

写作文本中的意义逻辑及其计算

赵兴龙¹, 何克抗²

(1.中央电化教育馆,北京 100031; 2.北京师范大学 现代教育技术研究所,北京 100875)

摘要 :本文从写作文本的角度探讨意义逻辑是语言和思维关系讨论深入的表现。意义逻辑是将客体的意义同化到特定语言情境中时所产生的逻辑。文章在分析意义逻辑及其特征的基础上,探讨了意义逻辑的计算模型。意义逻辑可以通过套件的显性化和层次化构建计算模型,并通过“句子—判断”“句群—推理”的方法进行计算。意义逻辑及其计算在写作文本分析中具有非常重要的实践意义,为语言材料中提取出现场思维状态提供了一种可计算的途径,使得不可观察的思维状态可以用计算的方式进行理解和解释。

关键词 :意义逻辑;形式逻辑;写作文本;作文分析;语言运用

中图分类号 :G434 **文献标识码** :A

一、研究问题的提出

语言和思维的关系一直以来是学术界关注的问题。研究者认为语言和思维是相关的,语言是思维的一个预测器^[1]。所谓思维的理论实际上就是对思维状态的一种理解,思维状态包括信念、欲望和知识,这些促使我们去解释和预测别人的行为^[2]。语言对于思维之所以很重要是因为思维状态是不可观察的,比如“跑”这个动词的意义可以通过观察“跑”这个事件来学习,而“想”这种反映内部活动或状态的动词不可能通过观察的手段获得^[3]。思维状态到底是通过什么样途径映射到语言材料当中的?我们怎么样从语言材料中析取现场思维状态?本文提出意义逻辑这个概念来专门解决这些问题。

二、意义逻辑及其特征

“意义”这个词语在英文中的表示是“Signification”,李其维教授认为客体的意义就是它能被用来做什么,将客体同化到一个动作格式过程中产生的逻辑就是意义逻辑^[4]。意义逻辑中最基本的逻辑——数学关系是意义蕴含。“意义蕴含”的界定为:如果 q 的一个意义 m 存在于 p 的意义之中并且这种意义 m 是传递性的,那么 p 蕴含 q ^[5]。

那究竟什么是意义逻辑呢?首先来看下面的这个非常有意思的例子:

“我带了一些点心去上班,并告诉我的同事乔治这些点心放在我办公桌的抽屉里,他自己可以随便吃。过了一会,我注意到有几只蚂蚁正爬向办公桌,我随即决定把这些点心放到一个高一点的柜子上。那天下午,当我去邮件室的时候,看见乔治走向我的办公室,我说到:“点心在那个柜子上”。其实,我几乎

毫不费力的想到乔治想吃点心,但是他却错误地认为点心在抽屉里,结果肯定是找不到。如果我告诉他点心在什么地方,我将改变他的“认为”,满足他吃点心的渴求。”^[6]

这段话实际上包含了1个直言推理和3个假言推理:

推理一:直言推理

要想吃到点心就必须知道点心现在准确的位置

乔治想吃点心

所以,乔治就必须知道点心现在准确的位置

推理二:假言推理

如果没有人告诉乔治点心位置会改变,乔治就认为点心还在抽屉里

我并没有告诉乔治点心已经放在柜子上

所以,乔治始终认为点心放在抽屉里

推理三:假言推理

如果不知道点心现在放置柜子上,就肯定找不到

乔治认为点心放在抽屉里而不是柜子上

所以,乔治肯定找不到点心

推理四:假言推理

如果有人事先告诉乔治点心的位置,将改变乔治的“认为”

我事先告诉乔治点心的位置

所以,我将改变乔治的“认为”

从上述推理中可以非常清晰地了解乔治为什么找不到点心的一种思维状态和这个文本中的“我”能根据乔治的思维状态来解释和预测乔治行为的一种思考原理,进而说明“思维理论”的基本内涵。有意思的地方恰恰就在这里,其实这就是生活中再普通不过的一个场景,当将这个场景外显出来之后就需用4个推理才能表示清楚,非常复杂。如果日常生活中

使用这样复杂的推理进行交流的话,可能反而阻碍了交流的正常进行,但是事实上这种内部的复杂推理确实存在着,我们却很少或几乎没有去自觉进行回想^[7]。

要界定意义逻辑,我们需要先来回顾一下人工智能领域中这种逻辑的研究。意义逻辑在人工智能领域中的术语表达是“缺省逻辑”。在人工智能领域中,率先开始较为深入系统地研究意义逻辑的学者是 Reiter。Reiter 举了一个鸟的例子来说明这种缺省逻辑:

“一般都有这样的逻辑形式‘大多数的 P 是 Q 的’或‘大多数的 P 具有 Q 属性’。比如说这个例子‘除了企鹅、鸵鸟、马耳他猎鹰等,大多数的鸟都能飞’。那么对于现在给定的一种鸟,除非它可能是上述三者中的一种,否则它就是能飞的。”^[8]

但是这种表述方法并不能推演出一般的鸟都能飞这个事实,只能试图去证明——如果具有了“飞”的基因,这一定是鸟。那么就鸟这个例子而言,具有“基因”能“飞”在什么条件下是可以缺省的(即默认的呢)? Reiter 采用了这样的默认规则:如果 x 是一只鸟并且“一直假定”x 能飞,那么就能推断出 x 能飞。“一直假定”的意思是在相反情况的任何信息都缺失。依据上面的例子来说就是,像企鹅、鸵鸟、马耳他猎鹰等等这些反例都不存在,给定的任何一种鸟都是能飞的。其实这里的“一直假定”就是一种缺省规则。

缺省逻辑具有非单调特征。传统一阶逻辑具有单调性特征,它的公式表达为 if A、B 两个集合存在并且 $A \vdash w$, then $A \cup B \vdash w$ (\vdash 在逻辑学中的意义为“推导自”,这个公式的意思是 w 推导自 A,那么 w 也能推导自 A 与 B 的并集),也就是说当新的信息 B 被发现时,A 仍然保持原有的有效性。而对于缺省逻辑而言则不同,Reiter 认为对于:MA/B 而言,B 可能是可信的。如果随后的非 A 被发现的话,那么对于:MA/B 而言,有非 A 存在的 B 则不可信了^[9]。“在一般的日常生活中,人们可以假定某些命题总是正确的,除非有特别的关于例外的声明”^[10],当新的事实被认识即前提增加时,结论并不会增加甚至有可能还要被推翻^[11],由于这种推理非常适合语用分析,有时候学者也将它称为语用推理。不论是缺省推理也好,还是语用推理也罢,都是非单调逻辑,都基于形式化从不完全知识并将其转换到相对完全知识的知识状态^[12],这是非单调推理的第一个显著特征。非单调推理的第二个显著特征是可错性。“这类非演绎逻辑一般都是可错的,它们允许某种逻辑跳跃来达到一些好的猜测。”^[13]综合起来,本文结合李其维

教授的定义以及人工智能领域中缺省逻辑的研究成果,将意义逻辑界定为:将客体的意义同化到特定语言情境中时所产生的逻辑。意义逻辑具有以下四个方面的特征:

1. 可错性

对于意义逻辑而言,可能是非常标准的三段论逻辑体系,但也可能是日常生活中非常随意的交谈话语。这些话语推理结果的正确与否不具有非常严格意义上的形式逻辑特征,有时候出现的话语可能就是在“彼时彼景”中是可以理解的,在常识知识或在“此情此景”中不可理解甚至可能是错的。只要语义理解的双方都能“心知肚明”,那么这种推理在意义逻辑的研究中是得到承认的。

2. 片断情景

意义逻辑依靠“彼时彼景”中的上下文场合,这种上下文场合不像小说文本中的情景场合那样的完整和步步到位、丝丝入扣,而是一种短时间甚至几乎是瞬间的交际或交互情景,具有非常好的逻辑跳跃性。可能只有两句话的对话,却包含了对话双方能够理解的多重意思。

3. 多种推理形式的综合

意义逻辑中的推理不仅仅遵循着归纳推理的基本要义,而且还能通过演绎推理以及类比推理等多种推理形式来发生作用。所以意义逻辑中的推理包含了非单调推理的特征与实例,同时也包含着其他推理的特征与实例,这和人类本身思想的复杂程度有着直接的联系。因为在人类的推理中,要想对事物、现象、状态以及事物与事物之间的关系做出一种统一的规则描述不可能仅有归纳推理一种形式,是多种推理形式交互演算的综合。多种推理形式的综合为分析特定的意义逻辑单元提供了线索,一个意义逻辑单元中的推理必然也是多种推理形式的综合。

4. 形义不一

王维贤将推理的逻辑形式的“严格性”与语言表现形式的“灵活性”进行了意义上的鉴别,并认为这种语言形式可能借助省略、语境、次序颠倒等表现手法呈现出“表里不一、意在言外”^[14]。实际上,逻辑推理和语言表达在同一情景中是一致的。逻辑推理必然是严格的,即便存在某种“常识”的错误或经验的局限;语言表达本身是灵活的,可能言简意赅或形式残缺,但是通过表现手法和跳跃总能找到逻辑推理的结果。看似两者不能调和,其实却是某一意义逻辑单元分析的不同侧面而已。

三、意义逻辑的计算模型

(一)两种推理形式的分析

意义逻辑依赖于人的话语,而且拥有跟形式逻辑一样的“形式”,但结果未必就是真。也就是说,形式逻辑必须保证最后的结果为真值,而意义逻辑就不用这样严格恪守。意义逻辑是确实存在的,当前的问题是如何较为准确地将意义逻辑从话语中显现出来,恢复话语的“原貌”。通过这种“原貌”对话语分析,才是相对较为准确的途径;同时,也只有通过这种“原貌”的分析,才能准确分析说话者或写作者在语言运用过程中的推理特征以及复杂关系,从而对其内在差异状况做出相应的判断。

在形式逻辑中,条件推理是最为常见的一种形式,还是拿上面的推理来进行说明。在这个推理中,“要想吃到点心就必须知道点心现在准确的位置”是一个条件性命题,“乔治想吃到点心”是一个陈述性命题,“乔治就必须知道点心现在准确的位置”是由前两者推出的结论。在条件性命题中,“要想吃到点心”是前件,一般以 P 来表示。“必须知道点心现在准确的位置”是后件,一般以 Q 来表示。在陈述性命题中,“乔治想吃到点心”肯定了条件性命题中的前件 P,所以最后的结论是肯定了条件性命题中的后件 Q。这是一种典型的“正确肯定式”,即“有前件就有后件”。其实,依据 P 和 Q 的关系,这种条件推理一共有 8 种关系,分别为:肯定前件式(有 Q 和 \sim Q 两种结果)、否定前件式(有 Q 和 \sim Q 两种结果)、肯定后件式(有 P 和 \sim P 两种结果)和否定后件式(有 P 和 \sim P 两种结果)^[15]。在这 8 种关系中,肯定前件式和否定后件式的结论为真值,否定前件式和肯定后件式的结论为假值。

再看另一个例子,这个例子来自于日常生活中的话语:

A: 玛丽同一个瑞典人结了婚。

B: 啊,天啊! 玛丽和一个北极熊猎者结了婚。

在这个例子中,A 与 B 组成的话语中肯定是出现了条件推理,但是 B 的结论肯定不是 A 直接给推出来的。A 只是陈述性命题,这个推理中缺少了一个命题性条件,也就是省略了“前件”,这里使用 A' 来表示。添加 A' 将这个推理补充完整:

A': 所有的瑞典人都是北极熊猎者。

A: 玛丽同一个瑞典人结了婚。

B: 啊,天啊! 玛丽和一个北极熊猎者结了婚。

将前面的推理一与这个关于日常会话的例子进行比较,可以看出,日常会话中由于人脑的推理速度过快,往往在瞬间就可以得出结论,而且有时候根本不需要有前件的出现。而推理一中推理话语往往是按照条件推理的格式,前件、后件和结论三项一个都不能少,否则不能构成一个完整的推理结构。All-

wood 等人使用“隐藏”这一术语来表示话语前提的缺失^[16],如图 1 所示。

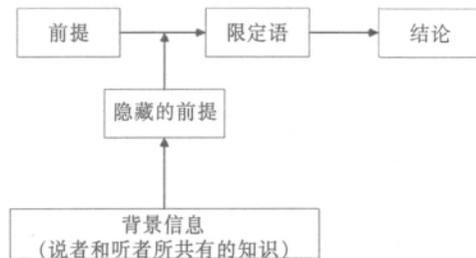


图 1 条件推理中“隐藏的前提”的示意图

(二)套件的计算模型

1. 套件的显性化计算

这里所说的套件,包括“隐藏的前件”或者是“隐藏的后件”。这些隐藏的前后件嵌套在话语交流的背景信息中,为了和“前件”“后件”的表述趋于一致,所以又将其称之为“套件”。实际上,对于话语交际中的绝大部分情况而言,套件确实是隐藏的。套件要能够可计算,首先就是要让“隐藏”像形式逻辑推理结构一样“显现”,才能准确地对套件中所要描述的意义及其意义之间的关系进行理解,这些理解来源于说者和听者(或者是读者与作者)所共有的背景信息;其次,从上面条件推理的 8 种形式中可以看出,肯定前件式和否定后件式的结论是真值,也就意味着在共同隐藏的背景信息中可能是前件,也可能是后件。如果是肯定前件式的推理,隐藏的一般为大前提;如果是否定后件式的推理中,隐藏的一般为小前提;再者,连接词的出现推理中也扮演着重要的辅助角色。“通过……”“使我……”“我一定……”“否则……”“如果……就……”等等,这些连接词只要出现在文本中,就已经表示了某种推理存在的可能性。一个理想的套件显性化的计算主要包括四个阶段:

第一个阶段:结果型话语捕捉

在写作文本中,粗略地可以分析出来两种不同的话语形式,一种是过程型话语,一种结果型话语。过程型话语主要用来描述整个事件的经过,结果型话语表达某种结果或表达自己的情感、观点等。这两种话语在写作文本尤其是儿童写作文本中表现得非常明显。比如下面的小短文共有七句话组成:

(1)过年的时候,回到老家,妈妈让我背唐诗。

(2)诗的里面有描写农民的,描写景色的,还有描写草的,还有描写荷花的好诗。

(3)我在家里很贪玩,很少时间背唐诗。

(4)妈妈一直催促我背,完了还要背给妈妈听。

(5)回到广州的时候我已经能背三十多首,觉得自己蛮好。

(6)可是真正到了古诗擂台赛的时候,我只背了九首,输给了同班的同学。

(7)我发现有些同学能背几十首唐诗,我离他们的成绩很远,以后一定要多思考,勤奋专心地读书,这样才能学习好。

其中第一句话到第四句话就是一种过程型话语,全部都是叙事的口吻。第五句、第六句和第七句话都包含了不同的推理过程,属于结果型话语,而这三句话中共包含了 5 个推理,分别是三段论(背诗多就是好,第 5 句)、三段论(背诗少就会输,第 6 句)、两个三段论(背诗多就成绩好,背诗少成绩就差很远,第 7 句前半句)、一个不完全归纳(一定要多思考、勤奋等,第 7 句后半句)。这样大致可以推断出写作文本中意义逻辑的推理机制往往潜藏在结果型话语中。所以,只要在作文文本中找到结果型话语的“藏身”之处,就基本上能够断定推理的出现位置。

第二个阶段:标签提取

标签提取包括两个方面:一是文本中提到的连接词,尤其是跟情景话题紧密相关的连接词;二是与情景话题紧密相关的内容核心词汇或短语。有了连接词和核心词汇或短语,就能将写作文本的关键内容简明扼要的提炼出来。比如在上面所举文本中第五个句子中从“以后”这个词语标签就能非常明确地判断出来这里的推理形式是“不完全归纳推理”。再比如一个小朋友在作品中这样写到:“又一次,我忘带了橡皮擦。我就问了颜嘉聪:“可不可以借我橡皮擦?”他说:“可以”。我心里很高兴,用完以后我就还给他了,还说了一声谢谢。我觉得颜嘉聪真是我的好朋友。”在这段话中,前半部分是大量的对话叙述,按照结果型话语的捕捉,推理应该集中在最后一句话上,而最后一句话中的“真是我的好朋友”这个内容核心词汇标签说明这段话中使用“不完全归纳推理”的形式,因为儿童眼中的好朋友就是相互借东西并且能够致谢的朋友。

第三个阶段:套件补足

根据连接词和核心词汇或短语的标签信息,找到相应的段落或句子,将隐藏的套件(前件或后件)按照演绎推理、归纳推理或类比推理的形式进行补足即显性化。

第四个阶段:推理识别

按照补足套件后的完整推理形式对推理的类型进行识别,这样就完成了套件显性化的计算过程。

2. 套件的层次化计算

套件显性化的过程只是套件计算模型的第一个步骤。套件显性化只是完成了对推理类别的识别,但是推理的数量如何进行量化又是需要计算模型解决

的另一个问题。简言之,就是需要确定在某一段话语中的推理类型的数量,确定数量的过程也就是套件层次化计算的过程。那么这个数量到底怎么样确定呢?

我们知道,层次化计算过程必然跟文本的长度有直接的关系,也跟文本的体裁有关系。然而有的文本虽然篇幅比较长,但是中间并没有涉及到意义逻辑的推理关系;有的文本平淡地叙述某件事情,与主题相关的逻辑推理倒是一大堆,这都是非常正常的事情。套件层次化计算的关键是找到能够跟作文文本主题紧密相关的意义逻辑推理,然后进行量化。套件层次化计算主要包括四个阶段:

第一个阶段:主题距离

所谓主题距离是指某一反映话语意义推理逻辑结构的内容体与该话语情景主题之间关系的远近程度。主题距离的出现是由意义逻辑中的“意义”所决定的,而不是由逻辑所决定的,这也是意义逻辑本身的特性。意义逻辑反映的内容是现实生活中发生的事件,而将这些事件凝聚在作文文本中用来表达某种思想或情感的过程是主题化的过程。内容是不是围绕着一个主题展开对于一个文本质量的优劣有着关键作用,那么显现出来的意义逻辑也需遵从主题来展开。但是由于写作者的写作表达能力存在差异,每一个表达意义逻辑结构的内容体并不总是与主题之间保持相当地一致,这就产生了远近关系。内容体不能围绕主题展开,可以认为内容体离主题的距离较远;反之,就认为内容离主题的距离较近或符合主题。与传统作文评价不同的是,传统的作文评价中如果作文偏离主题,基本上就否定了这篇作文的价值。但是本研究认为,当使用意义逻辑来表达主题距离的时候,距离只有远近,而没有严格的价值判断。偏离主题,只表示与主题的距离较远,而不能表示与主题没有关系。即使整篇作文偏离了所要求的主题,但是如果其中反映某种意义逻辑的内容体与主题的距离较近,这样的内容体也可以用来计算与主题的距离。

这里有一个例子可以说明:一个小朋友在作文中写道:“一天,上体育课的时候,我看见有一位同学他拿着一根绳子乱甩,这时老师下来了。老师看见那位同学拿着绳子乱甩,很生气地说如果你再乱甩的话,你以后也不用上体育课了。我知道,乱甩绳子是很危险的。”对于主题而言,这些句子都是情景话题之内。但是“老师看见那位同学拿着绳子乱甩,很生气地说如果你再乱甩的话,你以后也不用上体育课了。”这个句子虽然有一个假言推理,但是这个推理是由“老师”来发出的,并不是有文中的这个小作者“我”来发出的,而“我知道,乱甩绳子是很危险的”是

由小作者发出的一个归纳推理。所以在主题距离上,由老师发出的这个假言推理离主题的距离就要比“我知道,乱甩绳子是很危险的”这个归纳推理的距离远得多。在实际的计算中,这个假言推理是要被过滤的。

第二个阶段:有效距离识别

距离只有远近,而不存在价值意义。那么什么样的距离才能算是有效距离呢?也就是说,怎么样去识别一个内容与主题之间的距离是有效距离从而能将其内容体算入意义逻辑主题数量的范畴呢?实际上,有效距离识别的还是得来源于语义关系而不是语法结构。常用的识别有效距离的办法有两种:第一种办法是如果所提取的内容体中的核心关键词与主题具有较高的语义相似度,那么这个内容体与主题的距离就是有效的。这种识别方法的前提假设是句子或段落的关键意义能够被特定语境中的特征词所替代,一般在人工智能方面运用较多。但是,由于汉语语义的复杂性,这种语义相似度的计算还是不能非常顺利地运用到作文文本分析技术中。一方面是大量的研究还都集中在句子(或句类)层面上,人类自然语言的语义研究还仍旧是一个难以攻破的课题;另一个方面,大量研究在使用测试材料的时候,都是使用报纸、书刊等语料,这些语料的最大特点是语法正确、标点正确、句式相对完整。而作文文本就大大不同,作者在写作过程中不免要出现各种各样的语法错误甚至是语用错误,而这些错误通过计算机是很难完成的,但人能够轻而易举地完成,这也就是“意义”的独特之处。所以,采用人工判断虽然较为费力,但是在文本意义的理解上却是最合适不过的。有效距离识别主要通过人工识别的方法来判断某种逻辑结构的内容体与主题之间的距离有效性。这里采用评价者一致性信度来保证有效性程度。

第三个阶段:距离过滤

如果逻辑结构的内容体与主题之间的距离是有效的,那么就将这个逻辑结构添加到套件层次化的数量计算模型中。有效距离的逻辑结构内容体添加到套件层次化的数量计算模型中的过程,也就是距离过滤的过程。套件显性化和层次化的具体步骤示意图请见图 2。

四、意义逻辑的计算方法

(一)“句类—判断”分析

1. 句类与判断关系的辨析

作文话语分析中对于句子分析的着眼点并不在于什么是句子,而是句子如何进行分类。对于句子分类,到底选用哪一种分类作为本研究的参考标准呢?我

们知道,写作是用来交流知识、表达观点和情感的^[17],是儿童在不断使用和操练语言文字的过程中逐渐完成的^[18]。对于写作而言,母语儿童主要是运用语言而不是分析语言,是用写作这种方式来进行交流和自我表达而不是说明自己对某个句法成分或句式掌握的如何恰如其分;即使是一个非常老练的作家,一般也不会顾及自己一定得使用某种句子结构表达某个意思,句子结构以及句式的研究是语言分析家做的事情,并非是儿童语言运用中所使用的基本方式。所以,作文中句子分类的选择需要考虑的是语言运用的功能而非语言运用的结构。艾伟先生的分类是从属于语言运用的结构,在林杏光先生总结的六种分类中,按照结构特点、谓语性质、成分、所带词组、结构繁简分都是从属于语言运用的结构,而从语气分类的角度则考虑到了语言运用的功能,这四种类型的句子还渗透着某种情感的表达。本文采用了按照句子语气分类的疑问句、陈述句、感叹句和祈使句的分类方式。

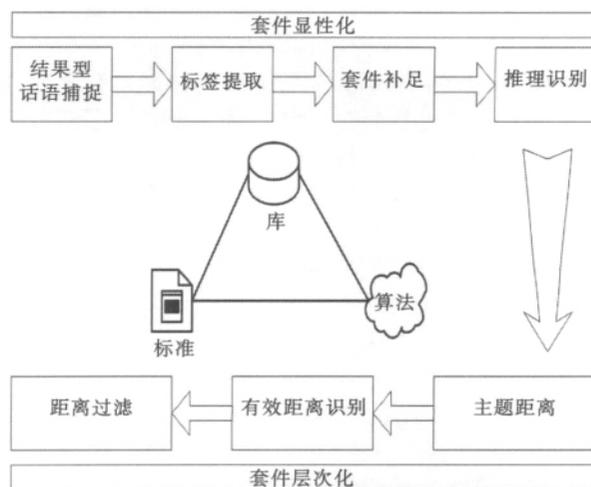


图2 意义逻辑中套件计算的模型表示

那何为判断?判断是思维的基本形式之一,肯定或否定某种事物的存在,或指明某一对象是否具有某种属性的思维过程。在形式逻辑上用一个命题来表达,表现形式相当于语言中的句子^[19]。语言文字研究者认为,判断是语句的思想内容,语句是判断的语言形式,任何判断都必须通过语句才能表达出来^[20]。但是任何判断都必须通过语句来表达,但并非任何语句都表达判断;同一语句可以表达不同判断;不同的语句表达同一个判断^[21]。判断分为四类:(1)直言判断,用来断定事物具有或不具有某种属性,也叫性质判断,一般又分为单称肯定判断、单称否定判断、特称肯定判断、特称否定判断、全称肯定判断、全称否定判断六类;(2)联言判断,用来断定几种事物情况共存的判断;(3)选言判断,断定在几种可能情况

下,至少有一种情况存在的判断,又分为相容选言判断和不相容选言判断;(4)假言判断是反映事物之间条件关系的复合判断,在假言判断中表示条件的判断叫做前件,表示结果的判断叫后件。假言判断又分为充分条件假言判断、必要条件假言判断和充分必要条件假言判断。

在判断与句类的关系辨析中,很多学者都做过细致的研究。命题是表示真假意义的判断,而“陈述句是可以翻译成形式逻辑中的命题的句子即可以判断是真是假的句子”^[22]。既然判断和陈述句之间有着这样的关系,对于一个作文文本,我们将其中陈述句提取出来就是命题,这些命题就是学习者思维加工的重要指示剂。

我们知道,祈使句不表示判断,疑问句也不表示判断,除非是双重疑问句等带有判断性质的句子。剩下的问题就是感叹句是否表达判断?这是一个非常关键的问题,将这个问题阐述清楚,有助于从作文文本中准确地提取出陈述句。一种观点认为感叹句不能表达判断。“一切没有判断内容的句子,无所谓是非的句子,如感叹句等都不能作为命题”^[23],其实这里否定了感叹句是命题,就已经完全否定了感叹句是判断,因为所有的判断都是命题。此外,有学者基于语音库从调阶、调域、重音和调尾四个方面分别对感叹句与陈述句的语调进行了研究,借以说明感叹句和陈述句划分它们之间的区别^[24]。如果将感叹句与陈述句直接区别开来,也就间接说明感叹句和陈述句不能同时都表达对某一事物情况的断定,即感叹句不能表达判断。对上述这种观点的质疑其实在 20 世纪的 80 年代就已经在公开发表的文章中初见端倪了。吕叔湘在讨论感叹句时认为“平常的言语大多兼有知识的和感情的成分”^[25]。朱德熙也认为感叹句表达情感的同时报道信息^[26]。实际上两位语言学大师的观点已经非常明确地表达了感叹句的实质:感叹句兼有表达情感和判断的双重任务^[27]。许爱琼认为感叹句一般都是可以表达命题的^[28]。她从情感—思想—语句三者之间内在的关系为切入点认为感叹句是表现人的情感的,总是在表达一定的思想内容,感叹句必须建立在陈述句的基础上对事物情况做出肯定或否定的断定,所以感叹句一般都表达命题(判断)。针对上述观点,她还给出了三类表达判断的感叹句类型:(1)陈述句加感叹语气的感叹句;(2)简化了的陈述句加感叹语气的感叹句;(3)借助语境构成的陈述句加感叹语气的感叹句。李广明也认为感叹句与其他句型的区别虽然在其传达的是感情信息,但这些感情信息的传递往往要借助于理智信息,依附于其他句类^[29]。

本文判断的计算过程中,采用了感叹句兼有表达情感和判断双重任务的观点。如果单纯从独立语句的角度考虑,感叹句是没有表达内容或者没有表达对某一事物的断定。汉语是以语境丰富著称,在特定主题情景中对某一事物的情感并不是凭空产生的,而是依托了一定的思想内容而激发出来的,正所谓触景生情而情动辞发。将感叹句简单定义为没有判断内容或无所是非的句子这种情况只有在感叹句没有语境支持单独出现的时候才得以成立,只要感叹句出现在具体的语境,其隐含的意义中必然有支持其表达思想内容的语义成分存在。比如“我很高兴”这个句子在写作文本中经常出现。如果将这个句子单独拎出来进行分析,它确实没有表现一定的思想内容;但是如果两个好朋友之间因为发生碰撞而相互谅解并且成为好朋友这样的一个语境中来看“我很高兴”这个句子,那么它表达的内容不仅仅是“我”当时的一种“高兴”的情感,而是“双方发生碰撞而相互谅解并且成为好朋友这件事情使得我很高兴”。增添了语境的感叹句无疑就增添了特殊的思想内容,使得感叹句中明显带有了陈述句的判断意味,所以将这类句子归为感叹句是合情合理的。在儿童作文文本分析的过程中,常出现的表达判断的感叹句有:

(1)她小声对我说:“你是不是没带音乐书呢?”我说:“是啊!”——在这一问一答两个句子中,前一个句子是疑问句,不表达判断;但是后一个句子虽然只有两个字“是啊”,但是它表达了作者对自己没有带音乐书这种行为的肯定,所以是一个典型的借助语境构成的陈述句加感叹语气的感叹句。根据上下文语境,这里隐含的命题是:“我没有带音乐书。”

(2)又如“我一边摇手,一边忙着说:不客气,不客气,我们同学之间是应该互相帮助的,谁有困难就应该帮助谁,老师不是经常这样教导我们的吗?”——在这个句子中,明显是个复合判断。但是“老师不是经常这样教导我们的吗?”这个双重否定句其实表达了积极肯定的意思,这个句子中虽然有问号,却表达了肯定的判断。

2. “句子—判断”的计算

(1)“文本到句子”转化的过程分析

在书面表达中,用句号、问号或叹号表示一个句子的结束。从作文文本到句子的转化过程相对比较简单,一般可以通过正则表达式或人工判断的方式来完成写作文本中句子的分离。比如某一写作文本由两个句子构成:

第一个句子为:

课间时,我和我的朋友去操场聊聊天,去散散

步,我还帮老师发本子呢!

第二个句子为:

课间是我的最爱,我要珍惜课间,不要浪费课间。

第一个句子的区分以叹号为标识,第二个句子的区分以句号为标识。

(2)“句子到判断”析取的过程分析

句子到判断的析取过程就是从句子中成功分离出判断的过程。这个过程分析分为两步:

第一步:断定该句子是不是陈述句。陈述句不一定表示判断,不是陈述句一定不表示判断。

第二步:断定该陈述句是不是表示判断的命题,如果是,应该是什么类型的判断。对于举例中的第一个句子而言,首先这个句子表述了一种“去聊聊天、去散散步、帮老师发本子”事实的存在,是一个陈述句;其次这个陈述句中所谈到的事实符合“P且Q”形式,所以这个陈述句表示了一个联言判断。对于举例中的第二个句子而言,这个句子也是一个陈述句,因为出现了“是”,所以是一个典型的直言判断。

对于句子与判断命题的投射分析,同样涉及到分析视角的问题。和词汇分析一样,关注的还是儿童在书面写作中是如何运用语言的,进而剖析儿童运用判断的一些特征。但与词汇计算不一样的地方,词汇有一个相对标准的汉语常用词频表,而句子没有这样的严格的频率区分的研究,句子没有哪类句子用的最多,哪类句子用的最少的说法,判断也没有哪类判断用的最多,哪类判断用的最少,诚然我们主观地推断这种频率现象对于句子而言是存在。所以在句子与判断的计算中,只计算句子的数量和判断的数量即可。

(二)“句群—推理”分析

1.句群与推理关系的辨析

对于什么是句群,界定也是非常复杂的,但是跳过句群的界定研究句群的运用,显然是不合适的。有学者认为句群是由前后衔接、表达中心语义的两个或两个以上句子组成的语言使用单位^{[30][31]},他认为句群具有向心性、衔接法、层次性、完整性和多句调性五个特点。

吴为章和田小琳在1984年《中学教学语法系统提要(试用)》中认为:

“句群是在语义上有逻辑关系,在语法上有结构关系,在语流中衔接连贯的一群句子的组合,是介于句子和段落之间的,或者说是大于句子、小于段落的语言表达单位。在连贯话语中,句群是相对独立的语义—句法单位,它以一定的方式为组合标志,可以从语流中切分出来。”^[32]

此外,他们还认为作为构成连贯话语的一个常

用单位,句群应该具有形式上的前后衔接性、意义上的相对完整性、语气上的不求同一性、风格上的语域一致性、表达上的指向中心性、组合上的先后有序性六个特点^[33]。

庄文中较为详细地归纳了这些组合手段^[34]。他认为句群的组合手段分为语法组合手段、语汇组合手段和修辞组合手段三类,其中语法组合手段通常包括语序组合、关联词组合、句式组合和起关联作用的实词组合四种;语汇组合手段通常使用相同的词语、同义的词语、近义的词语或对应的词语;修辞组合手段常用的有比喻、双关、仿词、设问和反复等。

基于上述的讨论,可以较为清楚地将句群与段落的异同区分开来,从句群的特点中,可以较为准确地从连贯话语中将句群提取出来。

对于句群的分类,和前述词语、句子一样,还是从儿童语言运用的视角进行划分。书面写作是一种言语表达方式,不但体现着儿童对语言文字的运用水平,同时也是儿童在特殊语境中对生活以及社会某种情感的流露和表达。所以,语用视角中的句群划分必然关注句群在一定语境表达中所起到的作用。从语言运用的视角来划分,句群一般分为主体句群、过渡句群以及插入句群三大类^[35]。其中主体句群分为记叙句群、说明句群、议论句群、对话句群、描写句群和抒情句群六类。

在上述的举例中,第一个句子“课间时,我和我的朋友去操场聊聊天,去散散步,我还帮老师发本子呢!”和第二个句子“课间是我的最爱,我要珍惜课间,不要浪费课间。”都不能单独构成句群,两个句子联合起来才能构成一个句群,因为这个句群主要说明人物的课间生活经历,所以是一个记叙句群。

那什么又是推理呢?推理是指从已有的知识推出新结论的过程,它是思维活动的一种重要形式^[36]。推理包含前提和结论两个重要的成份,前提是指已知的知识或推理所依据的知识,结论是指根据前提所得到的新的知识。推理一般分为演绎推理、归纳推理和类比推理。演绎推理是从一般到特殊,归纳推理是从特殊到一般,类比推理是从特殊到特殊。演绎推理有三段论、假言推理和选言推理等形式,归纳推理有完全归纳和不完全归纳等形式,类比推理则有三段演绎法、联言分解法、连锁推导法、综合归纳法、归谬反驳法等形式。

关于句群和推理之间的关系,吴为章等认为推理必然地表现为一定形式的复句或句群,但不是所有的复句或句群都是推理,这就需要识别,看看复句或句群中是否有前提和结论的推理关系,这个时候常常需要借助一定的语言形式标志,比如关联词语来

认定^[37]。一般而言,一个句群至少包含两个句子。有的学者承认推理与概念、判断一样,同语言密切联系在一起,但认为推理的语言形式为表示因果关系的复句或具有因果关系的句群,这样的说法是欠妥当的。因为在连贯的话语中,并不是每个推理都会采用严格的句法关联词语来表示的。推理的语言形式是多种多样的,不仅次序可以变更,项目也可以省略,或省略某个前提,甚至于连结论也可以省略^[38]。所以,从不同类型的句群中识别出推理是非常关键的。

2. “句群—推理”的计算

(1)“文本到句群”转化的过程分析

和词语、句子不一样,句群(尤其是来自儿童作文中的句群)没有什么特殊的判断标志,只能根据连接词和相关短语进行意义判断。从写作文本到句群的转化过程还是比较困难的,最为简单的判断方法以自然段为主进行放大或缩小。句群是构成自然段的基本材料,有时候句群可以独立成自然段,有时候几个句群支持一个自然段。比如以下这篇作文就有 2 个句群构成:

过年的时候,我在家背唐诗。诗里有描写农民,有描写景色,还有描写花草的。我背的很少,到了古诗擂台赛,我输给了同班的同学。我觉的以后一定要多读书,还要多长记性。

第一个句群到“还有描写花草的”结束,第二个句群到“还要多长记性”结束。

(2)“句群到推理”析取的过程分析

句群到推理的析取过程就是从句群中成功分离出推理的过程。推理析取的过程和判断析取的过程是不相同的:对于判断而言,我们明确知道非陈述句一定不表示判断,所以焦点会集中在提取陈述句上;而推理关系则不然,很多的推理并没有明显的关联词语,对于母语而言都是暗含在语义之中的。记叙句群中可能暗含推理,抒情句群中也可能暗含推理,这个复杂关系很难用一一对应的方法进行聚焦,所以从句群到推理的析取过程只能是对写作文本中的句群逐个进行判定才能确定推理的存在。

句群到推理的析取过程也就是上文中所提到的套件显性化的过程。按照套件显性化的四个步骤,句群到推理的析取算法具体如下:

(1)结果型话语

实际上,从儿童叙事型的作文文本中可以看出,一般而言体现推理的文本都在“谈感想”“谈体会”“谈收获”的结果部分。其他部分也有类似的推理,但是相对而言在文本中没有鲜明的突出出来。同时,由于人的语言的存在多义性和同义性,意义逻辑的捕捉本身就是非常复杂的工作,所以本研究在意义逻

辑部分主要捕捉作文文本中的结果型话语,这样便于统一分析思路。

(2)标签提取

第一步先是将作文文本中的标签全部提取出来,形成标签库,如下表所示,标签的提取是按照分词的词性来进行。但是意义逻辑中由于情景的不同,能够起到标志作用的标签词也有所不同。在“和同学之间发生的事情”这个话题情景中,能够起到标志作用的标签主要有名词、时间词和连词。上述例子中,名词标签有:“时”“朋友”“老师”“操场”“本子”;连词标签有“和”;时间词标签有“课间”。但是值得注意的是,在儿童写作文本中,有很多时候并没有明显的标签,这样就得靠一些表示意义的关键性短语来辅助判断。

各种不同类型句群中的连接词标签^[39]

句群类型	标签
条件句群	才、就、不然、不管、这样、只有这样、只要这样、除非如此
因果句群	因为、原来、所以、因此、由于、原因是、是因为、就因为、正因为如此、其结果、由此看来
转折句群	但是、可是、然而、不过、不料、可惜
选择句群	或者、还是、要不(么)、其中、是……还是、要么……要么、或者……或者
总分句群	总之、总而言之、总起来说、总起来看、归根到底、首先……其次……再其次
解证句群	像、如、例如、据说、比如说、举例说、这就是说、可以说、也就是说、意思是说、换句话说
目的句群	为此、为此目的、为的是、省得
假设句群	那、那么、否则、不然、如果这样
承接句群	于是、然后、接着、继而、从此、起先……后来
并列句群	也、又、还、同时、同样、与此同时、此外、相反

(3)套件补足

使用演绎推理、归纳推理和类比推理的推理方式将套件显性化即可。

(4)推理识别

推理识别这一个过程对于人工分析而言相对比较简单,只要按照演绎推理、归纳推理和类比推理的定义进行判定即可。在句群论述时提到的这个文本,表示推理的部分主要集中在第二个句群中。通过上述步骤的判定,第二个句群中主要找出的推理有两个:一个是背诗少输给全班同学的假言推理,一个是要多读书、多长记性的归纳推理。

五、研究结论及实践意义

从写作文本的角度探讨意义逻辑是语言和思维关系讨论深入的表现。本研究认为意义逻辑是将客体的意义同化到特定语言情境中时所产生的逻辑。意义逻辑可以通过套件的显性化和层次化构建计算模型,并通过“句子—判断”“句群—推理”的方法进行计算。意义逻辑及其计算在写作文本分析中具有重要的实践意义,为语言材料中提取出现场思

维状态提供了一种可计算的途径,使得不可观察的思维状态可以用计算的方式来进行理解和解释。比如对于“我说齐婧临抹布不够湿,齐婧临赶快水房头抹布”这个句子而言,“齐婧临”的两次出现犯了语句重复的毛病,同时丢掉了谓语句“到”,最为典型的是还出现了别字“头”。如果单从词汇和语法的角度来分析,这个句子的语病较多,在评价中的得分会相对较低。但是如果从语言的意义和上下文背景来分析,这个句子因果语义关系非常明确。这样在语法上尚显欠缺但能够表达真切语义关系的写作文本,仅依靠微观或宏观的分析是显然不够的。但是如果通过意义逻辑分析,就可以将写作文本中反映出来的因果语义关系非常清晰地表达出来。

受到汉语丰富语义的影响,意义逻辑的研究还处于起步阶段,需有很长的研究过程要经历。意义逻辑需要和宏观的叙事结构研究、微观的词汇研究相结合,才能更为完整的反映写作文本的真实意图,也才能通过写作文本对学习者的内在差异进行更为深入的探索。

参考文献:

- [1][2][6][7] Miller, C. A.. Developmental relationships between language and theory of mind[J]. American Journal of Speech-Language Pathology, 2006,15(2):142-154.
- [3] Gleitman, L.. The structural sources of verb meanings[J]. Language Acquisition, 1990, 1(1):3-55.
- [4][5] 李其维.发生认识论与逻辑——《走向一种意义的逻辑》导读[A].J. 皮亚杰 & R. 加西亚 (Ed). 走向一种意义的逻辑[C].上海:华东师范大学出版社,2005.3-28.
- [8][9] Reiter, R.. A logic for default reasoning[J]. Artificial Intelligence, 1980,(13): 81-132.
- [10] 袁毓林.基于认知的汉语计算语言学研究[M].北京:北京大学出版社,2008.117.
- [11] 黄华新,胡霞.缺省推理:认知语境的功能实现[J].信阳师范学院学报(哲学社会科学版),2004,24(2):1-4.
- [12] 林作铨,石纯一.非单调推理十年进展[J].计算机科学,1990,(6): 31-41.
- [13] 白硕.语言学知识的计算机辅助发现[M].北京:科学出版社,1995.1-2.
- [14] 王维贤,李先昆,陈宗明.语言逻辑引论[M].武汉:湖北教育出版社,1989.32.
- [15][36] 彭聃龄,张必隐.认知心理学[M].杭州:浙江教育出版社,2004.305-320.
- [16] Allwood, J., Andersson, L.-G., & Dahl, O. 语言中的心理学[M].北京:北京大学出版社,1977.
- [17] Scheuer, N., Cruz, M. d. l., Pozo, J. I., Huarte, M. F., & Sola, G.. The mind is not a black box:children's ideas about the writing process[J]. Learning and Instruction, 2006,(16): 72-85.
- [18] 赵兴龙,王冰洁,张俊.技术促进语言运用的五个假设[J].中国电化教育,2011,(4): 13-20.
- [19][20][21] 刘江.逻辑学推理和论证[M].广州:华南理工大学出版社,2004.23-25.
- [22][26] 朱德熙.朱德熙文集第一卷[M].北京:商务印书馆,1999.31.
- [23] 贾可荣,张彦铎.人工智能[M].北京:清华大学出版社,2006.32.
- [24] 陈虎.基于语音库的汉语感叹句与感叹语调研究[J].汉语学习,2007,(5):45-55.
- [25] 吕叔湘.中国文法要略[A].俞晓群和沈昌文(Ed).叔湘全集[C].沈阳:辽宁教育出版社,2002.312.
- [27] 李铁范.感叹句研究综述[J].云南师范大学学报,2005,3(4): 49-56.
- [28] 许爱琼.感叹句与命题的逻辑关系[J].武汉师范学院学报(哲学社会科学版),1984,(4): 109-113.
- [29] 李广明.感叹句及其分类[J].天水师专学报(哲社版),1994,(Z1): 72-78.
- [30] 文文.句群——现代汉语语法教学研究之二[J].兰州教育学院学报,1986,(2): 117-131.
- [31][34] 庄文中.句群[M].北京:人民教育出版社,1990.15-30.
- [32][33][35][37][38] 吴为章,田小琳.汉语句群[M].北京:商务印书馆,2000.18-171.
- [39] 盛玉麒.现代汉语网络课程[DB/OL]. <http://www.yyxx.sdu.edu.cn/chinese/index.htm>,2011-04-07.

作者简介:

赵兴龙:助理研究员,博士,研究方向为教育技术基本理论与实践(zhaoxinglong@moe.edu.cn)。

何克抗:所长,教授,研究方向为教育技术基本理论与实践(hekkbnu@163.com)。

收稿日期 2012年10月12日
责任编辑 宋灵青