

模糊语词的概率解释

邢鲲

摘要: 模糊语词在日常语言中随处可见, 人们使用它们也没有什么障碍。然而, 如何解释模糊语词却给学者们带来不小的困难, 迄今为止未能达成共识。当前的主要理论如认知主义、超赋值主义、多值逻辑(真值度)和语境主义等都面临各自的问题。在评述这些理论的基础上, 本文提出一种多值逻辑和概率方法的融合理论。该理论能克服多值逻辑和其它理论的主要缺点, 符合人们的语言常识和语言习惯, 并能很好地契合今天的人工智能应用。

关键词: 模糊语词; 不确定性; 多值逻辑; 概率; 人工智能

中图分类号: B81 **文献标识码:** A

1 引言

古典逻辑中, 每个语词(谓词)都是精确的(**precise**): 它的含义是一集个体或个体序列; 所有论域中的个体或个体序列, 要么在该集合中, 要么不在。([17], 24-32 页) 日常语言中, 很多语词的含义是模糊的/含混的(**vague**): 某些个体或个体序列是否在该语词所解释的集合中是不确定的。比如说, “他长得很高”中的“很高”是模糊的, 因为 1.8 米的人算不算很高在很多语境下是不完全确定的。这样的对象被称为该语词的边界情况(**borderline cases**)。可以说, 只要存在这样的边界情况, 某个语词的含义就是模糊的。¹ 依此理解, 容易确立如下事实:

事实 1. 模糊语词或语词的模糊性在日常语言中是一种普遍现象。

除了高、矮、胖、瘦、好、坏、贫、富等公认的模糊语词, 日常语言中其它语词也多多少少具有模糊性。([23, 26]) 实际上, 哪怕被认为最清楚的专名, 也并非具有绝对精确的含义。比如说, “上海”指的是一个地理区域, 在地图上能画出这

收稿日期: 2025-06-24

作者信息: 邢鲲 安阳师范学院马克思主义学院
xingkundeyx@gmail.com

致谢: 非常感谢审稿人对本论文的评审!

¹模糊语词缺少统一的定义, 存在边界情况是公认的最为清楚的特征。学者们还讨论过其它特征如外延没有预先定义好的清晰边界、容忍一定程度的变化、会导致连锁悖论等。然而, 这些特征或者本身不够清晰, 或者是边界情况存在所导致的结果。因此本文不将其视为定义的一部分。更多讨论参见 [27] 或 [17] 第 3 章。

个区域的边界。可是，现实中真能精确确定其每一处边界吗？大概做不到。而且，现实的边界总是有宽度的，那么在边界线的中心，究竟是不是属于上海呢？更不用说，在“上海下雨了”这种表述中，“上海”很可能不是指整个上海区域，也许指的只是其不知道精确边界的某部分。再比如，“小明”指称某个人，那这个人究竟什么时候诞生，什么时候消失呢？每个人都是从受精卵逐步发育成人形的胎儿然后离开母体，这个过程中小明什么时候从无到有诞生了？同时，每个人都会走向生命的终结，变成没有生命的尸体，然后变成尘土。这个过程中哪个时刻小明不存在了呢？能划出精确的时间边界吗？恐怕没有人能做到。因此，上海或小明这样的专名所表达的概念同样存在边界情况，即具有模糊性。²

事实 2. 除非不确定性被消除，否则语词的模糊性不能被彻底消除。

语言使用的目的是为了表达人们的认知结果。如果认知中存在不确定性，却只允许使用精确语词来表达，那么很多认知结果就无法表达。彻底消除模糊语词或语词的模糊性，意味着每句陈述都必须是完全精确的，这在充满不确定性的日常生活中是做不到的。当某个人问你多高或多重时，没有人能准确无疑地知道自己究竟有多高或多重。³ 如果只能用精确的语言进行陈述，结果就是这类最普通的问题都无法回答，只能说不知道。那么，语言就失去了日常交流的作用。因此，日常语言中模糊语词或语词的模糊性是无法彻底消除的，不能像数学或科学词汇一样先精确定义再使用。关于日常语言，常识或科学研究还支持如下事实：

事实 3. 日常生活中人们理解或使用模糊语词没有特别的困难，也不需要很高的智力或费力的思考，因此也不需要被消除。

事实 4. 语词的含义是语言使用群体的一种共同的约定俗成。

事实 5. 语词的含义是人们从语言使用的日常经验中学习得到的。

一个合格的解释模糊语词的语义学理论应当符合事实 1-5。

2 模糊语词的主要理论及其评述

如前所述，日常语言中绝大多数语词都是模糊的。因此，对模糊语词的解释是日常语言的语义学理论不可或缺的。当代比较有影响力的模糊语词理论包括：([17, 26, 27])

²在某些语境下（如法律中），可以强行规定清晰的边界，去掉边界情况。但这只适用于这些语境，其它语境下还是模糊的。何况，法律的某些边界规定是否合理还是有争议的。

³测量总是存在误差，即使仪器测出来某人身高 1.71 米，这里也包含了不确定性，因为很可能实际身高是 1.715 米，但仪器忽略了 0.005 米的误差。何况人的身高一天之内还是不断变化的。所以，即使你告诉别人你身高 1.71 米，1.71 米这个数字也不会被理解为你的精确身高，而是带有误差的一个模糊语词。

1. 认知主义 (epistemicism)
2. 超赋值主义 (supervaluationism)
3. 多值逻辑 (many-valued logic), 又称为真值度理论 (truth degree theory)
4. 语境主义 (contextualism)

其中, 认知主义的代表学者包括威廉姆森 (Timothy Williamson) 和索伦森 (Roy Sorensen) ([19, 20]); 超赋值主义的代表学者包括范恩 (Kit Fine) 和基菲 (Rosanna Keefe) ([2, 7, 8]); 多值逻辑的代表学者包括爱丁顿 (Dorothy Edgington) 和史密斯 (Nicholas J. J. Smith) ([1, 17]); 语境主义的代表学者包括坎普 (Hans Kamp)、拉夫曼 (Diana Raffman)、索姆斯 (Scott Soames)、夏皮罗 (Stewart Shapiro) 等 ([6, 14, 16, 18])。

为了准确理解这些理论的区别和各自的问题, 我们先看古典逻辑中如何刻画语词的精确含义。简单起见, 假设语言中只有真值连接词 \neg, \vee, \wedge , 个体常号 a, b, c, \dots 和一元谓词 F, G, H, \dots 。⁴ 语言中符号的含义和语句的真假由某个古典解释 (classic interpretation) $M = (W, I)$ 确定, 其中 M 满足如下条件: ([17], 第 24–31 页)

1. W 是一个非空集合, 称为论域;
2. I 是一个赋值函数⁵:
 - 对每个个体常号 a , $I(a) = a^I \in W$;
 - 对每个一元谓词 F , $I(F) = F^I$ 是 W 到 $\{0, 1\}$ 的函数;
 - 对每个原子语句 Fa , Fa 的真假

$$I(Fa) = F^I(a^I) \quad (1)$$

- 对任意语句 A, B , 其真假 $I(A), I(B)$ 满足:

$$\begin{aligned} I(\neg A) &= 1 \quad \text{iff} \quad I(A) = 0 \\ I(A \vee B) &= 1 \quad \text{iff} \quad \max\{I(A), I(B)\} = 1 \\ I(A \wedge B) &= 1 \quad \text{iff} \quad \min\{I(A), I(B)\} = 1 \end{aligned}$$

在评述各个理论时, 本文所关心的不是它们关于连锁悖论或其它某些哲学问题的解决方案或其中存在的问题。这些在前人的文献中已经有比较充分的讨论。本文更关心的是, 这些理论是否符合前面提到的关于日常语言的基本事实, 以及将这些理论实际应用到理解日常语言中会出现的问题。在人工智能中, 一个合格的语义学理论应当可以用来教会一个机器理解日常语言的模糊性, 从而让机器学会使用模糊语词。相比而言, 解决连锁悖论只是附带的结果。

⁴ 蕴含号、多元谓词、量词和个体变号对于理解本文的内容没有帮助, 因此不考虑。

⁵ 这里直接用数字 1 代表真, 0 代表假, 且用集合的刻画函数代表集合本身。

2.1 认知主义

认知主义对模糊语词的解释和精确语词是一样的。当人们说“张三很高”时，“很高”被关联到一个预期的古典解释（intended classic interpretation）。因此，任意一元模糊谓词 F 的解释仍然是论域 W 到 $\{0, 1\}$ 的函数。不同的是，对于精确语词人们准确知道它的解释函数是什么，而对于模糊语词则不能准确知道。换言之，属于边界情况的原子语句 Fa 仍然或真或假，只是人们不知道它究竟是真是假。（[26, 27]；[17]，第 34–35 页）

因此，认知主义有两个基本主张：（一）模糊谓词的语义分界点是存在的；（二）我们无法知道该分界点在哪儿。⁶ 第一个主张是非常反直观反常识的：光谱中存在一个精确的点，在那里红色变成橘色；仅仅失掉一根头发就能让我们从不秃头变成秃头；人们在某个高度不能算高个子，但再长高百分之一毫米却成了高个子。（[7]，第 64 页）然而，如此反直观的结论，认知主义却从来没有证明过，这是该理论最大的问题。（[27]）认知主义的主要工作在于论证第二个主张。然而，第二个主张是显而易见的事实，根本不需要证明。因为如果我们能知道分界点在哪儿，那么语词的含义就精确了，模糊性就不存在了。

从认知科学的角度来看，语言是组织、处理和传递信息的一种工具。（[3]，第 3 页）因此，一个合格的有所建树的语义学理论必须告诉我们，使用语言时究竟传递了哪些信息。比如说，小张告诉你，小明长得很高。虽然不能从这句话推出小明究竟有多高，但和不说这句话相比，小张肯定传递了更多的信息，而听到这句话的人显然也知道了更多的信息。不然的话，小张费力说这句话干嘛呢？因此，语义学理论必须告诉我们，当人们说出包含模糊语词的一句话，相比不说这句话时究竟传递了哪些更多的信息。然而，认知主义在这方面什么都没做。因此，它对于增进人们对模糊语词的理解没有帮助，可以说毫无建树。

2.2 超赋值理论

主流的超赋值理论认为：模糊语词的边界情况不具有真值。因此，它的第一个修改是，将古典解释中对谓词的赋值函数从全函数（total function）变成偏函数（partial function）：对于任意一元谓词 F ， $I(F) = F^I$ 是 W 的子集 U 到 $\{0, 1\}$ 的函数， $V = W - U$ 是没有真值的边界情况，即对任意 $v \in V$ ，原子语句 Fv 没有真值。对于复合语句的真值，超赋值理论的修改是：先将赋值函数从偏函数扩展

⁶虽然都支持上述两个主张，威廉姆森认为模糊性是认知中产生的（[20, 27]），索伦森则认为是语义特征使真者间隙（truthmaker gap）引起的（[19, 24]）。后面的多值逻辑则认为模糊性是形而上的，是世界本身所具有的。（[17]，第 70 页）然而，模糊性究竟是形而上的，是认知中产生的，还是意义构造中产生的，都无关紧要，因为最终的结论都是语义分界点不可知，边界情况的归属不确定，否则模糊性就不存在了。因此，模糊性究竟从哪儿产生对于语言的使用者来说不用关心，本文也不讨论这个问题。真正需要关心的是承认这种不可知或不确定性后还能做什么。

为全函数，因此得到很多不同的古典解释；如果在所有这些古典解释下，某个复合语句为真，那么它就超赋值为真；如果在所有这些古典解释下，某个复合语句为假，那么它就超赋值为假；其它情况则真假不确定或没有真假。在这种定义下，复合语句的真值不再完全由它组成部分的真值所决定。（[17]，第 76–79 页）

如果将赋值函数从偏函数扩展为全函数时不加限制，那么很容易推论得出，在上述定义下只要复合语句中包含边界情况，即其中包含的某个原子语句没有真假，那么除非该复合语句是古典逻辑中的重言式或矛盾式，否则它也没有真假。因此，在具体问题中，这种扩展通常要施加某种限制，得到某类可容许古典扩展（admissible classical extension），或称为可容许精确化（admissible precisification）。比如说，假如身高 1.80 到 1.90 米之间是“很高”的边界情况。当某个扩展将 1.85 米解释为“很高”，那么再将 1.85 到 1.90 米之间的身高解释为“不是很高”一般是不容许的。（[17]，第 79–81 页）

不论超赋值理论对于连锁悖论的解决方案如何，实际应用超赋值理论理解日常语言的模糊性时会遇到如下困难：（[25, 27]）

1. 如果可容许古典扩展中有真有假，那么该边界情况仍然没有真假。换言之，这时所有这类边界情况都是没有区别的。然而，实际应用中，即使都是“很高”的边界情况，1.83、1.85、1.88、1.90 米还是很不一样的。但这种区别超赋值理论没有刻画。
2. 超赋值解释下，需要通过扩展得到古典解释。具体应用中，究竟哪些扩展是可容许的？这些扩展是怎么得到？由谁来进行扩展？
3. 只有知道所有可容许扩展才能确定某个句子的真假。然而，可容许扩展非常多，很多时候还是无穷的。这在实际应用中人们是做不到的。那日常生活中人们是怎么毫不费力地学会并使用模糊语词的呢？
4. 高阶模糊性：对于某个具体的谓词而言，究竟哪些是它的边界情况？能够完全确定它的边界情况吗？换而言之，在构造解释时，能完全确定 U 和 V 吗？似乎做不到。
5. 非真值函项性：超赋值理论放弃了真值的函数式复合，不少哲学家认为这违反了常识。比如说，常识认为， p 或 q 为真当且仅当 p 为真或 q 为真。超赋值理论则认为， p 或 q 为真无法得出 p 为真或 q 为真。这涉及到超赋值理论中的超真和超假究竟如何理解的问题。

概率论实质上是一种特殊的超赋值理论，因为每一个古典赋值都可视为一个可能世界或可能情境，而超赋值为真/假就是在所有可能世界/情境上为真/假，即概率为 1 或 0。然而，概率论比超赋值理论更为精细，因为它将并非超真也并非超假的情况也进行了程度的区分；同时，概率值不是理解为句子的真假程度，而

是理解为信念的可信度 (credibility)。([13], 第 3-5 页) 后面会看到, 当采用概率这种特殊的超赋值理论时, 上面这些问题都可以得到很好的解决。

2.3 多值逻辑

多值逻辑是真值超过两个的逻辑理论的统称。下面主要讨论模糊逻辑 (fuzzy logic), 其它多值逻辑的讨论参见文献 [17] 第 50-59 页。模糊逻辑将二元真值 $\{0, 1\}$ 替换为实数区间 $[0, 1]$, 并保留复合语句的真值函数性。于是, 语言中符号的含义和语句的真假由某个模糊解释 (fuzzy interpretation) $M = (W, I)$ 决定, 其中 M 满足如下条件: ([17], 第 59-71 页)

1. W 是一个非空集合, 称为论域;

2. I 是一个赋值函数:

- 对每个个体常号 a , $I(a) = a^I \in W$;
- 对每个一元谓词 F , $I(F) = F^I$ 是 W 到 $[0, 1]$ 的函数, 被称为 F 在 M 中的隶属度 (degree of membership) 函数;
- 对每个原子语句 Fa , 隶属度 $F^I(a^I)$ 理解为 Fa 的真假程度或真值度 (degree of truth)

$$I(Fa) = F^I(a^I) \quad (2)$$

- 对任意语句 A, B , 其真假程度 $I(A), I(B)$ 满足:

$$\begin{aligned} I(\neg A) = i & \text{ iff } I(A) = 1 - i \\ I(A \vee B) = i & \text{ iff } \max\{I(A), I(B)\} = i \\ I(A \wedge B) = i & \text{ iff } \min\{I(A), I(B)\} = i \end{aligned}$$

很容易看到, 除了真值从 $\{0, 1\}$ 变为 $[0, 1]$, 模糊解释和古典解释的方法基本相同。因此, 古典逻辑是模糊逻辑的特例。然而, 正是由于继承了古典逻辑的真值函数复合方式, 模糊逻辑出现了一个严重的问题。假设命题 p 的真值为 0.5, 即 $I(p) = I(\neg p) = 0.5$, 那么:

$$I(p \wedge \neg p) = \min\{I(p), I(\neg p)\} = \max\{I(p), I(\neg p)\} = I(p \vee \neg p) = 0.5 \quad (3)$$

这在几乎所有实际例子中都是严重违反直观或常识的。([27]; [11], 第 53 页)

模糊逻辑的这个问题源于它的真值函项性: 复合命题的真值由其组成部分的真值通过某种函数的方式完全决定。我们可以改变函数复合的具体方式, 比如不再用 \max 或 \min 的复合方式, 但很快会遇到类似的问题或者其它困难。([11], 第 53-65 页; [9], 第 3 章)

包含模糊语词的句子真的具有真值函项性吗? 范恩、威廉姆斯等学者认为并非如此。假设你看到一个点 X 的颜色介于红色和黄色之间, 但不确定是哪种颜色,

这时“X 是红色”“X 不是红色”和“X 是黄色”都不确定为真，也不确定为假。然而，复合句“X 是红色或不是红色”和“X 是红色或 X 是黄色”确定为真，“X 是红色且不是红色”和“X 是红色且 X 是黄色”确定为假。因此，真值不是函数式复合的。当然，也有学者如史密斯等反对这种直观。（[17]，第 85 页）

模糊逻辑的第二个问题是，究竟如何指派真值度？1000 根头发的人秃头的真值度是多少？2000 根呢？模糊逻辑给出了一些计算真值度的方法，如专家共识等，然而这种方法的应用范围很局限。（[9]，第 10 章）如果方法太复杂，则难以解释人们是如何学会的。因此，如果没有一个简单而系统的方法决定真值度如何计算，那么这个理论就难以应用甚至毫无用处。而如果能够准确赋予每个句子真值度，那么模糊性还存在吗？（[27]）

上面两个问题其实关系到一个更为根本的问题：什么是真值度？它究竟该如何理解？张三说的话的真值度是 0.25，这是什么意思？一种比较合理的理解是张三所说的全部话语中 25% 为真。然而，单个语句“张三秃头”的真值度为 0.25 显然不是这种理解。那么，究竟该怎么理解“张三秃头”和“张三很高”的真值度都是 0.25？实际上，不少哲学家认为，真和假没有程度的区分，因此某句话为真的程度是 0.25 是无法理解的。（[27]）

本文同样认为，真假是关于某个世界的形而上的判断，在这个世界中句子要么真要么假，没有程度的区分。有程度区分的是人的认知或信念中对句子真假的不确定性。不确定性一旦出现，意味着世界本身如何是无法确切知道，能谈论的只有对某个陈述相信的程度。因此，将隶属度理解为真假程度是不合适的，应当理解为语句的可信度，而可信度则需要用概率来刻画。（[13]，第 3-5 页）

有学者将模糊逻辑同时称为真值度理论和概率解释（[22]），但二者实际上是截然不同的。真值度理论中，隶属度被理解为句子真假的程度，并且是函数式复合的；概率解释中，隶属度被理解为句子的概率值或可信度，并且不是函数式复合的。后面会看到，一旦将隶属度理解为某种概率值，那么模糊逻辑的各种困难都将迎刃而解。

2.4 语境主义理论

语境主义是和前面各种理论相互独立的理论。换言之，前面的各种理论都可以和语境主义相结合，形成认知主义语境主义、超赋值语境主义、多值逻辑语境主义，甚至古典逻辑语境主义等。（[17]，第 113-122 页）

语境主义理论不是简单地认为模糊语词的解释随语境而变化。这其实是所有理论所公认的事实。比如说，张三有 1000 万财产。一群普通人中，“张三比较富裕”是成立的；但一群亿万富豪中，“张三比较富裕”就显然不成立了。语境主义的核心在于处理边界情况时逐步递进的两条原则：

- 自由原则 (Freedom): 如果 P 是一个模糊谓词, a 是它的一个边界情况, 那么人们可以自由地断言 Pa 或 $\neg Pa$ 。([17], 第 115 页; [16], 第 10 页)
- 力量原则 (Power): 在所述情形下, 如果一个人断言 Pa , 那么 Pa 因这一断言而为真, 即 a 在 P 的外延中。([17], 第 115 页)
- 适应原则 (Adjustment): 在所述情形下, 不仅 a 在 P 的外延中, 和 a 的 P -性质无法区分的个体都在 P 的外延中。([17], 第 116 页; [18], 第 211 页)

举例来说, 假设小明身高 1.8 米, 处于“很高”和“不是很高”的边界情况。在某个语境下, 小张断言“小明很高”。这时, 小明很高这句话在该语境下就是真的, 即小明在谓词“很高”的外延中。同时, 身高和小明难以分辨的其他人也自动在谓词“很高”的外延中, 即这些人很高的断言也是真的。

人的某种断言确实构成了语境的一部分, 会影响语词的含义; 模糊语词的边界情况自然也可以由人的断言进行精确化。然而, 在上述例子中, 假设不止小张一个人进行了断言, 还有小王断言“小明不是很高”, 小李断言“小明很高”, 小赵断言“小明不是很高”。在这四个人同时进行了断言的语境下, “小明很高”究竟是真是假呢? 现实生活中, 多个人对同一对象进行不同的断言才是更常见的, 因为语言是一群人共同使用的。难道只考虑一个人断言的语境, 不考虑多个人断言的语境? 实际上, 认为某个人的断言就能决定某个语词的含义违反了事实 4, 即语词的含义不是由某个人决定的, 而是由语言的使用群体共同约定的。

语境主义还有另外一个问题: 对于边界情况, 某个人的断言决定其真假。可是, 究竟哪些情形是边界情况呢? 如何预先判定呢? 模糊语词之所以是模糊的, 很多时候它的边界情况也是没有共识, 不确定的。比如, 小明的例子中, 1.78 算不算边界情况呢? 只有在完全确定某个情形属于边界情况的情况下, 语境主义的理论才能用, 但在任意语境下完全确定某个语词的边界情况恐怕是不可能做到的。如果认为边界情况就是源于某些人说是而某些人说不是, 那么就回到前面那个问题了。模糊语词的概率解释可以很好地解决语境主义理论的这两个问题。

3 模糊语词的概率解释

如前所述, 语言是组织、处理和传递信息的一种工具。因此, 一个语义学理论应当告诉我们, 使用语言时究竟传递了哪些信息。比如说, “小明长得很高”这句话究竟告诉了我们什么呢? 由于“很高”这个模糊语词包含很强的不确定性, 所以从这句话无法推论得出小明的精确身高。然而, 和不说这句话相比, 陈述人肯定传递了更多的信息, 听到这句话的人显然也知道了更多的信息。那么, 这句话究竟传递了哪些更多的信息呢?

抛一个硬币, 落地前人们不知道结果是正面朝上还是反面朝上。我们可以停

留在这儿，就说我们不知道结果怎样，也没有人知道。认知主义就是这样一种态度。然而，科学家显然有更积极的态度，就是研究不确定性中的某种确定性：通过多次抛硬币可以得出，正面朝上和反面朝上的概率都是约 50%。有了这个结果，虽然还是不知道下一次抛硬币究竟是正面还是反面朝上，但我们能预测，抛了 100 次后，大概会有 50 次正面朝上，50 次反面朝上。这就比单纯说不知道提供了更多的信息。由于不确定性几乎无处不在，这样的故事在当代科学研究中反复出现，在很多领域已经成为主要的研究方法。

语词的模糊性也一样。模糊性没有什么神秘或特别之处，它就是一种不确定性：听到包含模糊语词的句子后，我们不确定对象所具有的准确属性究竟是什么。听到“小明长得很高”，我们不知道小明究竟有多高；听到“小明秃头了”，我们也不知道小明究竟还剩多少头发。然而，这种不确定性中一定存在某种确定性，不然从“小明长得很高”这句话中得不到任何信息，即说这句话和不说这句话就没有区别了。⁷

如何刻画模糊性中的不确定性，从而得到其中的确定性呢？虽然学者们发明了很多刻画不确定性的理论，它们在某些应用领域各有所长（[4, 11]），然而，最容易理解的，和常识最一致的，得到数学家最充分研究的，具有最坚实理论基础的，同时具有最广泛且成功的应用的只有概率论和以概率论为基础的统计学。（[12]，第 9 章；[5]，第 1 章）因此，除非被证明完全不适用，不用概率论刻画模糊性中的不确定性，反而从头开始创造形形色色的各种理论，才是最奇怪的。

因此，本文采用概率方法刻画模糊性中的不确定性。模糊性带来的不确定性主要包括两个方面，因此需要解决的问题也包括下面两种：

1. 理解：听到带有模糊语词的语句时，人们能最大程度地知道什么信息？比如听到“张三很高”时，虽然不知道张三的准确身高，能否知道张三身高 1.8 米的可能性有多大？
2. 表达：知道一个事物某方面的特征时，如何恰当地用带模糊语词的句子表达？比如说，看到一个人身高约 1.8 米，在某个语境下究竟说他比较矮，一般高，比较高，很高，还是特别/非常高呢？

与读者设想的可能不一样，第 2 个问题其实更简单，第 1 个问题的解决是基于第 2 个问题的。所以，下面先讨论第 2 个问题。看到一个人身高约 1.8 米，究竟说他不高，一般高，比较高，很高，还是特别高呢？这个问题首先取决于语境：我们是在 NBA 篮球运动员里面说，在荷兰人里面说，在中国北方人里面说，在中国南方人里面说，还是在一群 12 岁的男孩子里面说？显然，对于不同的群体，恰当的说法不是一样的。因此，首先要在语境中确定所谈论比较的对象有哪些，本

⁷按照信息论，得到信息就是减少不确定性的过程，即增加确定性的过程。（[10]，第 1 章）

文称之为子论域。子论域确定后，语词含义的确定遵循前面提到的如下事实：

- 事实 4：语词的含义是语言使用群体的一种共同的约定俗成。
- 事实 5：语词的含义人们从语言使用的日常经验中学习得到的。

对于大多数日常语词而言，约定俗成不是大家真的坐下来一起约定，而是看大家共同的使用习惯。因此，需要观察所在的语言使用群体日常生活中究竟是如何使用这些语词的。实际上，这种观察几乎每天都在进行：看到某个人身高约 1.8 米，然后看到身边某个人说“他长得很高”，于是就学习得到了“很高”这个词的一次使用经验。假如遇到 100 次这样的经验，看到 1.8 米身高的人后，其中 5 次人们说他“特别高”，70 次说他“很高”，20 次说他“比较高”，3 次说“一般高”，2 次说他“不高”，那么就能总结得到：对于约 1.8 米身高的人，这个群体中约 5% 的人说他“特别高”，约 70% 的人说他“很高”，约 20% 的人说他“比较高”，约 3% 的人说他“一般高”，约 2% 的人说他“不高”。以后遇到 1.8 米身高的人时，在同样的语境下一般也会说这个人“很高”，因为群体中最多的人是这么说的，除非故意不遵循群体的使用习惯，但这样就会让听众产生误解，达不到有效交流的目的。对于 1.8 米身高的人可以这么学习，对于其它身高的人也可以这么学习。当各个身高的使用经验都学到了，那么就可以说学会了这些词在这个群体中的使用习惯，也可以说学会了这些词的含义。

上面实际上是一种抽样统计的方式得到某个语言使用群体中语词的使用习惯，进而得到其含义的过程。这样做假定了人们对于语词的使用习惯服从某种概率分布，而抽样统计则是从日常经验中学习得到这种分布的近似结果。基于这一假设，可以得到如下定义。假设语言使用群体为 Z ，给定语境为 C 。语言中符号的含义和语句成立的概率由某个概率解释（probabilistic interpretation） $M^{ZC} = (W, I)$ 决定，其中 M^{ZC} 满足如下条件：

1. W 是一个非空集合，称为论域；
2. 对每个谓词 F ，给定语境 C 确定了 F 所谈论的子论域 $W_F \subseteq W$ ，同时确定了 F 所谈论的子论域元素的某种特征 T_F ，即存在特征函数 $f_F : W_F \rightarrow T_F$ 。⁸
3. I 是一个赋值函数：
 - 对每个个体常号 a ， $I(a) = a^I \in W$ ；
 - 对一元谓词 F ， $I(F) = F^I$ 是 T_F 到 $[0, 1]$ 的隶属度函数；
 - 对每个原子语句 Fa ，如果 $a^I \in W_F$ ，隶属度 $F^I(f_F(a^I))$ 理解为 Fa 成

⁸比如说，假设谓词 F 是“很高”。某语境确定该谓词谈论的子论域是所有篮球运动员，同时确定该谓词所谈论的是篮球运动员的身高特征。假设谓词 F 是“谷堆”。某语境确定该谓词谈论的子论域是所有堆放的谷物，同时确定该谓词所谈论的是谷物的颗粒数目（或许再加上谷物堆放的形状）。古典逻辑中，特征和类都是谓词，有相同的解释方法。是否需要把类和特征区分，本文暂不讨论。

立的概率:

$$P(Fa) = F^I(f_F(a^I)) \quad (4)$$

- 对任意语句 A, B , 其概率 $P(A), P(B)$ 遵循概率论的三条公理: ([13], 第 3-5 页)

(a) $0 \leq P(A) \leq 1$

(b) $P(A) = 1$

A 是重言式

(c) $P(A \vee B) = P(A) + P(B)$

$A \wedge B$ 是矛盾式

上述定义中, 隶属度函数 F^I , 即每个原子语句 Fa 成立的概率决定于群体 Z 如何使用谓词 F , 一般来说是不可能准确知道的。科学研究中通常使用抽样统计方法, 用统计得到的数值近似地表示这些概率值。比如说, 看到一个 1.85 米的人 a , 100 个人中有 80 个说这个人“很高”, 那么我们认为“ a 很高”成立的概率是 0.8; 看到一个 1.75 米的人 b , 100 个人中有 20 个说这个人“很高”, 那么我们认为“ b 很高”成立的概率是 0.2。当然, 如果语言使用群体 Z 不同, 这些概率值会不同; 子论域 U 不同, 这些概率值往往也不同; 即使 Z, U 都相同, 随着时间的推移这些概率值也会不一样, 因为人们的使用习惯很可能改变了。但不论如何, 隶属度函数 F^I 是可以通过抽样统计人们的语言使用习惯而直接得到的, 有非常明确而简单的解释。模糊逻辑中的真值度就没有这个优点。假如有一集共享子论域 U 和特征 T 的谓词 $\{F_1, \dots, F_n\}$ (比如前面提到的比较矮, 一般高, 比较高, 很高, 非常高), 其中每个谓词的隶属度函数都知道了, 如何选择其中最合适的谓词进行表达就很容易了: 对任意对象 $a \in U$, 选择 F_i 使得 $P(F_i a)$ 的值最大即可, 因为群体中最多的人是这么说的。

下面讨论模糊语词的理解问题: 听到有人说“张三很高”, 能否知道张三身高 1.8 米的可能性/概率有多大? 如果知道了谓词“很高”的隶属度函数, 再加上一个条件, 就能推出想要的结果。下面用一个简单的例子说明如何做到。假如一个硬币抛 100 次, 50 次正面朝上, 50 次朝下。假如有人对这些结果做“抛得好”的评价, 且该谓词的隶属度函数是这样的: 正面朝上时, 有 80% 的概率听到“抛得好”; 正面朝下时, 有 10% 的概率听到“抛得好”。下面的问题是: 当听到“抛得好”时, 有多大概率是正面朝上呢? 计算很简单: 抛 100 次硬币, 总共听到 $50 \times 0.8 + 50 \times 0.1 = 45$ 次“抛得好”, 其中一共有 $50 \times 0.8 = 40$ 次正面朝上。因此, 听到“抛得好”, 正面朝上的概率是 $40 \div 45 \approx 0.89$ 。

如果这个硬币不是公平的, 比如说抛 100 次, 80 次正面朝上, 20 次正面朝下。计算结果还是一样吗? 显然是不一样的: 100 次抛硬币, 总共听到 $80 \times 0.8 + 20 \times 0.1 = 66$ 次“抛得好”, 其中一共有 $80 \times 0.8 = 64$ 次正面朝上。因此, 听到“抛得好”, 正面朝上的概率是 $64 \div 66 \approx 0.97$ 。所以, 模糊语词的理解不仅和该语词的隶属度

函数相关, 还和该语词相关联的子论域中元素特征的概率分布相关。下面是一般的结果⁹:

命题 3.1. 假设 F 是一个一元谓词, 语境确定的子论域为 $W_F = \{a_1, \dots, a_m\}$, 特征为 $T_F = \{t_1, \dots, t_n\}$ 。已知 W_F 中 T_F 的概率分布: $P(t_i) = p_i$ ($\sum_{1 \leq i \leq n} p_i = 1$)。假设给定概率解释 $M = (W, I)$ 下, 已知隶属度函数 $F^I : T_F \rightarrow [0, 1]$ 。¹⁰ 假设有人说个体 a_i 具有性质 F , 即 Fa_i 。那么, a_i 具有特征 t_i 的概率为:

$$\frac{F^I(t_i) \times p_i}{\sum_{1 \leq i \leq n} F^I(t_i) \times p_i} \quad (5)$$

证明. 这实际上是贝叶斯公式的一个应用。假设命题 A_i 为: 某个体 x 具有特征 t_i , 命题 B 为: 有人说 x 具有性质 F 。题设实际上告诉我们:

$$P(B|A_i) = F^I(t_i) \quad (6)$$

$$P(A_i) = p_i \quad (7)$$

待求的是 $P(A_i|B)$ 。根据贝叶斯公式:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} \quad (8)$$

$$= \frac{P(B|A_i)P(A_i)}{\sum_{1 \leq i \leq n} P(B|A_i)P(A_i)} \quad (9)$$

$$= \frac{F^I(t_i) \times p_i}{\sum_{1 \leq i \leq n} F^I(t_i) \times p_i} \quad (10)$$

□

上面只考虑了特征离散的情况。有读者可能会问: 像身高之类显然是连续的特征变量, 难道不考虑了吗? 在实际应用中, 尤其是在抽样统计中, 连续变量会被离散化。比如身高, 我们不可能统计实数轴上每个精确身高值的个体出现的频率, 而只能统计每个小的身高区间的个体出现频率。这是因为, 测量不可能完全精确, 很多时候人无法识别也不会在意细微的区别。比如身高相差 1cm 大概都很难察觉出来。因此, 当某个人说身高 1.8 米时, 他实际上说的很可能是区间 1.8 ± 0.005 米。因此, 只考虑离散特征对绝大多数日常应用都是足够的。有读者可能会担心,

⁹有些读者可能会觉得, 下面的计算实在太简单, 不像一个理论。但不要忘了, 理解日常语言的模糊语词是绝大多数人都能胜任的, 智力很一般都没问题。如果需要复杂的理论或复杂的计算, 普通人又怎么不需要费力就能理解模糊语词呢? 因此, 计算简单恰恰是本文理论的一个显著优点, 因为它能解释为何人们能轻易学习并掌握它。实际上, 贝叶斯公式的计算可能都太难了, 很多时候人们都只能凭感觉猜一个接近的结果, 而非计算得到更精确的结果。

¹⁰假设不知道特征函数 $f_F : W_F \rightarrow T_F$, 不然就能直接知道每个 a_i 的特征, 不需要再推了。

连续特征离散化会不会丢失一些信息。这种担心是不必要的。连续特征离散化会丢失信息，指的是我们已经知道连续的分布函数是什么，但为了简化或实际应用而使用离散化表示方法。这个过程确实会丢失信息。但统计的情况是，连续的分布函数是什么本来不知道，我们只能收集离散的数据，得到离散的分布，然后用离散的分布推测连续的分布。这个过程不会丢失信息，反而人为添加了信息，因为这时连续分布比离散分布多的信息不是直接知道的，而是推测出来的。比如说，平面上统计得到两个点： $(1,1)$ ， $(2,2)$ 。这两个点属于哪个连续函数是不知道的。我们可以猜测它们在一条直线上，然后得出这条直线的方程是 $y = x$ 。连续函数 $y = x$ 的确包含了比这两个点更多的信息，但这些信息是我们猜测出来的，是人为添加的，很可能是错的。当然，如果能证明假设成立，即证明这两个点确实是同一直线上的点，那么用连续函数就更准确。但这种证明实践中很难做到，所以直接用统计得到的离散分布函数往往更合适。类似的，一个模糊语词比如“很高”在某个群体某个语境中究竟是什么意思，即它究竟对应哪个身高-概率连续分布函数，其实是不知道的。能够知道的只有 A 说 B 很高时，B 是 1.82 米左右；C 说 D 很高时，D 是 1.80 米左右等。我们是根据这些数据统计得到一个离散的身高-概率分布函数，而不是从一个已知的连续函数通过丢掉某些信息来得到这个离散函数，因此这里没有信息丢失的问题。

在模糊语词的概率解释中，每个模型 M^{ZC} 都加了上标群体 Z 和语境 C 。加上标 Z 意味着群体不同，模型可以不同。因此，在实际应用建模中，首先要考虑的是针对哪个群体建模，然后再在这个群体的语言使用习惯中进行统计得到模糊语词的概率分布。不同群体得到的统计数据不同，得到的概率分布不同，最终得到的模型也不同。如果有多个模型，究竟采用哪个要看实际应用。一般而言，如果数据量差别不大，群体越小模型越准确，选择越优先。实际上，这里的群体 M 可以是单个人，因为每个人其实都有自己独特的语言使用习惯，可以进行单独的统计和建模，只看有没有必要这么做。如果某个人很重要，比如是自己的家人、朋友、恋人或上司等，那么很可能我们就会根据这个人的使用习惯单独建立一个语词的解释模型，这样当这个人说出某个语词时，才不会产生大的误解或偏差。比如说，假设大众一般称 1.8 米左右的人很高，但某个人就是喜欢称 1.7 米左右的人很高，称 1.8 米左右的人特别高。那么熟悉这个人后，就需要对他单独建模。下次他再说某个人很高时，我们会理解为那个人只有 1.7 米左右，而非按照大众的习惯理解为 1.8 米左右。针对每个个体单独建模是最准确的语词解释方法，但这需要对这个人有很多了解以积累很多数据。这是为何对某个人越熟悉，越能准确理解他说的话。但如果对某个人不熟悉，数据太少，那就只能用更大群体的数据和模型进行解释，准确性很可能也会随之下降。

还有读者可能会问：在实际交流过程中，说话者可能会策略性地使用模糊语

词,例如谈判中刻意模糊“合理价格”的范围。类似的语用策略是否会影响前述的概率分布?对于大群体和通常语境下的建模,这样的策略一般不会影响统计得到的概率分布。首先,语言的大多数使用往往是一种无意识行为,会自然遵循习得的语言经验;为避免交流困难或误解,人们通常也不会有意识地违反习得的语言经验;因此,只有少数人在少数特殊情况下,才会故意偏离语词的通常含义。其次,即使谈判等特殊情况,语词的使用也不能大幅度偏离正常含义,否则就是欺骗或欺诈了。比如说,买鞋时问卖家这鞋质量怎么样,卖家说很好,结果买回家穿几次就坏了。这时我们不会认为鞋“很好”这个词的含义因为卖家的这个说法而改变了,而只会认为卖家说谎了或搞错了。上述讨论说明,统计得到的大多数数据是正常的,只有极少数数据是不正常的;如果真的出现大的偏离,实践中通常会先剔除明显异常的数据再计算结果,这也是统计学的常用方法之一。然后,如果建模的群体比较大,那么其中某个个体的数据就只占很少的部分,对整体不会有大的影响;如果建模的群体比较小,比如为某个人或几个人单独建模,那么这个模型就只适用于这个人或这几个人,不能用于其他人,对于理解其他人的话也不会有影响。最后,如果某类特殊情况经常出现,需要重点对待,那么可以视为某种特殊语境,单独为它建模。在本文模糊语词的概率解释中,每个模型 M^{ZC} 都加了上标语境 C 。这意味着语境不同,模型可以不同,只看有没有必要这么做。在谈判、购物等语境下,有经验的人都会比较警觉,一般不会按照通常的理解去解释对方的话语,这可以视为特殊语境下对语词含义单独建模的结果。

4 概率解释对主流理论的继承和修正

概率解释虽然简单,但足以继承主流理论的优点,并克服其各种缺点。下面逐一分析。首先是认知主义。如前所述,认知主义最大的问题是,承认存在不确定性但不进一步分析,没有新的建树,对于理解模糊语词没有帮助。概率解释显然克服了这一缺点。同时,概率解释并不主张模糊语词的语义分界点存在,故也不会出现类似的反直观结论。

威廉姆森为认知主义辩护的最大理由是,相比其它理论认知主义可以保留古典二值逻辑。([27]; [20], 第186页)概率解释同样保留了古典二值逻辑,因为概率值并不是真值,而只是信念的可信度。某个可能世界或情境中命题的真假仍然遵循古典二值逻辑,而概率刻画的是人们对命题是否成立的一种主观信念。这种信念主要来源于人们的日常经验。对于模糊语词,其解释中的概率值来源于日常生活中对该语词的使用经验,即抽样统计人们是如何使用这一语词的。在此理解下,概率解释可以进一步解决前面提到的超赋值理论的所有困难:

1. 如前所述,概率论是一种特殊的超赋值理论:在所有可能世界/情境下为真

时概率为 1，对应超真；在所有可能世界/情境下为假时概率为 0，对应超假；其它情况概率大于 0 小于 1，对应真假不定或没有真假。相对主流超赋值理论，概率论对真假不定的边界情况也进行了细致的区分，能更精细地刻画模糊语词中的不确定性。

2. 主流超赋值理论中，一个可容许扩展得到一个古典解释，但这些扩展究竟是怎么得到的呢？概率解释可以合理地说明这一点：这些扩展是通过每个人的语言使用得到的。当某个人断言边界情况 1.75 米很高，1.70 米不是很高时，实际上就进行了一次扩展，虽然这种扩展通常是不完整的。为何这样扩展而非那样扩展是由每个人自己的经验所决定的。哪些扩展是可容许的则是由整个语言使用群体共同决定的：当某人说出一个明显不合理的论断时，比如 1.60 米很高，其他人往往会予以纠正。
3. 主流超赋值理论中，需要知道所有可容许扩展才能确定句子的真假（除非是真假不定或没有真假）。这对于实际应用是一种灾难，因为这样的可容许扩展非常多，甚至是无穷的。概率解释不会有这个问题，因为实际应用不要求得到精确的概率值，而是通过抽样统计得到近似的概率值即可，抽样的多少也可以根据实际情况增减。
4. 对于主流超赋值理论出现的高阶模糊性，概率解释也不存在，因为概率解释不需要事先确定一集边界情况 V 。实际上，概率解释中的精确概率值被认为是不知道的，因此不需要也不会断言边界情况的界限在哪。能做的是通过抽样统计得到近似的概率值。
5. 概率解释不具有真值函项性，因为概率值根本就不是真值，而只是人们对命题是否成立的一种主观信念。信念的可信度不是真值函数式复合的。“ A 或 B ”的概率为 1，即“ A 或 B ”100% 可信，显然无法推出“ A ”100% 可信或“ B ”100% 可信（把 B 换成 $\neg A$ 更明显）。

本文认为，真和假都是相对于单个可能世界或可能情境的形而上的断言。因此，真假本身没有程度的区分，这时古典逻辑成立，“ A 或 B 为真”能推出“ A 为真”或“ B 为真”。然而，一旦考虑不确定性，就需要考虑不同的可能世界或可能情境，就是在考虑认知中信念和可信度的问题，不再是真假的问题。可信度不遵循古典逻辑规则，而是遵循概率规则。混淆形而上的真假和认知上的不确定性或信念的可信度是主流理论出现分歧的主要原因之一。

如此理解后，模糊逻辑的第一个问题就解决了。因为隶属度函数不再解释为真假程度，而是解释为概率/可信度，不遵循任何真值函数式复合规则，因此不会得到模糊逻辑那种严重违反直观或常识的结果。概率的复合规则已经得到数学家的反复证明和科学研究的反复验证，不会出现这样的情况。模糊逻辑的其它问题，比如如何指派真值度，什么叫真值度等问题，一旦将真值度替换为概率值，都可

以由概率论来回答，这里不再赘述。

有学者认为，概率论是线序理论，所以无法刻画多维度的模糊性，需要更复杂更一般的偏序结构布尔多值语义等其它理论。([22]) 这点实难认同。以“聪明”为例。人们经常说一个人不聪明、聪明、很聪明、非常聪明等，同时也会说一个人比另一个人聪明等。说这些话时，大概很少假定是某个特殊维度才成立，而是整体的线性比较。这说明聪明是可以线性比较的。即使聪明有很多维度，也完全可以综合各种维度得到线性比较的结果，不然智商量表怎么得到线序的分数的呢？其它概念同样如此。因此，偏序结构的布尔多值语义在应用中是不必要的。同时，布尔多值语义和模糊逻辑一样，其真值是函数式复合的，因此同样面临模糊逻辑的一些问题。而且，该理论复杂而不直观，不是普通人能理解和掌握的，这样无法解释人们是怎么轻松学会理解模糊语词的。

最后看语境主义。语境主义认为人的断言可以决定模糊语词的语义，概率解释同样承认这一点，因为概率值是通过抽样统计人们的语言使用习惯而近似得到的。不同的是，语境主义认为，单个人的单个断言就能决定某个模糊语词在某个语境下的含义，概率解释则认为并非如此：语言使用群体中多数人的断言共同决定模糊语词的含义。这样就不会出现语境主义中不同的人同时进行相反的断言时，命题的真假或语词的含义反复改变的问题了。同时，概率解释认为，一个情形是否是边界情况，同样是由语言使用群体中多数人的断言共同决定的。边界情况是不能预先知道的。这样就不会遇到语境主义中需要事先决定哪些情形是边界情况，然后再对边界情况进行断言的问题了。

相比其它理论，概率解释更符合日常语言的基本事实：抽样统计得到近似的概率值是很容易的，不需要很高的智力或费力的思考，不需要经过复杂的推理或运算，是从语言使用的日常经验中学习得到，也体现了语词含义是语言使用群体的一种共同约定俗成，是由语言使用者的使用习惯共同决定。

5 结语

模糊语词的解释问题，哲学家的讨论和实际应用似乎有点脱节。当人们理解“秃头”时，真的是把头发数清楚然后判断是否是秃头？头发根数多的一定比少的不秃头？当人们理解“谷堆”时，真的是完全根据其中颗粒的个数来决定，而不是加上形状等其它因素？颗粒个数多的一定比少的更像谷堆？答案显然是否定的。实际上，人们根本就无法去数头发的数量或谷物的数量。¹¹当然，连锁悖论本身也很有意思。如何用本文的理论解决连锁悖论，读者可以查阅文献 [21]，这里不

¹¹如果哲学家只是把这个问题当作纯理论讨论的一个问题，或者纯智力训练，而不追求对模糊语词的实际理解或应用，那没问题。不过，这时就不要怪人们说哲学讨论脱离实际或哲学讨论没什么用了。

再赘述。

实际应用中是如何理解模糊语词的呢？人工智能的发展已经给出了答案。模糊语词的概率解释其实早在人工智能中得到广泛应用了。人通过看到某个人的头部然后判断其是否秃头；机器则通过图像识别（image recognition）判断某个人是否秃头。在图像识别中，模糊语词的含义是通过某种监督学习（supervised learning）得到的。以“秃头”为例。科学家先找一些人的头像，有秃头的，有不秃头的，有秃头多的，秃头少的，然后找人标记这些头像是否秃头，或者更进一步标记秃头的程度。同一张图像很可能会找很多人进行标记。对于边界情况，有的人标记秃头，有的标记不秃头，有的标记秃头程度高，有的标记不高。标记好之后，建立一个神经网络进行训练，训练过程就是一种统计机器学习的过程。学习好后，该神经网络就存储了人们关于秃头的语言使用习惯或经验。之后任给一张人的头像，该网络就能输出他是否秃头的概率，或者进一步输出属于某种秃头程度的概率。当然，随着神经网络越来越大，一个网络可以同时学习很多词的含义，而非仅仅“秃头”一个词。（[15]，第22章）人学习语言或许也经历了类似的过程，不过证明这一点是心理学的工作，超出了作者的能力，这里不再讨论。

参考文献

- [1] D. Edgington, 1999, “Vagueness by Degrees”, in R. Keefe and P. Smith(eds.), *Vagueness: A Reader*, pp. 294–316, Cambridge, MA: MIT Press.
- [2] K. Fine, 1975, “Vagueness, Truth and Logic”, *Synthese*, **30(3-4)**: 265–300.
- [3] D. Geeraerts and H. Cuyckens, 2007, *The Oxford Handbook of Cognitive Linguistics*, Oxford: Oxford University Press.
- [4] J. Y. Halpern, 2017, *Reasoning about Uncertainty (second edition)*, Cambridge, MA: MIT Press.
- [5] E. T. Jaynes, 2003, *Probability Theory: The Logic of Science*, Cambridge: Cambridge University Press.
- [6] H. Kamp, 1996, “The Paradox of the Heap”, in U. Monnich(ed.), *Aspects of Philosophical Logic*, pp. 225–277, Cambridge: Cambridge University Press.
- [7] R. Keefe, 2000, *Theories of Vagueness*, Cambridge: Cambridge University Press.
- [8] R. Keefe, 2008, “Vagueness: Supervaluationism”, *Philosophy Compass*, **3(2)**: 315–324.
- [9] G. J. Klir and B. Yuan, 1995, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, New Jersey: Prentice Hall PTR.
- [10] R. McEliece, 2004, *The Theory of Information and Coding*, Cambridge: Cambridge University Press.
- [11] J. Paris, 1994, *The Uncertain Reasoner’s Companion: A Mathematical Perspective*, Cambridge: Cambridge University Press.

- [12] J. Pearl, 1988, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Francisco: Morgan Kaufmann Publishers.
- [13] J. Pearl, 2009, *Causality: Models, Reasoning, and Inference*, Cambridge: Cambridge University Press.
- [14] D. Raffman, 2014, *Unruly Words: A Study of Vague Language*, New York: Oxford University Press.
- [15] S. J. Russell and P. Norvig, 2020, *Artificial Intelligence: A Modern Approach*, New Jersey: Pearson Education.
- [16] S. Shapiro, 2008, *Vagueness in Context*, Oxford: Oxford University Press.
- [17] N. J. J. Smith, 2008, *Vagueness and Degrees of Truth*, Oxford: Oxford University Press.
- [18] S. Soames, 1999, *Understanding Truth*, Oxford: Oxford University Press.
- [19] R. Sorensen, 2001, *Vagueness and Contradiction*, Oxford: Oxford University Press.
- [20] T. Williamson, 1994, *Vagueness*, London: Routledge.
- [21] 邢锬, “渐进式论证、连锁型滑坡论证和连锁悖论的概率分析”, 逻辑学研究, 2023年第2期, 第16–32页。
- [22] 张立英, 张君, “内涵与外延之辩: 基于含糊性语义解释演进的分析”, 逻辑学研究, 2021年第2期, 第68–87页。
- [23] 伍铁平, 模糊语言学, 上海: 上海外语教育出版社, 1999年。
- [24] 王晶, “索伦森‘使真者——间隙认知主义’及其困境”, 河北学刊, 2020年第6期, 第40–47页。
- [25] 陈明益, “含混性与超赋值论”, 哲学动态, 2014年第8期, 第104–109页。
- [26] 陈明益, “从逻辑哲学观点看含混性问题”, 逻辑学研究, 2015年第3期, 第64–87页。
- [27] 陈波, “模糊性: 连锁悖论”, 哲学研究, 2014年第1期, 第111–118页。

(责任编辑: 袁之)

A Probabilistic Interpretation of Vague Words

Kun Xing

Abstract

Vague words are ubiquitous in natural language, and people rarely have difficulty in using them. However, the interpretation of vague words have brought philosophers much more difficulties, and there is still no consensus in this problem. Today the mainstream theories include epistemicism, supervaluationism, many-valued logic (truth degree theory) and contextualism, but all of them face their challenges. Based on the reviews of those theories, this article suggests a mixed theory of many-valued logic and probabilistic method. The theory is aimed at overcoming the difficulties mainstream theories have, conforming most people's common sense and conventions in language-using, and following the practices in Artificial Intelligence.