

# 阅读理解中因果关系类选项的研究\*

王元龙<sup>1</sup>, 李茹<sup>1,2</sup>, 张虎<sup>1</sup> 王智强<sup>1</sup>

(1.山西大学 计算机与信息技术学院, 山西 太原 030006;

2. 山西大学 计算智能与中文信息处理教育部重点实验室, 山西 太原 030006)

**摘要:** 针对阅读理解选择题中因果关系类选项, 提出了基于因果关系网的因果关系支持度分析方法。首先通过线索短语从阅读材料中抽取因果事件对, 并计算事件对之间因果关联强度, 综合利用抽取到的因果事件对与其对应的因果关联强度构成因果关系网; 然后综合考虑了选项中的每个词在文档中的重要性和整个文档中的区分能力, 采用 TF-IDF 方法分别从原文中检索与选项中因事件和果事件相关的句子; 最后基于因果关系网和抽取到的相关句计算选项的因果关系支持度。为了验证该方法, 实验采用了 769 篇模拟材料和 13 篇北京高考语文试卷材料 (包括原文与选择题) 作为测试数据集, 实验结果表明本文方法的准确率相比于 Baseline 方法提高了近 11 个百分点。

**关键词:** 自然语言处理; 因果关系网; 阅读理解; 语义相似度计算

中图分类号: TP391

文献标识码: A

## Causality and Its Applications in Reading comprehension

Wang Yuanlong<sup>1</sup>, Li Ru<sup>1,2</sup>, Zhang Hu<sup>1</sup>, Wang Zhiqiang<sup>1</sup>

(1.School of Computer and Information Technology, ShanXi University, Taiyuan, ShanXi 030006, China ;

2. Key Laboratory of Computation Intelligence and Chinese Information Processing of Ministry of Education, ShanXi University, Taiyuan, ShanXi 030006, China)

**Abstract:** For the choice of causal relation types in Chinese reading comprehension, this paper proposes a causal relation support analysis method based on causal network. Firstly, the causal events are extracted from literature by clue phrases, and the causal relation between the events is computed, and the causal network is constructed from the causal events and the causal relation. Then, the TF-IDF method is used to retrieve the related sentences from the document, and the importance of each word in the document and distinguish between the whole document are considered in the method. Finally, the causality network and related sentences are combined to analyze the causal support of the option. In order to verify the method, the 769 simulation materials +13 Beijing entrance examination materials (including the source text and the selected title) is collected as a test set. In this test, this method can obtain better result than the Baseline method.

**Key words:** natural language processing; causality network; reading comprehension; semantic similarity

---

\* 收稿日期: 定稿日期:

**基金项目:** 国家 863 计划项目 (2015AA015407); 国家自然科学基金 (61373082); 山西省自然科学基金项目 (201601D102030)

**作者简介:** 王元龙 (1983—), 男, 博士, 讲师, 主要研究方向为自然语言理解、机器学习; 李茹 (1963—), 女, 博士, 教授, 博士生导师, 主要研究方向中文信息处理、信息检索; 张虎 (1979—), 男, 博士, 讲师, 主要研究方向计算语言学。王智强 (1987—), 男, 博士研究生, 主要研究方向为自然语言处理。

## 1 引言

阅读理解 (Reading Comprehension, RC) 的任务是能够读懂单篇文档, 理解文中的语义且做适当的处理用于回答设计的问题, 它是自然语言处理和知识表示的一个重要应用领域, 驱动着该领域相关技术的发展<sup>[1-2]</sup>。近年来, 随着人工智能的浪潮, 阅读理解技术也越来越受到了众多自然语言处理研究机构及人工智能研究机构的关注<sup>[1]</sup>。阅读理解任务涉及到词法、句法、语法、语义和语用等多方面的信息, 需要综合运用文本的表示、分析、理解、推理等自然语言处理与理解技术, 相当具有挑战性<sup>[3]</sup>。

目前, 阅读理解任务有两种类型: 一种是 Multiple-choice 类型<sup>[4-6]</sup>, 该类型阅读理解旨在从原文中的多个句子推测出选项的正确与否; 另一种为 Cloze-style 类型<sup>[7-11]</sup>, 该类型阅读理解的答案是原文中的一个词, 需要理解原文的语义回答问题中所缺的答案。本文关注高考语文中的文学作品阅读理解任务, 重点解决选择题类型的问题。该任务与以上两种类型阅读理解任务不同, 选项是对原文的高度概括, 答案无法从原文中直接找到。因此需要先对选项进行深度理解。

郭少茹等<sup>[12]</sup>针对高考语文中科技文题型进行了分析, 将选择题求解归结为语义一致性分析任务, 注重的是背景材料与选项的语义一致性分析, 提出了句子语义相关度研究方法。由于文学作品阅

读理解中的选项通常比科技文阅读理解中的选项更加复杂, 因此简单分析语义相关度对文学作品阅读理解任务效果不明显。针对不同类型的选项分别给出不同的解决方案是文学作品阅读理解任务有效的解决思路。本文通过对近 13 年北京高考试题以及 769 篇模拟题选项进行分析 (每篇阅读理解包含一道选择类型题), 把选项分为四种类型:

(1) 因果关系类选项, 有明显的因果线索词。如, 2016 年北京高考试题 “A. 作者产生神秘感的原因是看见演唱老腔的是白发白眉老汉等一群关中农民。”

(2) 主旨, 线索, 写作手法类选项, 该类选项指原文运用了哪些写作原则、规律和方法来表达原文的内容, 主要有表达方式, 表现手法及行文结构等。如 2012 年北京高考试题 “D. 本文可以看作是《绝项》的创作谈, 交代了小说主人公的生平背景, 同时也折射出作者的心路历程。”

(3) 思想情感类选项, 该类选项指出作者借以表达的思想、感情及态度。如, 2014 年北京高考试题 “E. 文章既蕴含着历史感兴, 也渗透了现实关怀, 表达了作者对提升民族文化素质的热望。”

(4) 其它类选项, 无明显上述特征的选项。

通过对近 13 年北京高考试题以及 769 篇模拟题选项的分类统计, 因果关系类选项占全部选项的比重约为四分之一, 如表 1 所示。因此本文重点针对因果关系类选项提出相应的解决方案。

表 1 因果关系类选项比例统计

	选项总数	因果关系类选项个数	所占比例
模拟试题	3785	1050	27.7%
北京高考试题	63	16	25.4%

针对因果关系类选项, 本文提出一种汉语阅读理解中选择题的因果关系支持度分析方法。首先确定因果选项及选项中的因变量和果变量; 然后通过线索短语从文学作品中抽取因果事件对, 计算事件对之间因果关联度, 利用因果关联强度与抽取到的因果事件对构成因果关系网; 最后结合因果关系网及相关句综合分析选项的因果关系支持度, 判断因果关系类选项是否正确。

## 2 相关工作

针对阅读理解任务提出的相关技术也可以分为两种: 一种是基于特征的方法<sup>[4,5,12,13]</sup>, 另一种为基于神经网络模型的方法<sup>[6-11]</sup>。基于特征的方法需

要设计适合的特征来表征文本, 目前用到的特征有: 句法特征, 词频特征, 词距离特征, 框架语义特征, 词向量特征等。而神经网络模型方法是一种端到端的方法, 直接用阅读理解语料训练复杂的模型, 其中深度学习模型 (如, RNN 模型) 是应用最多的模型之一。神经网络模型方法是解决阅读理解任务的主流方法, 但由于数据集稀疏性和复杂性, 效果不及基于特征的方法。因此本文将阅读理解任务分为四种类型, 目的是针对不同类型选项设计不同的特征, 采用基于特征的方法寻求阅读理解任务的解决方案。

本文重点解决文学作品阅读理解中的因果关

系类选项判断问题。因果关系是文本理解中非常重要的一类语义关系，因此对因果关系的研究由来已久。Garcia<sup>[14]</sup>通过分析表达因果关系的动词实现了一个 COATIS 系统，该系统可以抽取显式因果关系。Girju 等<sup>[15]</sup>通过 WordNet 搜寻因果关系动词，实现了特定事件因果关系的自动识别。Khoo 等<sup>[16]</sup>认为除了动词之外，连词、形容词、副词及子句均能表达因果关系，采用模板匹配的方法抽取显示的因果关系。通常采用模板匹配的方法需要大量的人工标记，适应性不强及限制了因果关系的应用研究，因此利用统计机器学习的方法从文本中抽取因果关系就成为主流方法。Marcu 等<sup>[17]</sup>分析相邻句子间的词对概率，提出了采用贝叶斯方法抽取因果关系的方法。杨竣辉等<sup>[18]</sup>将文本划分为事件，提出了基于语义事件的因果关系抽取方法。

张志昌等<sup>[19]</sup>面向阅读理解 why 型问题设计了话题间因果关系识别，该方法采用话题中的普通名词、动词、形容词或者副词之间因果关系概率的乘积来表示话题间存在的因果关系概率。该方法采用线索短语模板从大规模语料中抽取由单个词组成的因果关系对，抽取模板如图 1 所示。其中，*c* 表示原因，*e* 表示结果。

综合上述研究，抽取因果关系的目的是识别因变量和结果变量之间的统计相关性，因变量和结果变量通常被句子中的词替代。但对于文学作品而言，通常单个词所表达的语义不能完全表达因果关系，因此本文通过对句子做依存句法分析，采用三种主要的关系类型短语（动宾关系、定中关系及状

中关系）作为因变量和结果变量，如“敷上\_藏药一）苏醒\_过来”。

[<e>...</e> because <c>...</c>]
[because<c>...</c> , <e>...</e>]
[<c>...</c> in order that <e>...</e>]
[<c>...</c> in order to <e>...</e>]
[<c>...</c> so that<e>...</e>]
[<e>...</e> for the reason that<c>...</c>]
[<c>...</c> ][therefore <e>...</e>]
[<e>...</e> ][thus <e>...</e>]
[<e>...</e>][hence <e>...</e>]

图 1 抽取含特殊线索短语的句子模板（文献[19]）

### 3 因果关系类选项问题的形式分析

在文学作品阅读选择题中，题干通常类似，均是对原文的理解正确或不正确，如“下列对文章内容的理解，不正确（或正确）的两项（或一项）是：”。因此本文忽略题干的作用，直接采用选项与原文理解是否一致的策略来选出最符合文意的选项。阅读理解中选择题可形式化的描述为一个三元组：

$$\langle D, C, A \rangle$$

其中 *D* 表示原文，*C* 为选项，在本文表示因果关系选项，*A* 表示答案（如果正解标为 1，否则被标记为 0）。

通过分析因果关系类选项，得出此类选项经常的错误类型，因此本文重点分析选项与原文文意是否存在以下错误，如表 2 所示。

表 2 因果关系类选项的错误类型

因果错误类型	错误说明	举例
原因不全	该类选项只给出部分原因	[北京高考试题/2016] A.作者产生神秘感的原因是看见演唱老腔的是白发白眉老汉等一群关中农民
因果颠倒	选项描述的因果变量与原文所描述的颠倒	[北京高考试题/2015] A.与雅致的“苏菲”相比，“梅”是不大动听的中文名字，因此令作者耿耿于怀
强加因果	因果选项在原文中没有因果关系	[北京高考试题/2013] A.饮食在作者关于故乡的感性中占有重要地位，根源于作者无法追怀的童年
原因或结果不准确	选项在原因部分与原文所描述不匹配	[北京高考试题/2011]C. 河西走廊的过客对祁连山的形容之所以见仁见智，是因为雪岭常被一派空濛笼罩，真相 难以认清

### 4 方法描述

对应于第 3 小节所提的具体因果选项与原文中相关句之间关系的几种错误类型，本文提出了一种

高考语文阅读理解中选择题的因果关系支持度分析方法。首先，确定选项中的因变量和果变量，根据因果线索词的类型来区分因果关系句的因果变

量；其次，抽取与选项相关的候选句，分别根据因变量和果变量在原文中找出其相关的候选句；最后，计算选项与相关句子之间的因果关系支持度。

#### 4.1 确定选项中的因变量和果变量

这一部分介绍选项因果变量的识别，首先在哈工大发布的中文篇章级句子间关系语料库（HIT-CDTB）中搜索因果关系线索词，共收集了116个因果线索词。然后参考HIT-CDTB体系将因果线索词分为原因在先类的词和结果在先类的词，如图2所示。根据图2所示的因果线索词分类来确定选项的因变量和果变量，例如，线索词“根源于”为结果在先类词，因此该线索词前面的部分被认为是果变量，后面部分被认为是因变量。

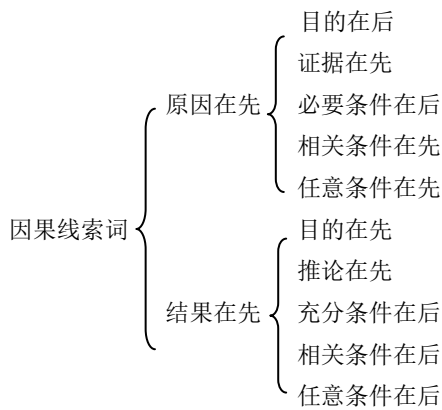


图2 因果线索词分类图

#### 4.2 相关候选句抽取

本文采用TF-IDF方法从原文中检索与选项相关的句子，综合考虑了选项中的每个词在文档中的重要性和整个文档中的区分能力。首先，对于原文按句号划分成句子集  $D = \{S_0, S_1, \dots, S_i, \dots, S_m\}$ ，选项和划分的句子均表示为由词组成的简单集合， $C = \langle u_0, u_1, \dots, u_i, \dots, u_n \rangle$ ， $S_i = \langle w_0, w_1, \dots, w_i, \dots, w_o \rangle$ 。具体TF-IDF计算方法如下：

$$\text{Rat}(C, S_i) = \frac{1}{n} \sum_{j=1}^n TF_j * IDF_j \quad (1)$$

其中： $TF_j$ 为选项中第j个词与原文中句子 $S_i$ 匹配的词数， $IDF_j$ 值采用式（2）计算得到：

$$IDF_j = \log\left(\frac{S}{S_{uj}}\right) \quad (2)$$

其中： $S$ 表示原文中句子总数； $S_{uj}$ 表示词 $uj$ 在原文中出现过的句子数。

#### 4.3 因果关系支持度分析

##### 4.3.1 构建因果关系网

这一部分重点介绍文学作品因果关系网的构建，因果关系支持度依据所构建的因果关系网计算获得。

1) 因果句子的抽取。与文献[19]类似，采用因果线索词从大规模语料中抽取因果句子。具体的抽取模板为一个三元组<线索词，[句1]，[句2]>。模板中的线索词为4.1小节所提到的线索词，[句1]为原因句，[句2]为结果句。根据线索词类型不同，[句1]和[句2]相对于线索词所处的位置不同。如果线索词类型为原因在先，那么[句1]在[句2]的前面，如“也许是因为本身的古典情怀[句1]，所以才对江南古院幽巷有着难以割舍眷恋[句2]”；如果线索词类型为结果在先，那么[句2]在[句1]的前面，如“朋友的可贵[句2]是因为曾一同走过的岁月[句1]”。

##### 2) 因果事件对抽取。

这一阶段是从因果句子中抽取因果事件对。文献[19]把因果句子看作为一个词的集合，该集合包含普通名词、动词、形容词或者副词，该方法集合中的元素较多，构建因果关系网计算量大，容易有重复的信息。文献[20]采用动词和名词集合来表示事件，该方法主要用来计算新闻标题之间的因果关系，对于文学作品中的普通句子，事件集合中有很多名词或动词，选择名词或动词本身也是一件很困难的事。本文针对文学作品中的因果句子采用名词短语表示因果事件，首先，对[句1]和[句2]做依存句法分析；然后，提取句子中的名词短语表示事件。名词短语的关系类型包括动宾关系（VOB）、定中关系（ATT）以及状中关系（ADV）。为了避免重复提取句子中名词短语，将抽取三种关系类型做排序，按优先顺序抽取相关的名词短语。

[例1]产生神秘感的原因是看见演唱老腔的是白发白眉老汉等一群关中农民。

看见老汉[VOB]—>产生神秘感[VOB]

[例2]天灾中沦为废墟了，更引发人们对艺术创造的热情

沦为废墟[VOB]—>艺术创造的热情[ATT]

##### 3) 构建因果关系网。

文献[18]把因果关联强度定义为因果事件同时出现的概率与两者析取出现的概率之比。本文借用文献[18]因果关联强度与抽取到名词短语事件构成因果关系网，因果关系网用三元组来表示为 $Net = [\text{因事件}, \text{果事件}, \text{关联强度}]$ ，例如，[残缺\_美，作为

经典, 2.66783669312e-05]。具体因果关联强度计算方法见式 (3):

$$P(T_1, T_2) = \prod_{w_1 \in T_1, w_2 \in T_2} P(w_1, w_2) \quad (3)$$

$$\begin{aligned} P(w_1, w_2) &= P(w_1 \wedge w_2) / P(w_1 \vee w_2) \\ &= \{s \mid w_1 \in \text{cause} \wedge w_2 \in \text{effect}\} / \{s \mid w_1 \in \text{cause} \vee w_2 \in \text{effect}\} \end{aligned} \quad (4)$$

其中  $\{s \mid w_1 \in \text{cause} \wedge w_2 \in \text{effect}\}$  表示在原因部分出现词  $w_1$  且在结果部分出现词  $w_2$  的句子数目; 同理  $\{s \mid w_1 \in \text{cause} \vee w_2 \in \text{effect}\}$  表示在原因部分出现词  $w_1$  或在结果部分出现词  $w_2$  的句子数目。图 3 为所构建因果关系网的部分示例, 图中节点表示抽取的名词短语事件, 有向边表示事件的因果关系。

### 4.3.2 因果支持度分析

接下来根据所构建的因果关系网来分析选项与相关句因果关系的的支持度。根据选项的因变量与果变量能否抽取到相关句, 判断具体选项与相关句子之间的关系是否一致可分为以下两种情形:

其中  $T_1$  表示因事件,  $T_2$  表示果事件, 分别包括两个词,  $w_1$  和  $w_2$  不为同一个词,

(1) 选项中的因变量及果变量都可以抽取到相关句, 这种情形需要计算抽取到的相关句之间的因果支持度, 如图 4 所示。

例 [北京高考试题/2004]

选项: 普希金作画从来不用画笔, 说明普希金的绘画出自他的自我意识。

抽取相关句子: ①而且他从来不用画笔作画, 只用写作工具鹅毛笔和墨水来画。②普希金的自画像大都出自诗人的自我意识, 正如诗人写诗大都以“第一人称”为出发点一样

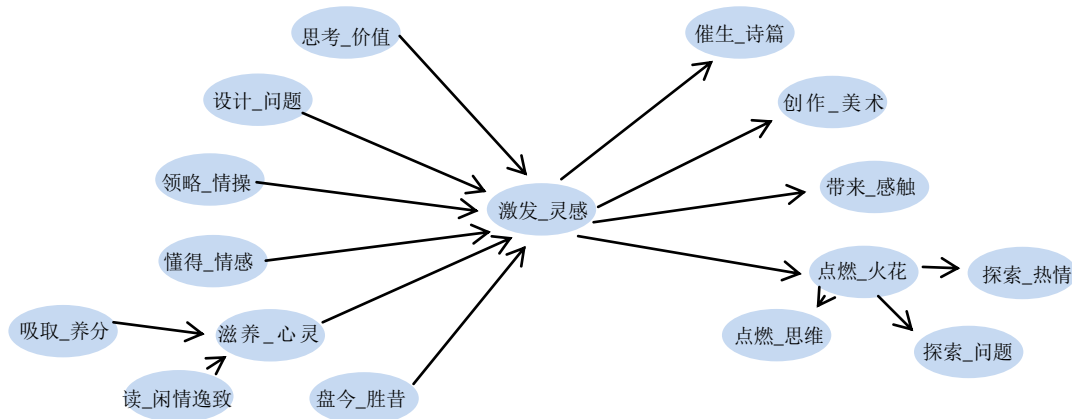


图 3 因果关系网示例

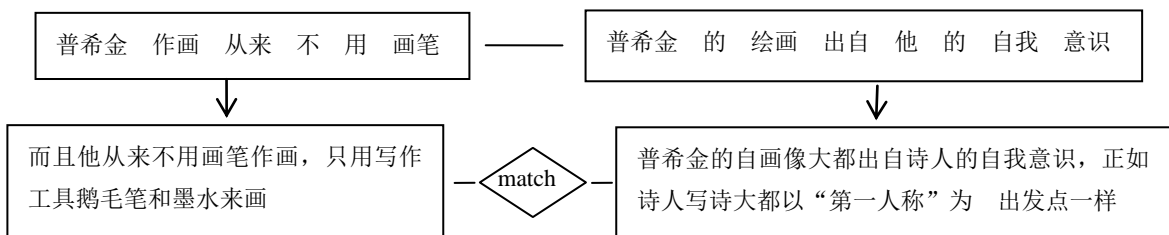


图 4 原文中存在与选项对应的因果变量 (情形 1)

这种情形需要计算相关句之间的因果关系支持度, 具体计算公式如下:

$$Causality(S_b, S_r) = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n P(T_i, T_j) + Rat(C, S_b) + Rat(C, S_r) \quad (5)$$

其中,  $P(T_i, T_j)$  表示因变量句子  $S_b$  中事件  $T_i$  和果变量句子  $S_r$  中事件  $T_j$  之间的因果关联度, 可以通过因果关系网获得,  $m, n$  分别为因变量句子和果变量句子中的事件个数。另外, 采用 TF-IDF 的值  $Rat$  来计算选项与相关句之间的相似度。

(2) 很多选项中的因变量或果变量抽取不到相关的句子, 这种情形需要采用选项的因变量或果变量补齐, 然后确定选项和相关句之间的因果支持度, 如图 5 所示。

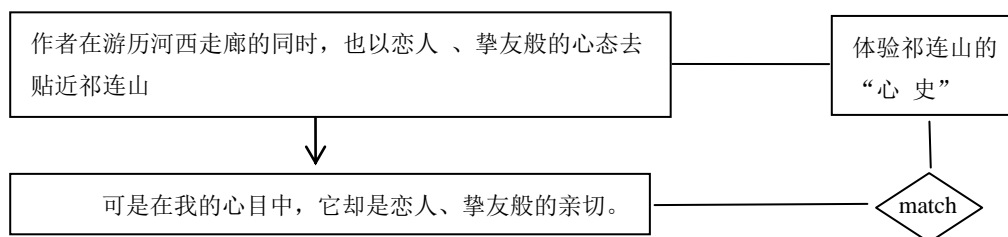


图 5 原文中缺少与选项对应的因变量或果变量 (情形 2)

这种情形需要计算相关句与选项之间的因果 关系支持度, 具体计算公式如下:

$$Causality(S_b, C) = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n P(T_i, T_j) + Rat(C, S_b) \quad (6)$$

其中,  $P(T_i, T_j)$  表示因变量句子  $S_b$  中事件  $T_i$  和选项中的果变量句子中事件  $T_j$  之间的因果关联度, 可以通过因果关系网获得,  $m, n$  分别为人因变量句子和果变量句子中的事件个数。

## 5. 实验

本文实验所用的语料包括用于抽取因果关系的 82367 篇文学作品阅读材料, 和测试材料 (13 篇北京高考语文试卷材料+769 篇模拟材料), 该语

例 [北京高考试题/2011]

选项: 作者在游历河西走廊的同时, 也以恋人、挚友般的心态去贴近祁连山, 试图体验祁连山的“心史”

抽取相关句子: ①可是在我的心目中, 它却是恋人、挚友般的亲切。② ‘ ’

料由山西大学中文信息处理课题组收集。本文所测试语料均含有一道选择题, 且每道选择题含有五个选项, 每个选项根据其是否符合原文被标记为 0 或 1 (如果符合原文的文意标为 1, 否则被标记为 0), 本文重点解决因果关系类型选项。文学作品阅读材料用来构建因果关系网, 共抽取了 425638 个因果关系句及 1042570 个三元组因果关系, 表 3 显示了关联度最高的 5 组因果关系对。

表 3 数据统计

因事件	果事件
上_本科线	没有_实现
思想_伟大	统治_世界
想念_父亲	满怀_期待
继续_前行	迎接_考试
感恩_梦	更加_明白

### 5.1 实验设置

为了验证本文方法在汉语阅读理解选择题中的有效性, 实验设置了相关句检索和答案选择两个阶段。相关句检索阶段, 本文采用 TF-IDF 方法从

原文中检索与每个选项相关的句子。答案选择阶段, 为了验证本文方法的优势, 实验设置了相关文献中常用方法做为 Baseline 进行比较分析。对于 Baseline 中的方法通过计算所有选项与原文相关句

的一致性给选项排序来判断选项正确与否，而本文方法仅关注因果关系类选项，通过式（5）或式（6）计算选项与原文相关句的因果关系支持度来判断选项正确与否。经过在测试语料中实验调整，本文设定因果关系支持度阈值为 0.83。Baseline 方法包括：

（1）基于词袋模型方法<sup>[21]</sup>

首先通过 82367 篇文学作品构建一个词表，对于选项和选项相关句子，分别计算其每一个关键词的出现次数，得到的对应向量分别为  $S_c=(f_1, f_2, \dots, f_n)$  和选项相关句子  $S_r=(f'_1, f'_2, \dots, f'_n)$ ，其中  $n$  为词表的大小。选项和选项相关句子的相似度利用两个向量之间的夹角余弦值来表示：

$$Sim(S_c, S_r) = \frac{\sum_{i=1}^n (f_i \times f'_i)}{\sqrt{\sum_{i=1}^n f_i^2 \times \sum_{i=1}^n f_i'^2}} \quad (7)$$

（2）基于知网语义（HowNet Semantic, HNS）的方法<sup>[22]</sup>

对于选项  $S_c$  和选项相关句子  $S_r$  中的任意单个词  $A_i$  和  $B_i$ ，基于 HowNet 计算其相似度  $S(A_i, B_i)$ ，取  $a_i = \max\{S(A_i, B_1), S(A_i, B_2), \dots, S(A_i, B_m)\}$ ， $b_i = \max\{S(A_1, B_i), S(A_2, B_i), \dots, S(A_n, B_i)\}$ ，选项和选项相关句子的相似度通过如下公式计算获得：

$$Sim(S_c, S_r) = \left[ \frac{\sum_{i=1}^m a_i}{m} + \frac{\sum_{i=1}^n b_i}{n} \right] / 2 \quad (8)$$

表 4 本文方法在测试集上召回率实验效果

测试集	召回率/%
高考题	18.7
模拟题	21.9

从表 4 中可以看出本文方法召回率偏低，原因主要为两点：（1）因果关系网覆盖范围小，本文所构建因果关系网中的事件都是从文学作品中的句子所抽取，而每一篇文学作品用词风格迥异，因此因果关系网不可能覆盖高考语文阅读的所有材料；（2）选项高度凝练，仅仅从相关句中无法理解其

（3）基于多维度投票的方法<sup>[12]</sup>

该方法将不同维度的度量作为投票依据，分别采用基于词袋模型、基于知网语义模型、基于 Word2vector 模型和基于汉语框架网的模型作为投票的维度，计算相关句子与选项之间的语义相关性。计算公式如下：

$$m(A) = \frac{1}{1-K} \sum_{B \cap H \cap W \cap C = A} m1(B)m2(H)m3(W)m4(C) \quad (9)$$

其中， $B$  表示基于词袋模型的词匹配策略， $H$  表示基于知网语义相似度， $W$  表示基于 Word2vector 模型的策略， $C$  表示基于汉语框架网的模型的方法。

5.2 实验结果及分析

考虑到高考语文题比模拟题更加规范，且本论文主要针对国家“863”研究发展计划项目中北京高考语文阅读理解所提的解决方案，因此对高考材料和模拟材料分开进行测试。其中北京高考语文阅读理解共 16 个因果关系类型选项和模拟材料共 1050 个因果关系类型选项。

首先，通过召回率  $R$  来作为本文所提方法的评价指标，计算公式如下：

$$R = \frac{r\_num(n)}{n} \times 100\%$$

其中， $r\_num(n)$  表示本文所构建因果关系覆盖的选项， $n$  表示因果关系类型选项总数。本文方法在两个测试集上召回率如表 4 所示。

选项的本意，如 2014 年北京高考语文阅读理解选择题：“A. 作者写本文的目的之一是纠正《现代汉语词典》中对“废墟”这一语词的错误理解”。理解选项中“本文的目的”需要理解全文或了解全文的主旨才能做到。

其次，由于选择题有准确的答案，对此本文采

用的另一个评价指标为准确率:

$$P = \frac{CorrectAnswer(k)}{k} \times 100\%$$

其中,  $CorrectAnswer(k)$  表示  $k$  个选项中正确的选项数目。表 5 分别给出了本文方法与相关方法在高考材料和模拟材料的实验结果比较。

通过表 5 可以看出, 本文方法解决因果关系类选项是有效的, 在高考题上的准确率与基于多维度投票方法相比提高了 11 个百分点。另外, 我们分析了本文方法错误的原因, 发现约有 20% 的错误选

项是由于相关句抽取不准确所致, 对于文学作品阅读理解来说, 很多选项具有高度的概括性, 很难抽取其与原文中相关的句子, 例如 2013 年北京高考语文阅读理解选项 “A. 饮食在作者关于故乡的感性中占有重要地位, 根源于作者无法追怀的童年”, “占有重要地位” 和 “无法追怀的童年” 都是对原文的高度概括, 很难从原文中准确抽取其相关句。

表 5 本文方法与相关方法的结果比较

实验方法	基于词袋模型方法	基于知网语义方法	基于多维度投票的方法	本文方法
高考题	46.1	50.0	55.7	66.7
模拟题	45.2	48.7	53.4	63.9

## 6 总结与下一步的工作

根据文学作品阅读理解选择题中选项的特点, 本文把选项分为四种类型。针对因果关系类型选项, 构建文学作品类因果关系网模型, 并依据因果关系网提出了因果关系类选项的因果支持度分析方法。实验结果表明, 本文方法召回率偏低, 而准确率与 Baseline 方法相比在两种测试集上均有所提高。为了提高召回率, 下一步工作将考虑构建普通事件因果关系网上一层的抽象因果关系网模型, 提高因果关系网覆盖范围, 下一步工作还需要考虑其它类型选项的解决思路。

### 参考文献

[1] Danqi Chen, Jason Bolton and Christopher D.Manning. A thorough examination of the CNN/Daily mail reading comprehension task[C]. In Proceeding of the Association for Computational Linguistics, 2016.

[2] 刘知远, 孙茂松, 林衍凯, 谢若冰. 知识表示学习研究进展[J]. 计算机研究与发展, 2016, 53 (2): 247-261.

[3] 林鸿宇, 韩先培. 机器阅读理解任务综述. 中国中文信息学会青工委学术专栏第 6 期, 2016.

[4] Mrinmaya Sachan, Avinava Dubey, Eric P Xing and Matthew Richardson. 2015. Learning answer-entailing structures for machine comprehension[C]. In Proceeding of the 53th Annual Meeting of the Association for Computational Linguistics, pages 239-249

[5] Adam Trischler, Zheng Ye and Xingdi Yuan. 2016. A parallel-hierarchical model for machine comprehension on sparse[C]. In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 432 - 441.

[6] Wenpeng Yin, Sebastian Ebert and Hinrich Schutze. 2016. Attention-based convolutional neural network for machine comprehension[C]. In Proceedings of 2016 NAACL Human- Computer Question Answering Workshop, pages 15-21.

[7] Yiming Cui, Ting Liu, Zhipeng Chen, Shijin Wang and Guoping Hu. 2016a. Consensus attention- based neural networks for Chinese reading comprehension. arXiv preprint arXiv: 1607.02250.

[8] Yiming Cui, Zhipeng Chen, Si Wei, Shijin Wang, Ting Liu and Guoping Hu. 2016b. Attention-over-attention neural networks for reading comprehension. arXiv preprint arXiv: 1607.04423.

[9] Karl Moritz Hermann, Tomas Kocisky, Edward Grefenstette, Lasse Espeholt, Will Kay, Mustafa Suleyman and Phil Blunsom. 2015. Teaching machines to read and comprehend[J]. In Advances in Neural Information Processing Systems, pages 1684-1692.

[10] Danqi Chen, Jason Bolton and Christopher D.Manning. 2016. A thorough examination of the CNN/Daily mail reading comprehension task[C]. In Proceeding of the 54th Annual Meeting of the Association for Computational Linguistics, pages



2359-2367.

[11] Rudolf Kadlec, Martin Schmid, Ondrej Bajgar and Jan Kleindienst. 2016. Text understanding with the attention sum reader network. arXiv preprint arXiv:1603.01547.

[12] 郭少茹, 张虎, 钱揖丽, 李茹, 杨陟卓, 马淑晖. 面向高考阅读理解的句子语义相关度研究[C]. 第十五届全国计算语言学会议, 2016.

[13] Matthew Richardson, Christopher J C Burges and Erin Renshaw. 2013. MCTest: a challenge dataset for the open-domain machine comprehension of text[C]. In Proceedings of the 2013 on Empirical Methods in Natural Language Processing, pages 193-203.

[14] Garcia Daniela. An NLP system to locate expressions of actions connected by causality links[C]. In Proceedings of the 10<sup>th</sup> European Workshop on Knowledge Acquisition, Modeling and Management. Catalonia, Spain, 1997:347-352.

[15] Girju Roxana. Automatic detection of causal relations for question answering[C]. In Proceedings of the 41<sup>st</sup> ACL Workshop on Multilingual Summarization and Question Answering, Sapporo, Japan, 2003:76-83.

[16] Khoo Christopher, Kornfilt Jaklin, Oddy Robert, et al. Automatic extraction of cause-effect information from newspaper text without knowledge-based inferencing[J]. Literary and Linguistic Computing, 1998, 13(4):177-186.

[17] Marcu Daniel, Echihabi Abdessamad. An

unsupervised approach to recognizing discourse relations[C]. In Proceedings of the 40<sup>th</sup> Annual Meeting on Association for Computational Linguistics, Philadelphia, USA, 2002:368-375.

[18] 杨竣辉, 刘宗田, 刘炜, 苏小英. 基于语义事件因果关系识别[J]. 小型微型计算机系统, 2016, 37(3):433-437.

[19] 张志昌, 张宇, 刘挺, 李生. 基于语题和修辞识别的阅读理解 why 型问题回答[J]. 计算机研究与发展, 2011, 48 (2): 216-223.

[20] Sendong Zhao, Quan Wang and Sean Massung. Constructing and embedding abstract event causality networks from text snippets[C]. In Proceedings of the 10<sup>th</sup> ACM International Conference on Web Search and Data Mining, 2017:335-344.

[21] 张志昌, 张宇, 刘挺等. 基于浅层语义树核的阅读理解答案句抽取 [J]. 中文信息学报, 2008, 22(1):80-86.

[22] 朱征宇, 孙俊华. 改进的基于《知网》的词汇语义相似度计算 [J]. 计算机应用, 2013, 33(8):2276-2279.

作者联系方式: 王元龙 山西省太原市坞城路 92 号山西大学 030006 15234154712  
ylwang@sxu.edu.cn