨天气，对农业生产有利；秋季天高气爽，舒适宜人。 《中国大百科全书. 中国地理卷》

11个标点句：

①北京冬季盛行西北风，②经常出现大风、降温、寒冷、干燥天气；③春季气温回升迅速，④春云量稀少，⑤多大风；⑥夏季当东南季风边缘摆动到北京附近时，⑦南来的暖湿空气与北京冷空气相遇，⑧形成7-8月高温多雨天气，⑨对农业生产有利；⑩秋季天高气爽，(11)舒适宜人。

2. 篇章表示理论

北京冬季盛行西北风,

经常出现大风、降温、寒冷、干燥天气;

春季气温回升迅速,

云量稀少,

多大风;

夏季当东南季风边缘摆动到北京附近时,

南来的暖湿空气与北京冷空气相遇,

形成7-8月高温多雨天气,

对农业生产有利;

秋季天高气爽,

舒适宜人。


2. 篇章表示理论

◆ 其它语篇理论

- 中心理论 (Centering Theory)
- 脉络理论 (Veins Theory)
- 言语行为理论 (Speech-Act Theory)

.....

本章内容

1. 概述
2. 篇章表示理论
-  3. 篇章关系分析
4. 篇章关系应用
5. 习题

3. 篇章关系分析

◆ 关于 CoNLL

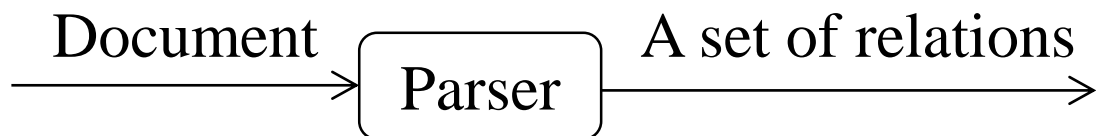
The SIGNLL (ACL's Special Interest Group on Natural Language Learning) *Conference on Computational Natural Language Learning*

- Shared Task (2016): Multilingual Shallow Discourse Parsing (<http://www.cs.brandeis.edu/~clp/conll16st/>)

属于浅层分析。

3. 篇章关系分析

◆ 基本任务

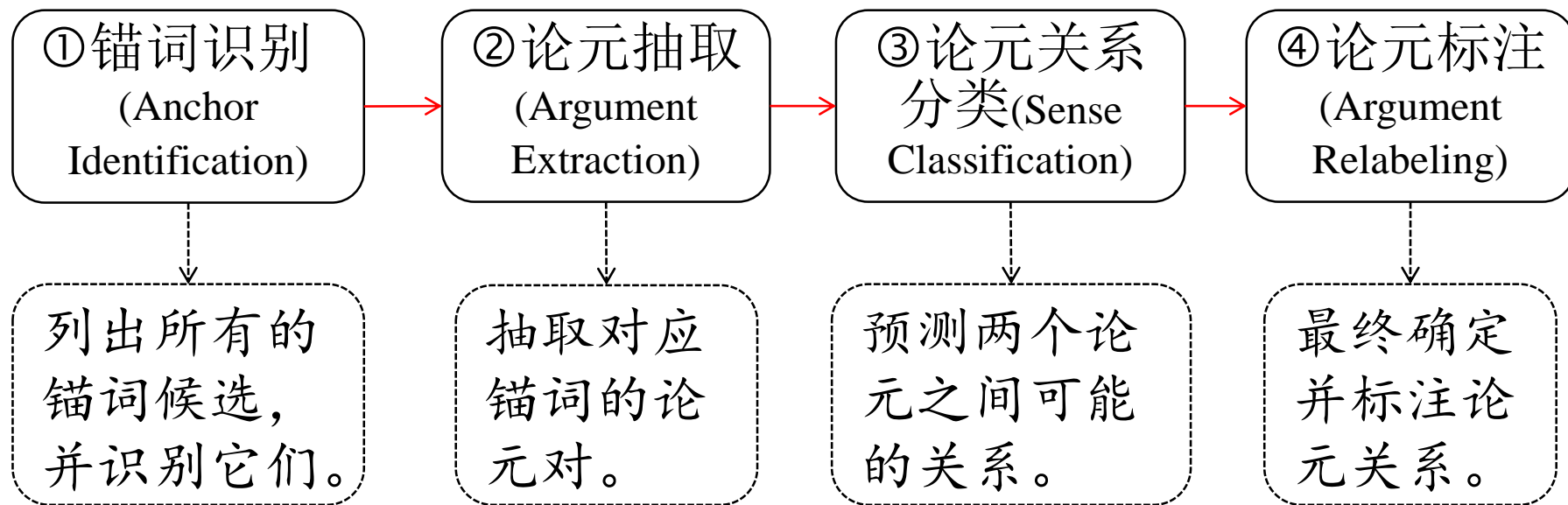


- 显式(有明确的关联词)或隐式关系
- 两个篇章论元 (Arg1 & Arg2) 及其它它们之间的关系
- Discourse Parser 的三大任务：
 - 关联词识别（两个论元之间的分界点）
 - Arguments 抽取（边界范围）
 - Arg1与Arg2之间的篇章功能类型判断

*有时将argument称作基本篇章单元(elementary discourse unity, edu)。

3. 篇章关系分析

◆ 基本的分析流程



◆ 实现方法

- 选取特征
- 构造分类器

3. 篇章关系分析

◆ 举例说明

① 锚词识别

- 显式关系：关联词表
- 非显式关系：标点符号
 - 句中(middle of sentence, MOS)
 - 句末(end of sentence, EOS)

将所有分号、逗号、冒号、句号、破折号、省略号、问号和感叹号作为候选。

尽管她的动作潇洒自如，但难度无法与罗莉相比，只获得9.875分，夺得银牌。

中国建筑市场近年来对外开放步伐进一步加快。据初步统计，目前在中国境内……

3. 篇章关系分析

- 用于锚点词识别的特征
 - 候选词本身
 - 候选词的词性
 - 候选词前一个词的词性
 - 在上下文子句的第一个词和最后一个词
 - 在上下文子句的前三个词和最后三个词
 - 在上下文子句的后三个词的词嵌入表示
 - 在上下文子句的前三个词和最后三个词的词嵌入表示
 - 前一个位置是否为标点
 - 后一个位置是否为标点
 -

3. 篇章关系分析

②论元对抽取

经过分析发现：

- 在大多数情况下, Arg1 和 Arg2 在同一个句子中, 或者分别在两个邻近的句子中;
- 一个论元由一个或几个连续的子句组成;
- 显式的 Arg2 与它的关联锚词在同一个句子中;
- 在大多数情况下, Arg1 和 Arg2 的跨度范围是邻近的, 它们之间没有其它子句。

两种情况：

- Arg1与锚词在同一子句内, 记作: SS;
- Arg1在锚词的前一个子句中, 有锚词的子句为Arg2。这种情况记作: PS。

3. 篇章关系分析

例如:

- 在显式的情况下，由 Arg1 的位置决定

【今年7月，在里海划分问题上俄以退为进，对其原先坚持的“只分边缘、中间共管”的立场作了重大调整，接受哈萨克斯坦提出的“按中心线划分海底、水域共享”的原则，并不顾伊朗、土库曼斯坦等其它里海沿岸国家的反对，同哈就划分里海北部地区达成协议。】 (SS) Arg1

- 在隐式的情况下，由锚点位置决定(句中 或 句尾)

【目前在中国境内承包工程的国外承包商已有一百三十七家，承包的工程达一百四十一项，.....】 Arg1
Arg2

3. 篇章关系分析

• 论元边界双向扩展



【今年7月，| 在里海划分问题上俄以退为进，| 对其原先坚持的“只分边缘、中间共管”的立场作了重大调整，| 接受哈萨克斯坦提出的“按中心线划分海底、水域共享”的原则，| C4并C6不顾伊朗、土库曼斯坦等其它里海沿岸国家的反对，| 同哈C5就划分里海北部地区达成协议。】

Clause1, Clause2, Clause3, Clause4, || Clause5, Clause6
← 并 →

{C4||C5}

{C4||C5,C6}

{C3,C4||C5}

{C3,C4||C5,C6}

{C2,C3,C4||C5}

{C2,C3,C4||C5,C6}

{C1,C2,C3,C4||C5}

{C1,C2,C3,C4||C5,C6}

3. 篇章关系分析

- 对种子论元对范围扩展，借助分类器选择最长边界的论元对
 - 方向：向左或者向右
 - 一个子句一个子句地考察
 - 利用词法、句法、位置等特征，借助分类器识别
 - 选取最长的边界

3. 篇章关系分析

③关系类型识别：分类

➤ 词法特征

- 连接词本身
- 连接词的词嵌入表示
- 连接词的前一个和后一个位置的标点符号

.....

➤ 句法特征

- Arg1 和 Arg2 节点本身
- 连接词节点的父亲节点
- 连接词节点的左右兄弟节点

.....

➤ 分类器

- SVM
- Bayes
- CRFs
- ME

.....

3. 篇章关系分析

④论元标识：分类

这一步的目的是进一步确定Arg1和Arg2孰前孰后。Arg1和Arg2的前后关系是根据语义决定的,而不是位置。

他生病了, 所以今天没来上班。

Arg1(原因) Arg2(结果)

他今天没来上班, 因为生病了。

Arg2(结果) Arg1(原因)

3. 篇章关系分析

◆ 分析性能

● CoNLL'2016 评测数据集

| 数据集 | #篇章数 | #句子数 | #关系个数 |
|-----------|------|-------|--------|
| 训练集(CDTB) | 455 | 6,332 | 10,240 |
| 开发集 | 24 | 349 | 383 |
| 测试集 | 30 | 348 | 455 |
| Blind(盲测) | 64 | 1,140 | 2,101 |

普通测试：测试集提前发个参赛队伍。

盲测：比赛前没有提供给参赛队任何测试信息。

3. 篇章关系分析


● 性能表现

- 在盲测集上分析性能 (F1值) 大幅度下降;
- Arg1和Arg2识别分别正确的F1值比同时两个论元都正确的F1值高 **10%** 左右。

X. Kang *et al.* An End-to-End Chinese Discourse Parser with Adaptation to Explicit and Non-explicit Relation Recognition. *Proc CoNLL'2016: Shared Task*. pp. 27-32

| | Task | Dev | Test | Blind |
|--------------------|---------------|--------|--------|--------|
| Explicit 显式 | <i>Conn</i> | 0.8356 | 0.7263 | 0.5627 |
| | <i>Arg1</i> | 0.5479 | 0.5587 | 0.3853 |
| | <i>Arg2</i> | 0.6849 | 0.6816 | 0.4444 |
| | <i>Both</i> | 0.4521 | 0.4916 | 0.2650 |
| | <i>Sense</i> | 0.7534 | 0.6480 | 0.4811 |
| Non-Explicit 隐式 | <i>Parser</i> | 0.4521 | 0.4859 | 0.2446 |
| | <i>Conn</i> | — | — | — |
| | <i>Arg1</i> | 0.6282 | 0.6266 | 0.5526 |
| | <i>Arg2</i> | 0.6798 | 0.6762 | 0.6017 |
| | <i>Both</i> | 0.5341 | 0.5379 | 0.4457 |
| | <i>Sense</i> | 0.5068 | 0.4987 | 0.4082 |
| All | <i>Parser</i> | 0.3982 | 0.3869 | 0.2712 |
| | <i>Conn</i> | 0.8356 | 0.7263 | 0.5627 |
| | <i>Arg1</i> | 0.6261 | 0.6328 | 0.5439 |
| | <i>Arg2</i> | 0.6932 | 0.6921 | 0.5843 |
| | <i>Both</i> | 0.5317 | 0.5418 | 0.4178 |
| | <i>Sense</i> | 0.5640 | 0.5333 | 0.4326 |
| | <i>Parser</i> | 0.4120 | 0.4089 | 0.2690 |

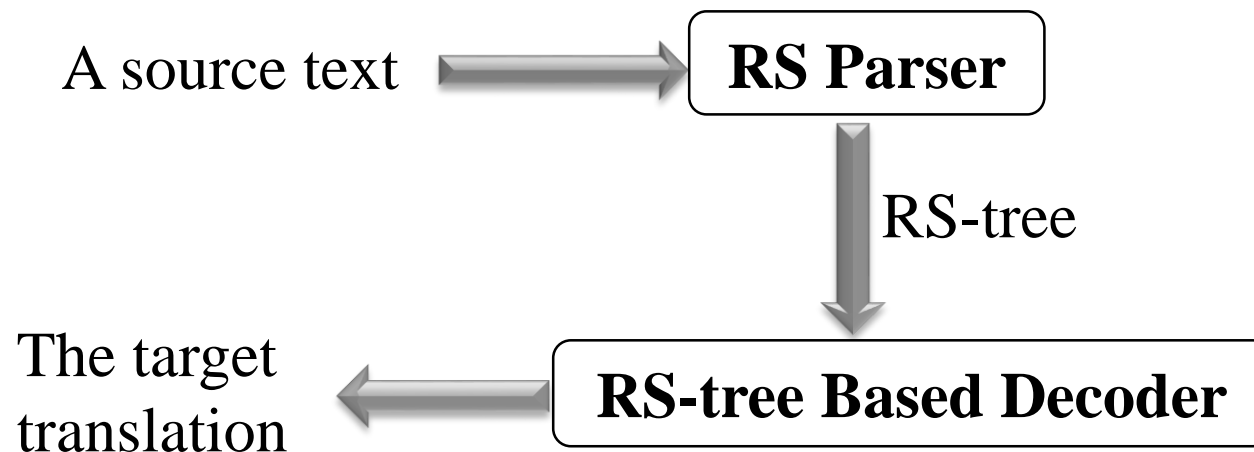
本章内容

1. 概述
2. 篇章表示理论
3. 篇章关系分析
-  4. 篇章关系应用
5. 习题

4. 篇章关系应用

◆ 以机器翻译为例

(1) RS-tree-to-String 统计翻译模型



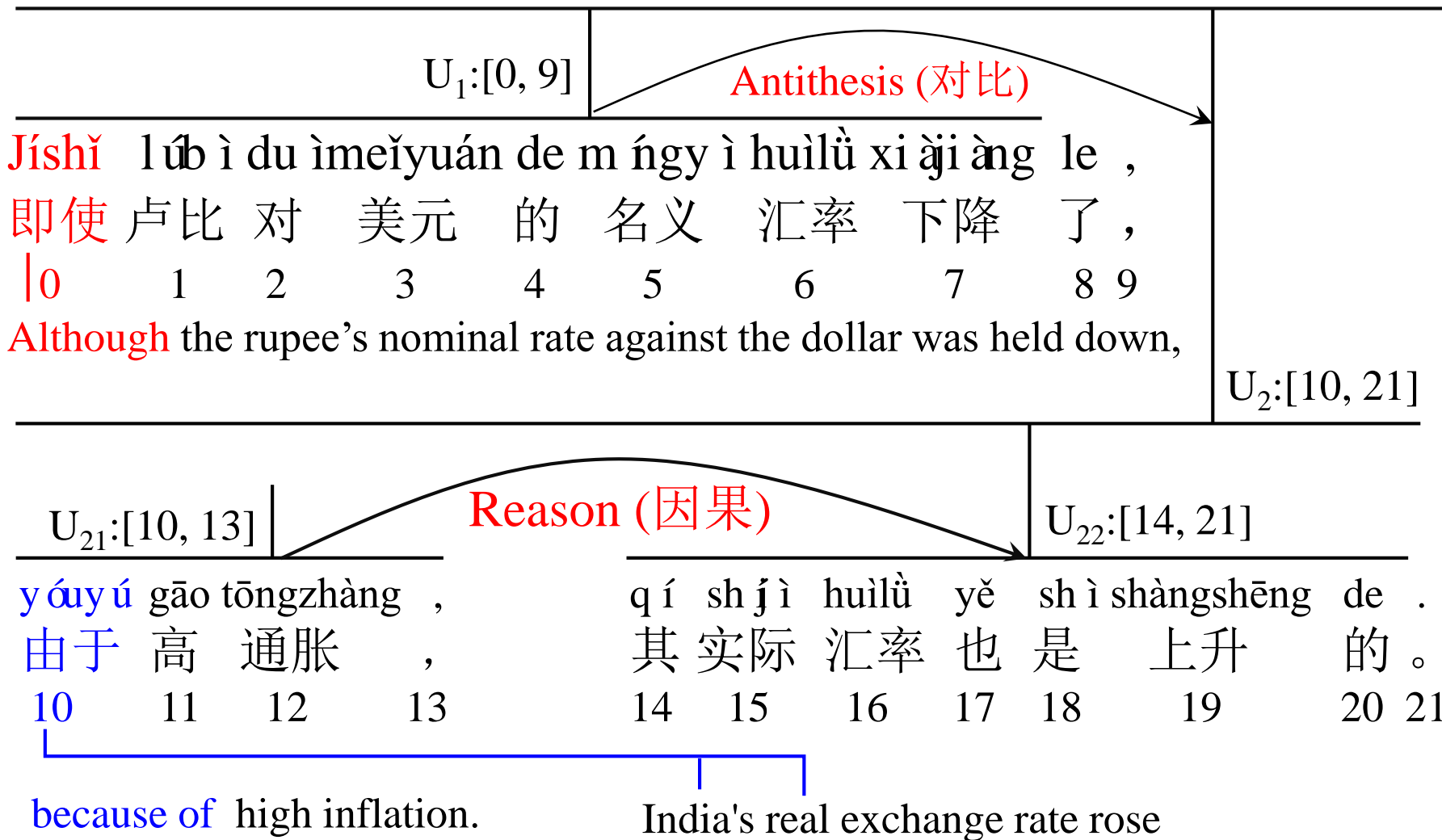
4. 篇章关系应用

● 举例说明

即使卢比对美元的名义汇率下降了，由于高通胀，其实际汇率也是上升的。

Although the rupee's nominal rate against the dollar was held down, India's real exchange rate rose because of high inflation.

4. 篇章关系应用



4. 篇章关系应用

- 定义翻译模板

$$relation :: U_1(\alpha, X) / U_2(\gamma, Y) \\ \Rightarrow U_1(tr(\alpha), tr(X)) \sim U_2(tr(\gamma), tr(Y))$$

$tr(U_{21}) = \text{because of high inflation.}$

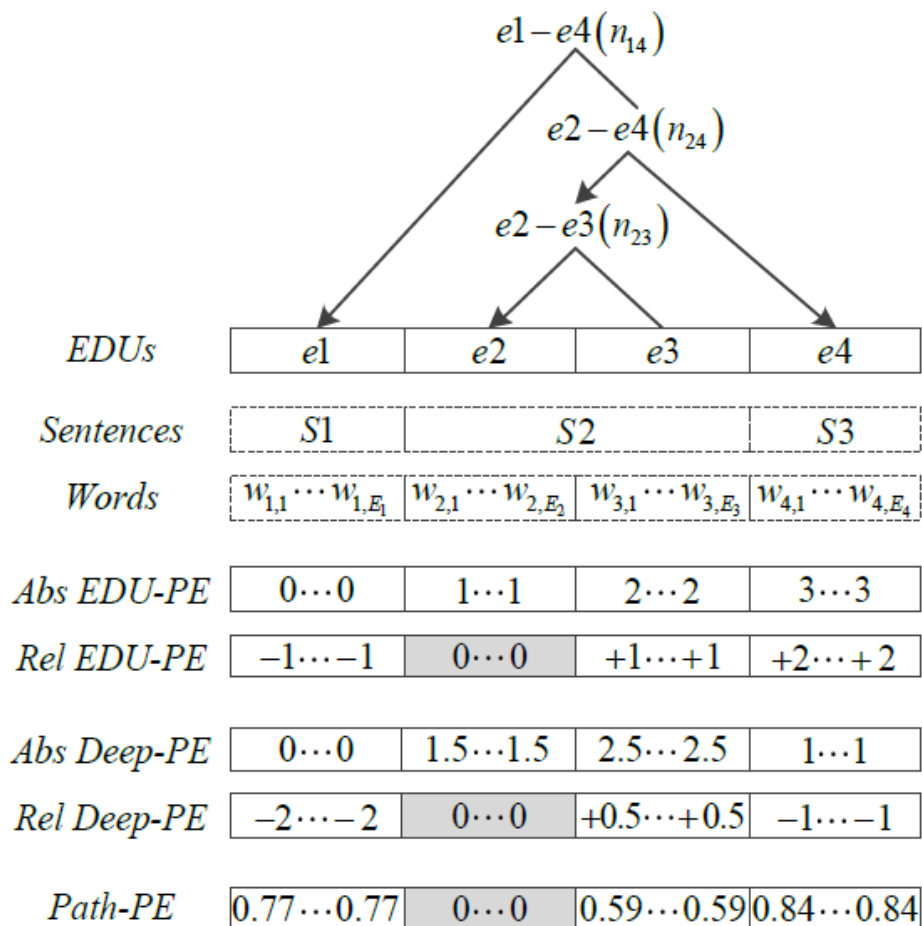
$tr(U_{22}) = \text{India's real exchange rate rose}$

Reason:: 由于 $[X]$ / $[Y] \Rightarrow [Y]$ / because of $[X]$

M. Tu *et al.* Enhancing Grammatical Cohesion: Generating Transitional Expressions for SMT. *Proc. ACL'2014*, pages 850-860

4. 篇章关系应用

(2)融合篇章结构信息的神经机器翻译



将EDU的位置和关系等信息融入Transformer-based 的端到端的神经机器翻译模型，包括：

- 绝对位置编码(Abs EDU-PE)
- 相对位置编码(Rel EDU-PE)
- 绝对深度位置编码(Abs Deep-PE)
- 相对深度位置编码(Rel Deep-PE)
- 路径位置编码(Path-PE)

亢晓勉等，融合篇章结构位置编码的神经机器翻译，智能科学与技术学报，2020, 2(2): 144-152

4. 篇章关系应用

(3) 基于大范围语境的神经机器翻译方法

● 针对的问题:

- 句子的语义往往不完整，需要周围句子甚至更大范围的语境知识帮助改善当前句子的理解，提升译文质量
- 已有的方法一方面不加区分地利用篇章范围中的句子，另一方面非常武断地选择篇章上下文的范围，例如整篇文章

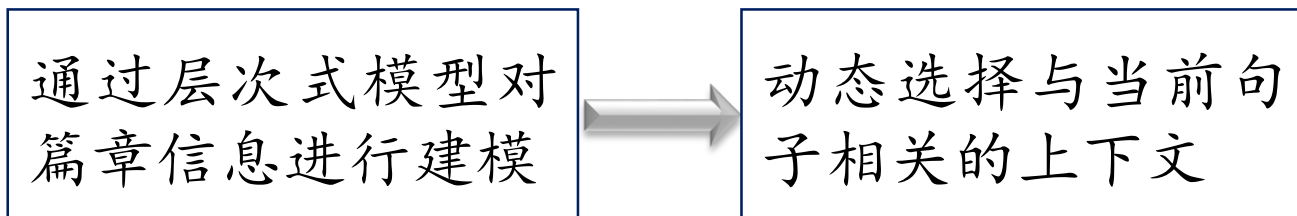
上下文: …… 赛车能带给我热情和动力，一段时间由于缺乏竞争对手，我有些懈怠了，其实我仍然非常热衷于这项运动。

Reference: ... I am still very fond of the sport.

NMT: ... I am still very enthusiastic about this movement.

4. 篇章关系应用

● 解决思路



- 不同的段落拥有不同的权重
- 不同的句子考虑不同长度范围的上下文
- 不同的上下文句子拥有不同的权重

● 实验结论

单纯地增加篇章上下文语境范围对译文的改进效果不明显，动态选择上下文语境的策略可以显著提升翻译质量。

IJCAI'2020
EMNLP'2020

本章小结

◆篇章分析的任务

◆篇章表示理论

- 词汇链、事件链、话题链、修辞结构理论(RST)

◆篇章关系分析方法

- 以修辞结构理论为例，主要采用分类的思路

◆篇章分析技术应用

- 以机器翻译为例

本章内容

1. 概述
2. 篇章表示理论
3. 篇章关系分析
4. 篇章关系应用

 5. 习题

5. 习题

1. 请下载 CoNLL'2016 Share Tasks评测的语料，结合该评测任务的说明，实现一个篇章关系浅层分析器。
2. 请设计实现一种端到端的篇章关系分析方法。
3. 请阅读修辞结构理论的相关文献。
4. 请阅读 UPenn LDC 篇章语料库的标注说明。
5. 请查阅学习关于汉语篇章表示的“广义话题结构理论”相关文献。
6. 设计实现一种针对客服对话的因果关系分析方法。

谢谢!

Thanks!

