

doi:10.14132/j.cnki.1673-5439.2020.06.011

基于中介真值程度度量的模糊语义翻译研究

何霞¹, 杜国平², 宗慧³

- 1. 中国社会科学院研究生院, 北京 102488
- 2. 中国社会科学院哲学所, 北京 100732
- 3. 南京航空航天大学 计算机科学与技术学院, 江苏 南京 211106

摘要:自然语言的语义具有与生俱来的模糊性,这增添了人们相互理解的困难;特别在使用不同语言的人群之间的交流中,这更易产生误解。中介逻辑揭示出语义模糊性的本质体现为语义的中介状态,中介真值程度的度量(MMTD)则从量的侧面度量这种中介状态。从词性上来看,语义的模糊性主要体现在形容词、动词和副词这三种词性上。文中运用MMTD,并根据各词性的独立性建立量化其模糊语义程度的度量函数,并在此基础上,提出叶词、词链接等概念,以建立量化语句模糊语义程度的方法,并将其运用于英、汉语句的互译。

关键词:模糊语义;中介逻辑;MMTD;词类;语句翻译

中图分类号: B81; H059; TP391.1 文献标志码: A 文章编号: 1673-5439(2020)06-0071-07

Research on fuzzy semantic translation based on intermediate truth degree measurement

HE Xia¹, DU Guoping², ZONG Hui³

- 1. Graduate School of Chinese Academy of Social Sciences, Beijing 102488, China
- 2. Department of Philosophy, Chinese Academy of Social Sciences, Beijing 100732, China
- 3. College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

Abstract: The semantics of natural language are inherently ambiguous, increasing the difficulty of mutual understanding. It is more likely to cause misunderstanding particularly in the communication between people using different languages. The intermediary logic reveals that the essence of semantic ambiguity is the intermediary state of semantics, and the measuring of medium truth degree (MMTD) is the tool for quantifying the intermediary state. From the perspective of part of speech, the ambiguity of semantics is mainly reflected in three parts of speech: adjectives, verbs and adverbs. A MMTD is used and a metric function is established for quantifying the degree of fuzzy semantics according to the independence of each part of speech. On this basis, these concepts, such as leaf words and word links, etc. are proposed and a method is established for quantifying the degrees of fuzzy semantics of sentences and is applied to the translation of English and Chinese sentences.

Keywords: fuzzy semantics; intermediary logic; measuring of medium truth degree (MMTD); part of speech; sentence translation

语言是人们交流思想的工具。自然语言与生俱来的语义模糊性使得交流双方在语义理解上容易产生歧义,甚至误解,从而限制了准确的交流。人们使用不同语言交流时,需要经过翻译的过程,这使得语义理解更具模糊性。这自然地引起了语言学、人工智能、逻辑学等众多学科学者们的重视,并对自然语言的模糊性、互译的准确性等方面进行了持久而深入的研究。

1965年,Zadeh提出用“模糊集合”作为表示模糊语义的数学模型,通过隶属度方法对模糊语义进行定量描述^[1]。中国学者对模糊语义进行了语言学的描述,文献[2]认为模糊语言是指表达事物类属边界或性质的亦此亦彼,即中介过渡性的词语;文献[3]指出模糊语义的基本特征是其表达的概念外延边界的模糊性与中心的明晰性;文献[4]指出语言的模糊性主要出现在语义连续统的两端的边缘部分;文献[5]认为语义的模糊性主要体现在词语的模糊义素上。在此基础上,一些学者还分别运用模糊数学的方法^[2,6]、模糊语义模糊度的测量方法^[5],以及模糊语义的隶属度方法^[7]对自然语言语义的模糊性进行了定量分析;文献[8-9]采用中介逻辑中中介真值程度的度量工具对中文副词的模糊语义进行度量;文献[10]采用中介逻辑中中介真值程度的度量工具对英文副词的模糊语义进行度量。在翻译方面,文献[11]指出,语言的互译性说明了自然语言中的模糊性恰好等于任何其他自然语言中的模糊性;由此,文献[12-13]分别利用模糊数学的方法^[12]、信息熵^[13]的方法,从定量的角度研究了语言的模糊语义的互译情况。在人工智能方面,文献[14]以中智词的名称分析了最模糊的自然语言话语,并根据中智理论和熵测度的概念开发了一个语义表示框架;文献[15]特别针对语义概念的集成,以更高的准确性和精度表示可视内容;文献[16]评估从斯洛伐克语合成语言到分析英语的机器翻译输出的质量时,使用所有错误率和准确性的自动测量方法。

词是语言的基本单位,语言的模糊性显著地体现在词的语义上。正如 Russel 所言^[17],除了逻辑与纯粹数学,具有准确意义的字是没有的。本文采用中介逻辑的模糊语义量化工具中介真值程度的度量(Measuring of Medium Truth Degree, MMTD)对汉英词及其链接的模糊语义进行量化研究,并根据量化结果应用到词及语句的翻译方法进行了讨论。

1 中介逻辑与 MMTD 简介

1.1 中介逻辑简介

1984年,朱梧楨和肖奚安认为,在认识论中有一条重要原则:对立面的相互转化过程中总有中介过渡现象,它是同一性在质变过程中的集中表现。他们指出:“既然对立面的相互转化普遍存在,则中介状态也必然存在。那么从量的侧面去研究中介对象或同一性质在质变过程中的集中表现就是不可避免的了”^[18]。在此中介原则下,他们创建了中介逻辑(Medium Logic, ML)^[18]。该逻辑以 Aristotle 开创的形式逻辑中的“矛盾”和“对立”概念作为哲学背景。若两个概念中的一个内涵否定另一个内涵,则称它们是一对矛盾概念,如:好与不好、正常与非正常;若两个概念都有其自身的肯定内容,并在同一内涵的一个更高级的概念中,它们之间存在着最大差异,则此两个概念就是一对对立概念,如:好与坏、大与小等。然而,自然界中的所有事物并不都是“非此即彼”,也存在(而且是大量地存在)“非此非彼”的现象,如黎明既不是黑夜也不是白昼,这“非此非彼”正是由此至彼或从彼至此的中介过渡。在真假二值的基础上,中介逻辑引入了第三值,该值是基于对对立面的中介状态的肯定,记为 M ,这样,中介逻辑的真值的集合 $truth = \{T, M, F\}$,它反映了对立物在转化过程中的中介状态,其语义深刻地反映了模糊性。中介逻辑的基本联结词为 \neg , \sim 和 $<$, 其真值表如表 1 所示。

表 1 ML 的连接词真值表

A	B	$\neg A$	$\sim A$	$A < B$
T	T	F	M	T
T	M	F	M	M
T	F	F	M	F
M	T	M	T	T
M	M	M	T	T
M	F	M	T	M
F	T	T	M	T
F	M	T	M	T
F	F	T	M	T

记 P 为一谓词(概念或性质), x 为任一变元, $P(x)$ 表示变元 x 完全具有性质 P 。符号“ \neg ”称对立否定词,读为“对立”,把谓词 P 的对立面记为 $\neg P$,则 P 与 $\neg P$ 就抽象地表示了一对对立概念;符号“ \sim ”称模糊否定词,解释并读作“部分地”,反映了对立物在转化过程中的中介状态,其语义深刻地反映了

模糊性;“<”是真值程度词,例如“ $x < y$ ”表示“ x 真值程度不强于 y ”。对于 P 和 $\neg P$,若变元 x 满足 $\sim P(x) \& \sim \neg P(x)$,则称 x 为 P 与 $\neg P$ 的中介对象,此处 $\&$ 为自然语言中的“并且”或“与”的简记。

1.2 中介真值程度的度量(MMTD)

MMTD 是基于中介逻辑对模糊现象进行定量计算的一种方法^[19]。记 X 是非空对象集合, f 是 X 的一维数值化映射。如图 1 所示,将一般数值化应用的数值区域划分为对应谓词真值的 3 个区域,即 $\neg P(x)$ 、 $\sim P(x)$ 、 $P(x)$ 。在对应 $P(x)$ 的“真”数值区域, α_T 是关于 $P(x)$ 的 T 标准度, $[\alpha_T - \varepsilon_T, \alpha_T + \varepsilon_T]$ 是 α_T 的邻域;在对应 $P(x)$ 的“假”数值区域, α_F 是 $\neg P$ 的 F 标准度, $[\alpha_F - \varepsilon_F, \alpha_F + \varepsilon_F]$ 是 α_F 的邻域。

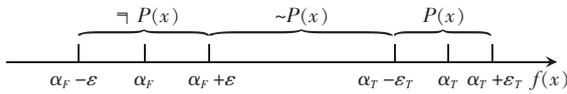


图 1 数值区域与谓词的对应关系

个体的真值程度可以通过计算距离比率函数 $h_T(f(x))$ 或 $h_F(f(x))$ 得到。采用 MMTD 方法处理模糊现象时,寻求 f 至关重要,它既可以采用已有知识,也可以根据应用的特点自主建立。

2 形容词和动词的模糊语义度量

模糊语义学研究的对象是语义中客观存在的大量模糊现象,主要体现在对词义的研究上^[6]。文中对形容词、动词和副词的模糊语义进行度量。

2.1 形容词模糊语义的度量

形容词用于表示事物形状、性质、状态等,是典型的具有模糊语义的一类词,例如,“可爱”的标准是什么,“年轻”的年龄范围是什么,都无法给出一个明晰的答案。

对某类形容词的模糊语义进行度量的依据和标准是该类形容词自身的“褒”和“贬”的语义区别。这时“褒”与“贬”之间存在非褒非贬的、中介过渡的形容词。需要说明的是,本文只是将形容词谓词定义为“褒”和“贬”,这不仅限于褒义形容词和贬形容

表 2 中文喜怒类形容词的情绪程度赋值表

x	震怒	愤怒	恼怒	生气	难过	不快	愉快	高兴	喜悦	大喜	狂喜
$f_{adj}(x)$	0	0.10	0.20	0.30	0.35	0.45	0.70	0.75	0.90	0.95	1.00

词,该谓词还包括形容词的正向和反向概念,譬如,“胖”和“瘦”是一对反义形容词,他们没有褒贬之分,但是有正反向之分,“胖”比“瘦”呈现的面积更大,所以,将“胖”归类为正向形容词,“瘦”归类为反向形容词。依照“褒义”对应“正向”,“贬义”对应“反向”,“胖”和“瘦”分别属于谓词“褒”和“贬”。下面建立基于 MMTD 的形容词模糊语义的度量方法。

首先采用逻辑真值对某类形容词的模糊语义定性。设形容词一元谓词 $C(x)$ 表示“ x 是褒义的”。如果 $C(x)$ 为 T ,那么 $\neg C(x)$ 为 F ,即“ x 是贬义的”; $\sim C(x)$ 为 M ,即“ x 是非褒非贬的”。

再采用函数对某类形容词的模糊语义定量分析。设 $f_{adj}(x)$ 为形容词模糊语义的度量表示,它的值域为 $[0, 1]$ 。那么形容词一元谓词的合式公式与 $f_a(x)$ 的对应关系如图 2 所示。

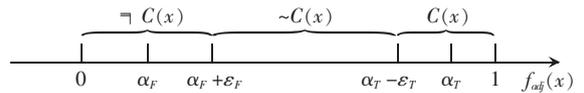


图 2 形容词谓词与数值区域对应关系

最后,根据图 2,建立形容词的 MMTD 函数。

定义 1 给定某类形容词的集合 A 。对于任意 $x \in A$,如果存在 $f_{adj}(x)$,那么相对于谓词 C 的真值程度度量函数

$$h_T(f_{adj}(x)) = \begin{cases} 0, & \neg C \\ \frac{d(f_{adj}(x), \alpha_F + \varepsilon_F)}{d(\alpha_F - \varepsilon_F, \alpha_F + \varepsilon_F)}, & \sim C \\ 1, & C \end{cases} \quad (1)$$

其中, d 是欧式距离,即

$$d(a, b) = \left[\sum_{i=1}^n (a_i - b_i)^2 \right]^{1/2}$$

在一维情况下, $d(a, b) = |a - b|$ 。

例 1 根据语义,设中文喜怒类形容词的有序集合: $A1 = \{震怒, 愤怒, 恼怒, 生气, 难过, 不快, 愉快, 高兴, 喜悦, 大喜, 狂喜\}$ 且喜怒程度如表 2 所示。

值程度。例如, $h_T(f_{adj}(\text{愤怒})) = 0$; $h_T(f_{adj}(\text{难过})) = 0.21$; $h_T(f_{adj}(\text{愉快})) = 0.71$; $h_T(f_{adj}(\text{大喜})) = 1$ 。

例 2 根据语义,设英文喜怒类形容词的有序

集合 $A_2 = \{ \text{furious, enraged, mad, irritated, resentful, angry, annoyed, upset, sad, unhappy, joyless, moody,} \}$

$\text{pleased, happy, joyful, overjoyed, thrilled} \}$, 且喜怒程度如表 3 所示。

表 3 英文喜怒类形容词的情绪程度赋值表

x	furious	enraged	mad	irritated	resentful	angry	annoyed	upset	sad
$f_{adj}(x)$	0	0.04	0.08	0.12	0.16	0.20	0.24	0.28	0.30
x	unhappy	joyless	moody	pleased	happy	joyful	overjoyed	trilled	
$f_{adj}(x)$	0.32	0.38	0.42	0.60	0.70	0.80	0.90	1.00	

设与 $\neg C(x)$ 对应的数值区域是 $[0, 0.08]$, 则 $\alpha_F = 0.04, \varepsilon_F = 0.04$; 与 $C(x)$ 对应的数值区域是 $[0.9, 1]$, 则 $\alpha_T = 0.95, \varepsilon_T = 0.05$ 。根据式(1), 可以计算出表 3 中任一形容词对于 $C(x)$ 的真值程度。如: $h_T(f_{adj}(\text{mad})) = 0$; $h_T(f_{adj}(\text{upset})) = 0.24$; $h_T(f_{adj}(\text{pleased})) = 0.63$; $h_T(f_{adj}(\text{thrilled})) = 1$ 。

2.2 动词的模糊语义真值程度度量

动词是陈述人或事物的动作、行为、心理活动或存现的词。以表示动作的动词为例, 由于动作行为之间难以截然划界, 往往导致动词的模糊性。例如, 现代汉语里“跑”: 两只脚或四条腿迅速前进。“迅速”本身是模糊的, 两脚向前移动的频率多大才算迅速, 这恐怕难以精确界定。

对表示动作范畴的动词的模糊语义进行度量的依据和标准是该类动词所代表的动作范畴大小的语义区别。需要说明的是, 本文只是将动词谓词定义为“大”和“小”, 是统一地对动作范畴的“大”, 包括动作引起的速度快慢, 形成的距离远近, 力度的轻重等正反对立概念所作的描述。使用“大”这一谓词来代表动作产生的速度快, 力度大, 距离远等的动

词, 使用“小”这一谓词来代表动作产生的速度慢, 力度小, 距离近等的动词, 这时“大”与“小”之间存在非大非小的、中介过渡的动词。设动词一元谓词 $B(x)$ 表示“ x 是大的”。如果 $B(x)$ 为 T , 那么 $\neg B(x)$ 为 F , 即“ x 是小的”; $\sim B(x)$ 为 M , 即“ x 是非大非小的”。

定义 2 给定某类动词的集合 V 。对于任意 $x \in V$, 如果存在 $f_v(x)$, 那么相对于谓词 B 的真值程度度量函数

$$h_T(f_v(x)) = \begin{cases} 0, & \neg B \\ \frac{d(f_v(x), \alpha_F + \varepsilon_F)}{d(\alpha_F - \varepsilon_F, \alpha_F + \varepsilon_F)}, & \sim B \\ 1, & B \end{cases} \quad (2)$$

例 3 根据语义, 设中文走跑类动词的有序集合 $V_1 = \{ \text{踱, 走, 阔步, 疾行, 奔走, 奔, 小跑, 跑, 奔跑, 飞奔, 飞跑, 狂奔, 冲} \}$ 。以速度的快慢程度建立该类动词模糊语义的线性度量表示: $f_v(x) = \text{nord}/(n-1)$ 。其中 nord 是序数, 它表示 V_1 中元素的次序, $0 \leq \text{nord} \leq n$ 。经计算, $f_v(x)$ 之值如表 4 所示。

表 4 中文走跑类动词的情绪程度赋值表

x	踱	走	阔步	疾行	奔走	奔	小跑	跑	奔跑	飞奔	飞跑	狂奔	冲
$f_v(x)$	0	1/12	2/12	3/12	4/12	5/12	6/12	7/12	8/12	9/12	10/12	11/12	12/12

设与 $\neg B(x)$ 对应的数值区域是 $[0, 2/12]$, 则 $\alpha_F = 1/12, \varepsilon_F = 1/12$; 与 $B(x)$ 对应的数值区域是 $[10/12, 1]$, 则 $\alpha_T = 11/12, \varepsilon_T = 1/12$ 。根据式(2), 可以计算出表 4 中任一动词对于 $B(x)$ 的真值程度。例如, $h_T(f_v(\text{阔步})) = 0$; $h_T(f_v(\text{奔})) = 0.38$; $h_T(f_v(\text{奔跑})) = 0.5$; $h_T(f_v(\text{狂奔})) = 1$ 。

例 4 根据语义, 设英文走跑类动词的有序集合 $V_2 = \{ \text{pace, walk, stride, rush, trot, streak, scamper, run, scurry, dash} \}$ 。同样以速度的快慢程度建立该类动词模糊语义的线性度量表示, 经计算, $f_v(x)$ 值如表 5 所示。

表 5 英文走跑类动词的速度赋值表

x	pace	walk	stride	rush	trot	streak	scamper	run	scurry	dash
$f_v(x)$	0	1/9	2/9	3/9	4/9	5/9	6/9	7/9	8/9	9/9

设与 $\neg B(x)$ 对应的数值区域是 $[0, 2/9]$, 则 $\alpha_F = 1/9, \varepsilon_F = 1/9$; 与 $B(x)$ 对应的数值区域是

$[7/9, 1]$, 则 $\alpha_T = 8/9, \varepsilon_T = 1/9$ 。根据式(2), 可以计算出表 5 中任一动词对于 $B(x)$ 的真值程度。例

如, $h_T(f_v(\text{stride})) = 0$; $h_T(f_v(\text{rush})) = 0.2$;
 $h_T(f_v(\text{scamper})) = 0.8$; $h_T(f_v(\text{scurry})) = 1$ 。

3 副词模糊语义的度量

作为独立的词类,副词主要用于限制、修饰动词和形容词性词语,表示程度、范围、时间等意义。副词的模糊性主要表现在两个方面:一是本身语义的模糊性,比如,“经常”、“有点”。二是修饰别的词语一起构成模糊概念,如“经常小跑”,“有点高兴”,等等。副词发生作用之后,使得被修饰词的语义增强或减弱。使用谓词 S 表示“增强”,那么副词一元谓词 $S(x)$ 表示“ x 是强的”, $\neg S(x)$ 表示“ x 是弱的”, $\sim S(x)$ 表示“ x 是非强非弱的”。

由于在值域为 $[0, 1]$ 的 f 中, $f(x) = 0.5$ 表明 x 是最模糊的,以副词看,这是语义减弱和语义增强的临界点和分水岭。设 $f_{adv}(x)$ 为副词模糊语义的度量表示,那么它与副词一元谓词的合式公式的对应关系如图 3 所示。

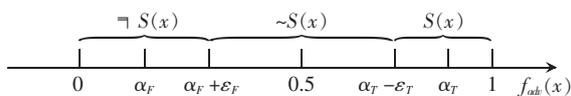


图 3 副词谓词与数值区域对应关系

表 6 中文频度副词的频度赋值表

x	从不	鲜少	偶尔	有时	#	经常	通常	总是
$f_{adv}(x)$	0	0.15	0.30	0.45	0.50	0.75	0.90	1.00

设与 $\neg S(x)$ 对应的数值区域是 $[0, 0.15]$, 则 $\alpha_F = 0.075$, $\epsilon_F = 0.075$; 与 $S(x)$ 对应的数值区域是 $[0.9, 1]$, 则 $\alpha_T = 0.95$, $\epsilon_T = 0.05$ 。根据式(3), 可以计算出表 6 中任一副词对于 $S(x)$ 的真值程度。例如, $h_T(f_{adv}(\text{鲜少})) = 0$; $h_T(f_{adv}(\text{有时})) = 0.8$; $h_T(f_{adv}(\text{经常})) = 1.6$; $h_T(f_{adv}(\text{通常})) = 1$ 。

例 6 根据语义, 设英文频度副词的有序集合 $Adv_{F2} = \{\text{never, seldom, occasionally, sometimes, often,}$

表 7 英文频度副词的频度赋值表

x	never	seldom	occasionally	sometimes	#	often	usually	frequently	generally	always
$f_{adv}(x)$	0	0.05	0.20	0.45	0.50	0.70	0.80	0.85	0.90	1.00

设与 $\neg S(x)$ 对应的数值区域是 $[0, 0.05]$, 则 $\alpha_F = 0.025$, $\epsilon_F = 0.025$; 与 $S(x)$ 对应的数值区域是 $[0.9, 1]$, 则 $\alpha_T = 0.95$, $\epsilon_T = 0.05$ 。根据式(3), 可以计算出 $h_T(f_{adv}(\text{seldom})) = 0$; $h_T(f_{adv}(\text{occasionally})) = 0.35$; $h_T(f_{adv}(\text{frequently})) = 1.88$; $h_T(f_{adv}(\text{generally})) = 2$, 等等。

定理 1 设某类副词的集合 $Adv = \{x_1, x_2, \dots, x_i, \#, x_{i+2}, \dots, x_n\}$, 且 $f(\#) = 0.5$, 则存在 $\sim S(\#)$ 。

与形容词、动词不同,副词属于二次修饰的词汇。建立其真值程度度量函数时需要注意到修饰词与被修饰词链接后的合理性和副词的独立性。

定义 3 给定某类副词的集合 Adv 。对于任意 $x \in Adv$, 如果存在 $f_{adv}(x)$, 那么相对于谓词 S 的真值程度度量函数

$$h_T(f_{adv}(x)) = \begin{cases} 0, & \neg S(x) \\ \frac{2 * d(f_{adv}(x), \alpha_F + \epsilon_F)}{d(\alpha_F - \epsilon_F, \alpha_F + \epsilon_F)}, & \sim S(x) \\ 1, & \sim S(\#) \\ 2, & S(x) \end{cases} \quad (3)$$

例 5 根据文献[20]及词典释义, 设中文频度副词的有序集合 $Adv_{F1} = \{\text{从不, 鲜少, 偶尔, 有时, 经常, 通常, 总是}\}$, 频度程度如表 6 所示。“从不, 鲜少, 偶尔, 有时”修饰动词后使其语义减弱, 属于 $[0, 0.5)$ 区间, “从不”为 0, # 为空缺位, 表示无修饰无语义变化; “经常, 通常, 总是”修饰动词后使其语义增强, 属于 $(0.5, 1]$ 区间, 0.5 为空缺位, 即表示无修饰无语义变化, “总是”为 1。

usually, frequently, generally, always}, 频度程度如表 7 所示。根据语义, “never, seldom, occasionally, sometimes”为修饰后使语义减弱的副词, 属于 $[0, 0.5)$ 这个区间, “never”为 0, # 为空缺位, 表示无修饰无语义变化; “often, usually, frequently, generally, always”为修饰动词后使其语义增强的副词, 属于 $(0.5, 1]$ 这个区间, 0.5 为空缺位, 表示无修饰无语义变化, “always”为 1。

例 7 根据文献[21]及词典释义, 设中文程度副词的有序集合: $Adv_{D1} = \{\text{绝不, 不太, 有点, 很, 非常, 特别, 极其}\}$, 其程度如表 8 所示。

表 8 中文程度副词的程度表

x	绝不	不太	有点	#	很	非常	特别	极其
$f_{adv}(x)$	0	0.10	0.30	0.50	0.75	0.85	0.95	1.00

设与 $\neg S(x)$ 对应的数值区域是 $[0, 0.1]$, 则 $\alpha_F = 0.005, \varepsilon_F = 0.005$; 与 $S(x)$ 对应的数值区域是 $[0.95, 1]$, 则 $\alpha_T = 0.975, \varepsilon_T = 0.025$ 。那么可以得到 $h_T(f_{adv}(\text{不太})) = 0; h_T(f_{adv}(\text{有点})) = 0.47$;

表9 英文程度副词的程度表

x	not at all	hardly	nearly	a little	somewhat	#	rather	very	extremely	absolutely
$f_{adv}(x)$	0	0.05	0.20	0.30	0.45	0.50	0.75	0.85	0.95	1.00

设与 $\neg S(x)$ 对应的数值区域是 $[0, 0.05]$, 则 $\alpha_F = 0.025, \varepsilon_F = 0.025$; 与 $S(x)$ 对应的数值区域是 $[0.95, 1]$, 则 $\alpha_T = 0.975, \varepsilon_T = 0.025$ 。根据式(3), 可以计算出表9中任一副词对于 $S(x)$ 的真值程度。例如, $h_T(f_{adv}(\text{hardly})) = 0; h_T(f_{adv}(\text{somewhat})) = 0.89; h_T(f_{adv}(\text{very})) = 1.78; h_T(f_{adv}(\text{absolutely})) = 2$ 。

4 词链接与语句翻译

4.1 词链接

副词用于修饰形容词和动词; 形容词主要修饰名词或代词, 及物动词指向由名词或代词为主要成分的宾语。因此, 副词、形容词、名词或代词, 副词、动词、名词或代词就形成了含有模糊语义的有序链接。显见, 名词或代词、动词、形容词均可能是被修饰的终极对象。称被修饰的终极对象为叶词。例如, 非常帅气的小伙, 努力学习数学, 经常小跑, 有点难过。这里的“小伙”、“数学”、“小跑”、“难过”就是叶词。

定义4 设词的集合 $W = \{w_i | 0 \leq i \leq n\}$, 则 $w_n * w_{n-1} * \dots * w_1 * w_0$ 记作词链接。这里 $*$ 是链接符; w_0 是叶词。

定义5 设词的集合 $W = \{w_i | 0 \leq i \leq n\}$, 且存在 $f_i(w_i)$ 。又设当 w_0 不是非及物动词和形容词时, $h_T(f_0(w_0)) = 1$, 那么词链接的真值程度度量函数

$$h_{T-mix}(W) = \prod_{i=0}^n h_T(f_i(w_i)) \quad (4)$$

根据定义4, 副词*形容词*名词(代词)、副词*及物动词*名词(代词)和副词*不及物动词是三种不同的含有副词的词链接。根据定义5, 设 $x \in Adv, y \in A, z \in V$, 含有副词的词链接的真值程度度量函数

$$h_{T-mix}(x, y, z) = \begin{cases} h_T(f_{adv}(x)) \times h_T(f_{adj}(y)) \times h_T(f_0(w_0)), & \text{形容词} \\ h_T(f_{adv}(x)) \times h_T(f_v(z)) \times h_T(f_0(w_0)), & \text{及物动词} \\ h_T(f_{adv}(x)) \times h_T(f_v(z)), & \text{不及物动词} \end{cases} \quad (5)$$

$h_T(f_{adv}(\text{很})) = 1.63; h_T(f_{adv}(\text{特别})) = 2$, 等等。

例8 根据文献[22], 设英文程度副词的有序集合 $Adv_{D2} = \{\text{not at all, hardly, nearly, a little, somewhat, rather, very, extremely, absolutely}\}$, 其程度如表9所示。

4.2 基于词链接的语句翻译

例9 环湖跑道上, 有人早晨经常小跑。

$x = \text{经常}, z = \text{小跑}$, 根据例3和例5, 并依据式(5), 本句的模糊语义真值程度 $h_{T-mix}(\text{经常}, \text{小跑}) = h_T(f_{adv}(\text{经常})) \times h_T(f_v(\text{小跑})) = 1.6 \times 0.5 = 0.8$ 。

从翻译标准“信”的角度, 在英文中寻求语义真值程度和中文最接近的词语, 根据上文研究, “经常”的语义和“often”最接近; “小跑”的语义和“trot”最接近。即: $h_{T-mix}(\text{usually, trot}) = h_T(f_{adv}(\text{usually})) \times h_T(f_v(\text{trot})) = 1.76 \times 0.45 = 0.79$ 。

因此, 此句的翻译为:

On the runway around the lake, people usually trot in the morning.

例10 一天的努力白费了, 李强不禁有点难过。

根据定义4, $x = \text{有点}, y = \text{难过}$, 本句的模糊语义真值程度为:

$$h_{T-mix}(\text{有点, 难过}) = h_T(f_{adv}(\text{有点})) \times h_T(f_{adj}(\text{难过})) = 0.24 \times 0.2 = 0.048$$

根据上文研究, “有点”的语义和“a little”最接近; “难过”的语义和“sad”最接近。即: $h_{T-mix}(\text{a little, sad}) = h_T(f_{adv}(\text{a little})) \times h_T(f_{adj}(\text{sad})) = 0.25 \times 0.22 = 0.055$ 。

因此, 此句的翻译为:

One day's effort wasted in vain. Li Qiang couldn't help but feel a little sad.

5 结束语

文中从中介逻辑角度定性分析了形容词、动词、副词的模糊性, 并采用MMTD建立了定量分析它们的模糊语义的真值程度函数。在此基础上, 建立词链接概念, 以量化语句的模糊语义, 并通过实例对中、英文语句的模糊语义进行了互译。文中第一次提出采用逻辑真值程度对词和语句的模糊语义进行翻译, 这对于人工智能, 机器翻译, 翻译评价都有启示作用。

参考文献:

- [1] ZADEH L A. Quantitative fuzzy semantics[J]. *Information Sciences*, 1971, 3(2): 159 - 176.
- [2] 苗东升. 模糊学导引[M]. 北京:中国人民大学出版社, 1987.
- [3] 黎千驹. 模糊修辞学导论[M]. 北京:光明日报出版社, 2006.
- [4] 伍铁平. 模糊语言学[M]. 上海:上海外语教育出版社, 1999.
- [5] 石安石. 语义论[M]. 北京:商务印书馆, 2016.
- [6] 张琪. 自然语言模糊语义的量化与实现[D]. 南京:南京邮电大学, 2008.
ZHANG Qi. The quantification and realization of fuzzy semantics of natural language[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2008. (in Chinese)
- [7] 张乔. 模糊语义学[M]. 北京:中国社科出版社, 2015.
- [8] 谢丁. 基于MMTD的模糊语义量化研究[D]. 南京:南京邮电大学, 2012.
XIE Ding. Research on fuzzy semantic quantization based on MMTD[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2012. (in Chinese)
- [9] 谢丁, 洪龙. 基于MMTD的中文副词模糊语义的量化研究[J]. *南京邮电大学学报(自然科学版)*, 2012, 32(4): 48 - 52.
XIE Ding, HONG Long. Quantification research on the fuzzy semantics of Chinese adverbs based on MMTD[J]. *Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition)*, 2012, 32(4): 48 - 52. (in Chinese)
- [10] 张晶颜. 基于MMTD的英文陈述句模糊语义的量化研究[D]. 南京:南京邮电大学, 2014.
ZHANG Jingyan. Quantitative research on fuzzy semantics of English declarative sentences based on MMTD[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2014. (in Chinese)
- [11] ROY A S. Handbook of the Philosophy of Science, Philosophy of Logic [M]. Amsterdam: North-Holland, 2007: 155 - 171.
- [12] 范守义. 模糊数学与译文评价[J]. *中国翻译*, 1987, 8(4): 2 - 9.
- [13] 邵璐. 文学中的模糊语言与翻译:以《达·芬奇密码》中英文本比较研究为例 [M]. 北京:商务印书馆, 2011.
SHAO Lu. Fuzzy Language in Literature and Translation—A Comparative Study the Da Vinci Code and Its Chinese Versions [M]. Beijing: The Commercial Press, 2011. (in Chinese)
- [14] COLHON M, SMARANDACHE F, VOINEA D V. Entropy of polysemantic words for the same part of speech [J]. *IEEE Access*, 2020, 8: 2975 - 2982.
- [15] BELKHATIR M. Fuzzy-logic-based integration of web contextual linguistic structures for enriching conceptual visual representations [J]. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2019, 3(4): 351 - 356.
- [16] DASA M, PETR H, MICHAL M, et al. Evaluation of machine translation quality through the metrics of error rate and accuracy [J]. *Procedia Computer Science*, 2020, 171: 1327 - 1336.
- [17] RUSSELL B. Human Knowledge: Its Scope and Limits [M]. Oxon: Routledge, 2009.
- [18] 朱梧槿, 肖奚安. 数学基础概论[M]. 南京:南京大学出版社, 1996.
- [19] 洪龙, 肖奚安, 朱梧槿. 中介真值程度的度量及其应用(I) [J]. *计算机学报*, 2006, 29(12): 2186 - 2193.
HONG Long, XIAO Xi'an, ZHU Wujia. Measure of medium truth scale and its application (I) [J]. *Chinese Journal of Computers*, 2006, 29(12): 2186 - 2193. (in Chinese)
- [20] 王逢鑫. 英语意念语法[M]. 北京:北京大学出版社, 1989.
- [21] 张志公. 现代汉语:上册(试用本) [M]. 北京:人民教育出版社, 1982.
- [22] 徐盛桓. 论荷恩的等级关系——新格赖斯会话含意理论系列研究之十 [J]. *外国语(上海外国语大学学报)*, 1995, 18(1): 11 - 17.
XU Shenghuan. Notes on Horn scale and other papers [J]. *Journal of Foreign Languages*, 1995, 18(1): 11 - 17. (in Chinese)