

概念语义与弗雷格谜题消解*

周北海

北京大学哲学系

zhoubh@phil.pku.edu.cn

摘要：弗雷格因同一替换律讨论而提出了涵义与指称的理论，这个理论后来引出了弗雷格谜题。弗雷格谜题的形成有多方原因，直接指称论对弗雷格理论批评是主要原因之一，以至于可以说，这是产生于直接指称论哲学立场的谜题。尽管如此，弗雷格理论确有不足。最重要的是，弗雷格理论只有关于涵义与指称的理论，即只有语言层面的理论，而缺少认知层面的理论。这个不足使得在弗雷格理论基础上解决同一替换律问题难有令人满意的结果，让“谜题”多添了几分“迷”的色彩。这里将给出一个新的方案：在弗雷格理论的基础上，增加有关概念的理论，以概念和内涵、涵义等这些概念的形式刻画为中心，建立可以消解弗雷格谜题的形式语义学，即概念语义。通过概念语义可以在不同层次上对弗雷格谜题的消解给出统一回答。

关键词：弗雷格谜题，涵义，内涵，概念，涵义语义，概念语义

1. 弗雷格谜题及有关背景

1879年弗雷格发表了《概念文字》，其中给出了一个一阶逻辑的公理系统，标志着现代逻辑的诞生。为给出该系统，弗雷格讨论了等式及相关问题。这些讨论和观点在形式方面的凝结有两点：(1) 给出表达式 $\vdash a=b$ ，表示的是“ a 这个符号和 b 这个符号有相同的概念内容，因此到处都可以用 b 替换 a 并且反之亦然”（王路 2006；20-22）。这也是弗雷格对同一替换律最初的表述。(2) 给出了关于等词的两条公理： $x=x$ ， $x=y \rightarrow (Fx \rightarrow Fy)$ 。1891年弗雷格在论文“函数与概念”中对原来的观点做了一些修改：将“概念内容”细化为涵义（sense）和指称（reference）；两个符号不同的专名因概念内容相同可以相互替换也比较明确地改为因指称相同而可以相互替换（王路 2006；64-66），这就是外延的同一替换律。1892年弗雷格发表了后来影响深远的著名论文“论涵义与指称”，专门讨论相关问题，比较完整地提出了涵义与指称的理论。这个理论有以下要点：(1.1) 专名（包括摹状词）

*基金项目：国家社会科学基金项目 07BZX047。

通常有涵义和指称¹，涵义决定指称；(1.2)不同的专名可以指称相同而涵义不同；(1.3) $a=a$ 是先验的或分析的，不提供新知识， $a=b$ 提供新知识，具有认知方面的意义。弗雷格用晨星和昏星的例子解释相关问题：“晨星”和“昏星”有自己的涵义和指称，它们的涵义不同，但指称相同；“晨星等于晨星”不提供新知识，“晨星等于昏星”提供了新知识，有认知价值。至此可以认为，在关于等式或等词问题上，弗雷格不仅在形式方面给出了一阶逻辑关于等词的逻辑意义的刻画，而且也合直观地解释了一些问题，给出了相应的哲学基础。所剩的只有“两小片乌云”²：(1)什么是涵义还不够清楚，没有清楚到在形式刻画方面有确切的体现，而且除了在依赖直观的论证中存在，涵义不像个体、真值那样在语义学中不可缺少；(2)外延同一替换律的成立是有条件的，同指称的个体词并不是“到处”都可以相互替换，比如在某些认知或内涵语境下就不可替换。晨星和昏星就是常用的例子：我们都知道晨星是晨星，但是并不一定都知道晨星就是昏星。这就是外延同一替换律应该失效的问题，以下也简称同一替换律问题。不过至此为止，弗雷格的理论至多是还有些不足，他所面临和试图解决的问题中还没显现有什么难题难到成为谜题³。所以后来出现了“谜题”有两个原因，恰好对应这两小片乌云。

首先是在涵义方面受到置疑。最严峻的挑战来自于后来克里普克加强和发展的专名理论，即直接指称论。该理论认为专名没有涵义，只有指称。按这个观点，既然专名只有指称而没有其他语义内容，于是很难解释

问题(1) $a=a$ 和 $a=b$ 为什么(怎么)会有认知方面的差异?

这就是所谓的弗雷格谜题。同样道理，也很难解释

问题(2)为什么同指称的专名在非外延语境下并非处处可以相互替换?

问题(2)即通常的同一替换律在非外延语境下应该失效的问题，是问题(1)在逻辑方面的体现。如果说问题(1)是哲学难题，那么问题(2)就是连带的逻辑学难题。问题(1)是否得到了正确的解答，连带地应该表现为问题(2)是否得到了合理解决。

根据叶闯(2007; 125)，Salmon 1986年以《弗雷格谜题》(Frege's Puzzle)为名出版专著，首次引进“弗雷格谜题”这一称谓，指如何解释 $a=a$ 与 $a=b$ 有不同认知意义的哲学难题。弗雷格谜题或繁或简有不同的表述。Bealer (1993; 18)将弗雷格谜题概括为：“ $[a=b]$ ，如果是真的，怎么会与 $[a=a]$ 有不同的意义(meaning)或认知价值?与此相关

¹ 有两点说明：(1)这里不谈没有指称的专名即空专名的情况。(2)弗雷格没有严格区分专名与摹状词，“晨星”和“昏星”是摹状词而不是专名，这个问题在后来的讨论中得到了澄清。但是这个澄清不是解决弗雷格谜题的关键。因为即使有这个澄清，弗雷格谜题依然还是谜题，而且可以把这个例子中的“晨星”和“昏星”分别换成“金星”和“太白星”，这可以看作专名，或者相应的英文单词“Phosphor”和“Hesperus”，这应该是“彻底的”专名，弗雷格谜题所提出的问题依然还是问题。因此本文不考虑这个区分，还是用弗雷格原来的例子。为表述方便，以下也用“个体词”概括专名和摹状词。

² 这里借用了19世纪物理学史上的“两朵乌云”的说法。

³ 这里把“谜题”理解为基本看不到解决出路的问题，一般的难题至少在表面上还有解决问题的出路。

的谜题是：为什么同指称的专名在命题态度句中的保真替换不成立？”

弗雷格理论提出或引发的问题导致了逻辑和哲学两个方面研究持续进行，经久不衰。首先是在逻辑学方面导致了内涵逻辑和内涵语义学的研究。上世纪三、四十年代，以考虑内涵推理为中心，Church 提出了的类型论式的内涵逻辑。四十年代，围绕内涵语义的形式刻画，卡尔纳普提出可能世界语义学式的内涵语义学，五十年代蒙太古进而提出类型论的内涵语义学。七十年代中期后，面对各种逻辑和形式语义学都未能解决同一替换律问题，Cresswell 提出了超级内涵逻辑 (hyper-intensional logic)。今天各种超级内涵逻辑的研究成为解决同一替换律问题的主要方向⁴。在哲学方面，关于弗雷格的理论，首先有罗素提出的专名与摹状词的区分以及著名的摹状词理论。四、五十年代有奎因关于同一性、模态语境下同一替换律问题的讨论⁵。五十到七十年代，可能世界语义学为模态逻辑提供了一个从方法论角度看几乎完美的形式语义，在技术方面得到了如完全性、可判定性等系列结果，将模态逻辑的研究提升到新的阶段，但是在哲学方面也引出了更多的问题。在其中一些问题的讨论中，克里普克提出直接指称论的专名理论，包括严格指示词理论、历史因果论的命名理论等，引起哲学方面研究涵义与指称等问题的热潮。弗雷格谜题也由此而凸显。时至今日，哲学方面的相关研究还在进行，出现了更多的谜题⁶，也出现了更多的理论和争论。

可以说，“涵义”本身不够清楚和直接指称论的置疑是弗雷格谜题产生的直接原因。其实从弗雷格理论立场看，可以理解为这是直接指称论的谜题，是他们要回答的问题。因为除指称外还有涵义， $a=a$ 与 $a=b$ 有不同认知意义的问题至少在直观上得到一定的解释，所以对于弗雷格涵义指称论来说，这个问题并不是谜题，尽管是一个难题。这个难题之难的表现是，如果弗雷格理论是正确的，为什么让涵义也像个体和真值一样不可缺少的各种内涵语义和内涵逻辑看起来并不成功，在同一替换律问题上的研究也迟迟得不到令人满意的进展。这增加了涵义指称论的可疑性，同时也增加了直接指称论的可信性。这意味着，也增加了“迷”的成分。正是在这个意义上说，涵义本身不够清楚和同一替换律问题成了“弗雷格谜题”成为谜题重要原因。

弗雷格谜题实际上是哲学立场的问题，在直接指称论立场上是个谜题，在弗雷格理论的立场上这是个难题。尽管难题与谜题有时难以分清，二者还是有本质的不同。我认为，如果弗雷格理论是正确的或其基本方向是正确的，就不应该只限于直观的解释，而应该再进一步，在这个理论基础之上给出相应的形式语义学，使得问题 (1) 和问题 (2) 在其中都得到解释和相应的形式刻画。尽管这还不足以消解直接指称论的弗雷格谜题，这是因为，如果坚持直接指称论的立场不变，那么无论怎样，弗雷格谜题都将是永远的迷，形式理论并不

⁴ 有关内涵逻辑的介绍可参见文学锋 (2009)。

⁵ 奎因年的著名论文《经验主义的两个教条》(1951) 对二十世纪哲学产生了重大影响。他所反对的第一个教条，分析命题与综合命题有明确区分，就是从分析同义词和同义词替换开始的。

⁶ 可参见 Salmon (1986), Fine (2007)。

能改变直接指称论的哲学立场，但至少可以提供一个合适的形式语义学，消解弗雷格理论的难题。本文将试给出一个基于弗雷格理论的形式语义学，在解决弗雷格理论难题的意义上消解弗雷格谜题。至于直接指称论的弗雷格谜题，或许只有在直接指称论对自己的理论有所修改后才有可能消除。⁷

2. 弗雷格理论：从形式的角度看

弗雷格理论涉及到事实、语言、认知和逻辑四个层面的问题。事实层面：被两个不同的名字所称呼的对象事实上可以是同一个对象，同一个名字也可以在不同情况下称呼不同的对象，有这些对象在事实层面的同一或不同的问题。语言层面：专名是否有涵义，什么是涵义，涵义与指称是什么关系，专名的指称是如何确定的等。认知层面：通过“ a ”，“ b ”，“ $a=a$ ”与“ $a=b$ ”这些表达式我们可以知道什么，有什么认知方面的意义等。逻辑层面：如何对外延同一替换律失效问题给出合理的解释。以上这四个层面只是相对区分。虽然弗雷格谜题表现为认知方面的哲学难题，但是与其他层次也有关联。作为弗雷格谜题消解来说，这四个层面又是紧密结合的，甚至可以说，是相互纠缠的，这增加了解决问题的难度，同时也说明，弗雷格谜题的消解不能“就事论事”，只限于问题(1)本身，而应该是不同层次的综合研究。本文将在这种“综合研究”的观点下谈论弗雷格谜题的消解。

至此本文关于弗雷格谜题的消解提出了三个出发点：基于弗雷格理论，建立形式语义学和多层次综合研究。综合此三点，实际上也是试图发展、完善弗雷格理论，消除该理论自身存在的问题，沿此路线达到弗雷格谜题的消解。以下将此种做法简称为弗雷格谜题消解诸方案中的“发展路线”。

弗雷格的理论在直观上简单、清楚，但是目前还没有形式语义学可以统一给出该理论几个要点的形式表达。以可能世界语义学为例，就是希望有下面的结果：存在模型 $M = \langle W, R_a, R_b, V \rangle$ ，存在 $w, w' \in W$ ，使得

- | | |
|---|------|
| (1) $M, w \models \text{晨星} = \text{昏星}$, $M, w' \models \text{晨星} \neq \text{昏星}$ | 事实层面 |
| (2) $M, w \models \text{晨星} \neq \text{昏星}$ | 语言层面 |
| (3a) $M, w \models K_a(\text{晨星} = \text{昏星})$, $M, w \models \neg K_b(\text{晨星} = \text{昏星})$ | 认知层面 |
| (3b) $M, w \not\models \text{晨星} = \text{昏星} \rightarrow K_b(\text{晨星} = \text{昏星})$ | 认知层面 |
| (4a) 并非“ $\models a = b \rightarrow K_a(a = b)$ ” | 逻辑层面 |
| (4b) 并非“ $M \models a = b \Rightarrow M \models K_a(a = b)$ ” | 逻辑层面 |
| (4c) 并非“ $\models a = b \Rightarrow \models K_a(a = b)$ ” | 逻辑层面 |

⁷ 弗雷格涵义指称论和克里普克的直接指称论各有不同的侧重，可以视为观众语言和角色语言两种不同语言的名称理论，详情请参见周北海（2010）。至于弗雷格谜题，首先这应该是观众语言的问题，直接指称论将角色语言的某些结论用于观众语言，有失偏颇。因此对于消解弗雷格谜题来说，弗雷格理论有优势，只是难题，而对于直接指称论来说，是谜题。这些观点和相关问题需要另文专门讨论。

其中“=”表示指称相同；“≠”表示涵义不相同；“ K_a ”表示“a知道”，“ K_b ”类似；“ \Rightarrow ”表示“从…可得…”。(1)是说：“晨星”和“昏星”在可能世界 w 上指称相同，在 w' 上指称不同。(2)：“晨星”和“昏星”在可能世界 w 上涵义不同。(3a)：在可能世界 w 上，认知主体 a 知道晨星就是昏星，而 b 不知道晨星是昏星。(3b)：在可能世界 w 上，事实上晨星等于昏星并不意味认知主体 b 知道晨星是昏星。(4a) - (4c) 分别表明不同层次的同时替换律不成立。(4a) 表明一般意义的同一替换律在认知语境中不成立，也是 (3b) 的一般形式；(4b) 表明模型有效条件下的同一替换律在认知语境中不成立；(4c) 表明有效条件下的同一替换律在认知语境中不成立⁸。由 (3b) 可得 (4a)，而且 (4a,b,c) 中也都有认知因素，这也表明认知层面和逻辑层面的问题是相互关联的。

(1) - (4) 分别对应于弗雷格理论所涉及的四个层面。其中 (2) 对于可能世界语义学来说不适用，因为可能世界语义学中没有表达涵义的部分。所以还将 (2) 列在这里，只是一种展示，表示一个能够解决弗雷格谜题的语义学应该也有这一层面的解释。对于“晨星等于昏星”这个例子来说，在通常的可能世界语义学中可以使得 (1)，(3a)，(3b) 和 (4a) 同时成立⁹，但是无法使得 (4b) 成立。这是因为根据可能世界语义学对于模态词的通常解释，从 $M \models a=b$ ，只能有 $M \models K_a(a=b)$ 。

在这个问题上可能世界语义学的不成功促进了后来的发展，是导致后来各种超级内涵语义学和超级内涵逻辑产生的重要原因。“具初步统计，目前解决超内涵问题的方案已有十几种之多”（文学锋 2009；45）。但是这些方案各有自己的不足和问题，弗雷格谜题还远没有解决。这里将给出一个新的方案：以涵义、概念和内涵等这些概念的形式刻画为中心，建立可以消解弗雷格谜题的形式语义学。具体做法将分为处理语言现象和认知现象两个阶段。在语言阶段，给出涵义语义，重点处理涵义和指称的关系；在认知阶段，建立概念语义，对弗雷格谜题做出正面解答。

3. 涵义语义：弗雷格理论的形式语义

在涵义语义中个体词的涵义是可能世界集到个体域的映射，个体词的指称是该映射在给定可能世界时的值（周北海，2008）。两个个体词的指称的等同，即事实层面的等同，是函数的值相同，涵义相同是函数本身相同。函数相同和某函数值相同是两个不同层次的问题，有关涵义和指称各类相同或不同的问题因此而自然可以区分。

形式语言 \mathcal{L}

符号 \mathcal{L} 有可数无穷的个体变元、个体常项、一元谓词符号，有等词 = 和 \equiv ，命题联结词 \neg ， \rightarrow ，量词符号 \forall ，辅助符号 (,)。被定义的符号有 \wedge , \vee , \leftrightarrow , \exists ，按通常定义。所

⁸ 模型有效的同一替换律： $M \models a=b \Rightarrow M \models \alpha \leftrightarrow \alpha[a/b]$ ($\alpha[a/b]$ ：将 α 中 a 的某处出现换成 b)。有效的同一替换律类似。

⁹ 这里涉及到偶然等同。关于偶然等同的赋值定义可参见周北海（1997；393-395）。

有变元的集合记作 Var 。

词 (1) 个体词：个体变元和个体常项都是个体词，分别用 x, y, z 和 a, b, c 表示，用 t 表示任意的个体词；(2) 谓词，用 P 表示。个体常项和谓词也称为非逻辑常项，简称常项，所有常项的集合记作 CT 。

公式 $\alpha ::= Pt \mid t=t' \mid t\equiv t' \mid \neg\alpha \mid \alpha\rightarrow\beta \mid \forall x\alpha \mid$

形式语言 \mathcal{L} 的语义

定义 3.1 一个 \mathcal{L} 框架是一个二元组 $\langle W, D \rangle$ ，其中 W, D 是任意非空集。

\mathcal{L} 框架中的 W 和 D 分别可以理解为可能世界集和个体域。

定义 3.2 设 $\mathfrak{F}=\langle W, D \rangle$ 任意 \mathcal{L} 框架。 ε 是对常项和谓词在 \mathfrak{F} 上的解释：对任意的常项 a 和谓词 P ， $\varepsilon(a)\in D^W$ ， $\varepsilon(P)\in \wp(D)^W$ 。 $\varepsilon(a)$ ， $\varepsilon(P)$ 又分别记作 a^ε 和 P^ε 。

定义 3.3 一个 \mathcal{L} 结构是一个二元组 $\langle \mathfrak{F}, \varepsilon \rangle$ (简称结构，记作 \mathfrak{G})，其中 \mathfrak{F} 是一个框架， ε 是 \mathcal{L} 常项和谓词在 \mathfrak{F} 上的解释。

定义 3.4 一个 \mathcal{L} 模型是一个二元组 $\langle \mathfrak{G}, \sigma \rangle$ (简称模型)，其中 $\mathfrak{G}=\langle \mathfrak{F}, \varepsilon \rangle$ 是一个 \mathcal{L} 结构， $\mathfrak{F}=\langle W, D \rangle$ 是一个 \mathcal{L} 框架。 σ 称为 \mathfrak{G} 上的指派 (简称指派)，是变元集到个体词涵义域的映射，即 $\sigma: \text{Var}\rightarrow D^W$ 。

设 $\mathfrak{F}=\langle W, D \rangle$ 是任意框架。结构 $\mathfrak{G}=\langle \mathfrak{F}, \varepsilon \rangle$ 又可记作 $\langle W, D, \varepsilon \rangle$ ；模型 $\mathfrak{M}=\langle \mathfrak{G}, \sigma \rangle$ 又可记作 $\langle \mathfrak{F}, \varepsilon, \sigma \rangle$ 和 $\langle W, D, \varepsilon, \sigma \rangle$ 。设 $\mathfrak{M}=\langle W, D, \varepsilon, \sigma \rangle$ 是任意的模型，以下用 $W_{\mathfrak{M}}$ ， $D_{\mathfrak{M}}$ ， $\varepsilon_{\mathfrak{M}}$ ， $\sigma_{\mathfrak{M}}$ 等表示该模型中的相应部分，用 $\mathfrak{F}_{\mathfrak{M}}$ ， $\mathfrak{G}_{\mathfrak{M}}$ 表示相应的框架和结构。

定义 3.5 设 \mathfrak{M} 是任意的模型， $s\in D_{\mathfrak{M}}^{W_{\mathfrak{M}}}$ 。 $\sigma_{\mathfrak{M}(s/x)}$ 是 $\sigma_{\mathfrak{M}}$ 关于变元 x 的变体，当且仅当， $\sigma_{\mathfrak{M}(s/x)}: \text{Var}\rightarrow D_{\mathfrak{M}}^{W_{\mathfrak{M}}}$ ，并且满足以下条件：

$$\sigma_{\mathfrak{M}(s/x)}(y) = \begin{cases} s, & \text{如果 } y=x \\ \sigma_{\mathfrak{M}}(y), & \text{否则} \end{cases}$$

以 $\sigma_{\mathfrak{M}(s/x)}$ 为指派的模型记作 $\mathfrak{M}_{(s/x)}$ 。

定义 3.6 设 \mathfrak{M} 是任意的模型， t 是任意的项。 t 在模型 \mathfrak{M} 中的解释 $t^{\mathfrak{M}}$ 是满足以下条件的映射：

$$t^{\mathfrak{M}} = \begin{cases} \varepsilon_{\mathfrak{M}}(a), & \text{如果 } t=a \\ \sigma_{\mathfrak{M}}(x), & \text{如果 } t=x \end{cases}$$

定义 3.7 设 $\mathfrak{M}=\langle W, D, \varepsilon, \sigma \rangle$ 是任意的模型。对任意的 $w\in W$ ，

- (1) $\mathfrak{M}, w \models Pt$ ，当且仅当， $t^{\mathfrak{M}}(w)\in P^\varepsilon(w)$
- (2) $\mathfrak{M}, w \models t=t'$ ，当且仅当， $t^{\mathfrak{M}}(w)=t'^{\mathfrak{M}}(w)$
- (3) $\mathfrak{M}, w \models t\equiv t'$ ，当且仅当， $t^{\mathfrak{M}}=t'^{\mathfrak{M}}$
- (4) $\mathfrak{M}, w \models \neg\alpha$ ，当且仅当， $\mathfrak{M}, w \not\models \alpha$
- (5) $\mathfrak{M}, w \models \alpha\rightarrow\beta$ ，当且仅当， $\mathfrak{M}, w \not\models \alpha$ ，或 $\mathfrak{M}, w \models \beta$
- (6) $\mathfrak{M}, w \models \forall x\alpha$ ，当且仅当，对任意的 $s\in D_{\mathfrak{M}}^{W_{\mathfrak{M}}}$ ， $\mathfrak{M}_{(s/x)}, w \models \alpha$

定义 3.8 设 $\mathfrak{M}=\langle W, D, \varepsilon, \sigma \rangle$ 是任意的模型， α 是任意的公式。对任意的 $w\in W$ ，称 α

在 w 上是真的，如果 $\mathfrak{M}, w \models \alpha$ ； α 在模型 \mathfrak{M} 上是有效的（也称 α 模型有效），记作 $\mathfrak{M} \models \alpha$ ，当且仅当，对任意的 $w \in W$ ， $\mathfrak{M}, w \models \alpha$ ； α 是有效的，记作 $\models \alpha$ ，当且仅当，对任意的模型 \mathfrak{M} ， $\mathfrak{M} \models \alpha$ 。

命题 3.1 设 \mathfrak{M} 是任意的模型， t, t' 是任意的个体词。

(a) $\mathfrak{M} \models t \equiv t'$ ，当且仅当，存在该模型中的可能世界 w ， $\mathfrak{M}, w \models t \equiv t'$ 。

(b) $\models t \equiv t' \rightarrow t = t'$ ，并且 $\not\models t = t' \rightarrow t \equiv t'$

证明从略。

由以上定义给出的形式语义称为涵义语义。由定义 3.7 可以看出，在涵义语义中，两个个体词的涵义相同是两个函数相同（3），而指称只能在可能世界中相同，是两个函数在相应变目下的函数值相同（2）。函数相同与某函数值相同有完全不同的性质。在涵义语义中，还表现为有命题 3.1(a) 所表示的性质：两个涵义（函数）相同与否其实与可能世界无关。这表明，涵义相同是对模型来说的，不是对可能世界来说的。与此相对照，指称相同是对可能世界来说的。涵义相同和指称相同属不同层次，有不同的语义性质。

结合命题 3.1(a)，由命题 3.1(b) 还可以看出，如果两个个体词的涵义相同，那么它们在任意可能世界中的指称相同，反之，在某可能世界中两个个体词的指称相同并不意味着它们的涵义相同，或者说，两个个体词的涵义不同，也可以在某世界上指称相同。这与直观上涵义与指称的关系相吻合。“晨星”和“昏星”的涵义不同，但是它们可以在某一个世界上指称相同。

以上结果表明，涵义语义是对于弗雷格理论所主张的个体词涵义和指称关系的形式刻画，符合其理论要点（1.1）和（1.2），但是还不涉及要点（1.3），因此还不足以在认知层面和逻辑层面考虑弗雷格谜题。这就是概念语义所要解决的问题。¹⁰

4. 概念语义：涵义域

时至今日，关于概念理论并不完善，从不同的视角、学科背景出发，对于概念有不同的观点和理论。这里列出关于概念的一些观点，后面将根据这些观点给出一个关于概念的形式语义学。这些观点（以下称为概念观点）有：¹¹

（1）概念有内涵和外延。

（2）概念由语词或词组等语言形式表达。

¹⁰ 可以说弗雷格关于涵义与指称有比较多的讨论和相对丰富的理论，但这还只是限于语言层面。在认知层面，弗雷格没有像在语言层面那样提供可以依据的理论，这是需要发展的地方。如果说涵义语义还有弗雷格理论的背景，那么，本文以下关于概念的理论及形式刻画等，是超出弗雷格理论的部分。弗雷格也讨论了概念，但是他所说的概念是与他的专用术语“思想”相关的概念，与本文所说的个人认知意义下的概念不同。

¹¹ 这些观点的理由可参见周北海（2004），Zhou and Mao（2010）。

(3) 概念的外延是概念所涉及的对象，也是表达概念的语词的指称。

(4) 概念的内涵，从思想层面看，是一组关于其外延的属性等方面的认识或知识；从语言的语形层面看，由一组相应的语言表达式，如语词或词组；从语言的语义的层面看，是这些语言表达式涵义。所以，内涵也可以视为涵义集。

(5) 一个概念在不同的语境（或环境等，也可以称为可能世界）下可以有不同的内涵，所以，概念是可能世界到内涵的映射。给定一个可能世界，可以得到该概念的当下内涵，即一个涵义集，其中的元素又称为内涵项。

(6) 不同的认知主体对于相同的对象（外延）可能会有不同认知或知识，这导致内涵方面的不同，由此带来了概念的个人性。比如，一个普通人关于老虎的概念与一个动物学家关于老虎的概念可能是不同的。这个不同主要不是外延的不同，而是关于老虎的认识或知识不同，是内涵方面的不同。

根据以上观点，可以在涵义语义上通过增加一些内容构建关于概念的形式语义学。首先考虑涵义域。因为概念是可能世界到内涵的映射，内涵又是由涵义构成，所以涵义域是关于概念的形式语义学的基础。

在涵义语义中，涵义是可能世界到指称的映射。个体词、谓词的赋值首先是涵义。不同的类型的语词的涵义表现为不同类型的映射。个体词、谓词的涵义分别是可能世界到个体域、个体域的幂集的映射。因此，它们的涵义域分别为： D^W ， $\wp(D)^W$ ¹²。对于任意框架 $\mathfrak{F}=\langle W, D \rangle$ ，集合 $D^W \cup \wp(D)^W$ 称为该框架的框架涵义域，记作 $SEN_{\mathfrak{F}}$ 或 SEN 。这是涵义语义已经提供的基础。

设 $\mathfrak{F}=\langle W, D \rangle$ 是任意框架，设 ε 是 \mathfrak{F} 上的某个解释。对于所解释的语言 \mathcal{L} 来说，结构 $\mathfrak{G}=\langle \mathfrak{F}, \varepsilon \rangle$ 可以视为该语言非逻辑常项的一个标准语义。令 $SEN_{\varepsilon}=\{X^{\varepsilon} : X \in CT\}$ 。 SEN_{ε} 是在框架 \mathfrak{F} 上基于解释 ε 所得到的所有 \mathcal{L} 非逻辑常项的涵义的集合，称为 \mathcal{L} 在结构 $\langle \mathfrak{F}, \varepsilon \rangle$ 上的标准涵义域，也简称结构涵义域。可以看出，在给定框架的情况下，由不同的解释 ε 和 ε' 得到的结构涵义域 SEN_{ε} 和 $SEN_{\varepsilon'}$ 会有相应不同，而且，对任意的解释 ε ，都有 $SEN_{\varepsilon} \subseteq SEN$ ，其中 SEN 是该给定框架的框架涵义域。

语言具有公共性，要求所有同一语言的使用者都要按标准语义使用语言，在这里的体现就是要求都要按 SEN_{ε} 所确定的意思使用语言。但是，另一方面，语言的使用者并不一定都能知道或掌握这个标准语义。这里将此种情况称为语言的个人性。 SEN_{ε} 体现了语言的公共性或标准性，个人性的情况在形式语义中也应该得到反映。

设 a 是一个语言的使用者，也是一个认知主体。设 X 是一个语词，在结构 $\langle \mathfrak{F}, \varepsilon \rangle$ 中的解释为 X^{ε} (SEN_{ε})，这就是 X 的标准语义。现将 a 所在结构 $\langle \mathfrak{F}, \varepsilon \rangle$ 中所知道的 X^{ε} 记为 $[X^{\varepsilon}]_a$ 。 $[X^{\varepsilon}]_a$ 与 X^{ε} 有可能不同。这个不同可以有以下一些情况：(1) 受认知能力所限或其

¹² 一般地，以 A 和 B 分别为定义域和值域的所有函数的集合记为 B^A 。

他原因，a 没有按 ε 给出的标准意思 X^ε 理解或使用 X （即不同的函数）；（2）a 知道 X ，但是不知道 X^ε （即 $[X^\varepsilon]_a$ 没有定义），比如我们听说过某个词，但是不知道这个词的意思是什么；（3）a 不知道 X （ X 不在 a 的词汇列表中），比如我们从来没有听（见）过某个词。这些情况表明，标准涵义域还不是具体认知主体的涵义域。认知主体的涵义域应该是该主体所知道的所有涵义的集合，一个他所不知道的涵义不能成为其涵义域中的元素。

在考虑语言及其给定解释的情况下，一个主体 a 知道涵义 X^ε 具有两个条件，a 不仅知道 X^ε 本身，还需知道 X 这个语言表达式，a 是通过 X 知道 X^ε 的，应该是语形和语义两方面二者的统一。以下用 $|X^\varepsilon|$ 表示这种形义的统一体，称为涵义 X^ε 的涵义名。涵义名增加了语形的因素。两个涵义名相同，如 $|X^\varepsilon| = |Y^{\varepsilon'}|$ ，当且仅当， $X = Y$ 并且 $\varepsilon = \varepsilon'$ 。设 X 和 Y 是两个不同的语词（ $X \neq Y$ ）。在某个解释 ε 下，可以有 $X^\varepsilon = Y^\varepsilon$ ，二者是同一个函数，但是仍然有 $|X^\varepsilon| \neq |Y^\varepsilon|$ ，即二者涵义名不同。在上述考虑下，主体 a 知道某个用 X 表达的涵义，可以视为 a 知道该涵义的涵义名 $|X^\varepsilon|$ 。

以下用 SEN_{ε_a} 表示在给定结构 $\langle \mathfrak{F}, \varepsilon \rangle$ 上 a 的涵义域，直观上 SEN_{ε_a} 是主体 a 所知道的所有涵义的集合，称为主体涵义域。根据以上讨论，首先，a 所知道的涵义应该是形义的统一，所以可设 SEN_{ε_a} 是涵义名的集合。其次，出于尽量简化问题，在此假设每个主体可以不知道某个语词的涵义，但是如果知道就是标准的¹³。在此假设下，主体涵义域和标准涵义域有关系：如果 $|X^\varepsilon| \in SEN_{\varepsilon_a}$ ，那么 $X^\varepsilon \in SEN_\varepsilon$ 。最后，综合上述分析和考虑，可以设主体 a 的涵义域为： $SEN_{\varepsilon_a} = \{|X^\varepsilon| : X^\varepsilon \in SEN_\varepsilon \text{ 并且 } X \in CT_a\}$ ，其中 $|X^\varepsilon|$ 是涵义名， CT_a 是 a 的语词集： $X \in CT_a$ 即 a 知道 X 这个语词。

5. 概念语义：从直观到形式

至此得到三种涵义域：框架涵义域 SEN ，结构涵义域 SEN_ε 和主体涵义域 SEN_{ε_a} 。框架涵义域是给定框架上所有逻辑可能的涵义组成的涵义域，也是变元的论域，因此也可称为逻辑涵义域。结构涵义域是相对于各主体来说的标准或公共涵义域。主体涵义域是认知主体所实际使用的涵义域。概念都是相对于某个认知主体的概念，主体涵义域是得到概念及其内涵的基础。下面在主体涵义域的基础上，结合前节所列的概念观点，作以下补充，也是对其中一些观点更精细的表述。

（i）设 $\mathfrak{F} = \langle W, D \rangle$ 是任意框架， $\mathfrak{G} = \langle \mathfrak{F}, \varepsilon \rangle$ 是 \mathfrak{F} 上的任意 \mathcal{L} 结构，a 是某一认知主体。由此可得 a 的主体涵义域 SEN_{ε_a} 及其幂集 $\wp(SEN_{\varepsilon_a})$ 。 SEN_{ε_a} 可以视为所有 a 所知道的涵义的集合。根据概念观点（4），内涵是涵义的集合，所以 $\wp(SEN_{\varepsilon_a})$ 就是 a 的内涵域，记作 INT_{ε_a} 。

¹³ 此举针对的是标准语义 X^ε 和个人理解 $[X^\varepsilon]_a$ 有可能不同的情况（1）和（2），避免对 $[X^\varepsilon]_a$ 进行过多讨论。

(ii) “概念是可能世界到内涵的映射”，应该更为严格地表述为，在框架 \mathfrak{F} 中，主体 a 的一个概念 c 是一个从 W 到 $\text{INT}_{\varepsilon_a}$ 的映射，即 $c:W \rightarrow \text{INT}_{\varepsilon_a}$ ，或 $c \in \text{INT}_{\varepsilon_a}^W$ 。

(iii) 根据概念观点 (3)，概念的外延“也是表达概念语词的指称”，在涵义语义的基础上，指称由涵义确定，对此可以视为每个概念都由某个涵义生成，以保证每个概念都有相应的外延：设 $|s| \in \text{SEN}_{\varepsilon_a}$ 是任意的涵义名， \mathcal{C} 是从 $\text{SEN}_{\varepsilon_a}$ 到 $\mathcal{P}(\text{SEN}_{\varepsilon_a})^W$ 的映射。对任意的 $|s| \in \text{SEN}_{\varepsilon_a}$ ， $\mathcal{C}(|s|)$ 称为由 $|s|$ 生成的概念，满足对任意的 w ， $|s| \in \mathcal{C}(|s|)(w)$ 。 $\mathcal{C}(|s|)$ 也记为 $|s|^{\mathcal{C}}$ 。如果 c 是由 $|s|$ 生成的概念，则 s 在 w 中的指称就是 c 在 w 中的外延。另外，根据 s 是个体词涵义 ($s \in D^W$) 或谓词涵义 ($s \in \mathcal{P}(D)^W$)，由 s 生成的概念 c 相应地也称为个体概念和类概念。¹⁴

(iv) 根据概念观点 (6)，概念具有个人性，因此概念的生成映射也应该因人而异，关于不同主体的概念映射应该可以不同。例如，如果考虑主体 a ， b ，相应地有概念生成映射 \mathcal{C}_a ， \mathcal{C}_b ，且二者可以不同。

现在可以考虑概念语义结构的基本“形状”。

根据上述讨论，一个概念语义的结构已经不是一个独立于语言的使用者（也是认知主体）的结构，而是一个架构在涵义语义基础之上同时又带有个人特点的结构，因此已经不存在原来独立于主体的结构，涵义域、内涵域和概念生成映射也应该做相应的调整：一个概念语义的结构是一个多元组 $\langle W, D, \varepsilon, A, \mathcal{S}, \{ \mathcal{C}_a : a \in A \} \rangle$ ，其中 $\langle W, D \rangle$ 是一个涵义语义框架， $\langle W, D, \varepsilon \rangle$ 是一个涵义语义结构。在此之上有框架涵义域 SEN ，结构涵义域 SEN_{ε} 。 A 是主体集。对任意的 $a \in A$ ，由于对于语言掌握的个人性， a 可以有与标准语义不同的涵义域。在技术上，可以由 \mathcal{S} 对每个主体分配一个涵义域来实现，使得 $\mathcal{S}(a) = \text{SEN}_{\varepsilon_a}$ ，称为主体涵义域映射。此后，还可以对任意的 $a \in A$ 得到 a 的内涵域 $\text{INT}_{\varepsilon_a} = \mathcal{P}(\text{SEN}_{\varepsilon_a})$ ，以及关于 a 的概念生成映射 \mathcal{C}_a 。 \mathcal{C}_a 是从 $\text{SEN}_{\varepsilon_a}$ 到 $\text{INT}_{\varepsilon_a}^W$ 的映射。对任意的 $|X^{\varepsilon}| \in \text{SEN}_{\varepsilon_a}$ ， $\mathcal{C}_a(|X^{\varepsilon}|) \in \text{INT}_{\varepsilon_a}^W$ 是一个函数。 $\mathcal{C}_a(|X^{\varepsilon}|)$ 也记作 $|X^{\varepsilon}|^{\mathcal{C}_a}$ 。对任意的 w ， $|X^{\varepsilon}|^{\mathcal{C}_a}(w)$ 本身是一个涵义名的集合，也可以视为一个涵义集，这就是在 w 上主体 a 用语词 X 表达的概念的内涵（“ a 用语词 X 表达的概念”以后简称为“ a 的概念 X ”）。

最后，还需对“知道”作一些说明。“知道”有多种意思，可以是单纯的知道或仅仅知道，比如，仅仅是知道有某个说法，但是并不接受，或不相信。还有一种知道，是具有接受意味的知道，知道什么意味相信什么或认定什么。这种知道将所知道的事情当作真实的情况，将表达知道什么的语句当作真语句，甚至将所知道的说法当作知识而加以接受。与弗雷格谜题有关的知道，就其本意来说，是具有接受意味的知道，既然弗雷格谜题涉及“ $a=b$ 提供新知识”这样的问题。在这个意义上，知道什么与概念内涵的丰富性有关，为从概念

¹⁴ 根据通名的四层语义理论 (Zhou and Mao, 2010)，通名（谓词）可以表达四种意思：类，涵义，概念，内涵。这里将同样的理解推广到了个体词上。

的形式语义学方面寻找弗雷格谜题认知现象的解释提供了基本的依据。

6. 概念语义：一个关于概念的形式语义

形式语言 \mathcal{L}^* \mathcal{L}^* 是 \mathcal{L} 的扩张。在 \mathcal{L} 的基础上，增加符号： K, a, b ；增加公式： $K_a\alpha, K_b\alpha$ ，其中 α 是形如 Pt 和 $t=t'$ 的 \mathcal{L} 公式。 $K_a\alpha$ 读作“ a 知道 α ”， $K_b\alpha$ 类似。

为了简明地说明问题， \mathcal{L}^* 是一个用于讨论弗雷格谜题的尽可能小的语言，其中只设有两个表示认知主体的符号 a 和 b ，可以类似地增加更多的表示认知主体的符号。对形如 $K_a\alpha$ 公式的限制，意味 \mathcal{L}^* 中没有形如 $K_a(\alpha \rightarrow \beta)$ ， $K_a\neg\alpha$ 和 $K_a\forall x\alpha$ 的公式，意为这里不讨论知道复合命题的情况； \mathcal{L}^* 中也没有形如 $K_a(t \equiv t')$ 的公式，意为在这里也不讨论“ a 知道两个个体词的涵义相同”的情况。

形式语言 \mathcal{L}^* 的语义

\mathcal{L}^* 的语义是 \mathcal{L} 的语义扩张，在 \mathcal{L} 的语义的基础上增加以下内容。

定义 5.1 设 $\mathfrak{F} = \langle W, D \rangle$ 是任意 \mathcal{L} 框架， ε 和 ε' 是 \mathfrak{F} 上的任意解释。对任意的 $X, Y \in CT$ ， $|X^\varepsilon|$ 称为涵义名； $|X^\varepsilon| = |Y^{\varepsilon'}|$ ，当且仅当， $X =_{\text{syn}} Y$ ¹⁵ 并且 $\varepsilon = \varepsilon'$ 。

定义 5.2 设 $\mathfrak{F} = \langle W, D \rangle$ 是任意 \mathcal{L} 框架， $\mathfrak{G} = \langle \mathfrak{F}, \varepsilon \rangle$ 是 \mathfrak{F} 上的任意 \mathcal{L} 结构， $i \in \{a, b\}$ 。

(1) 令 $\text{SEN}_{\mathfrak{F}} = D^W \cup \wp(D)^W$ ，称为框架 \mathfrak{F} 的框架涵义域，简记为 SEN 。令 $\text{SEN}_{\varepsilon} = \{X^\varepsilon : X \in CT\}$ ，称为结构 \mathfrak{G} 的结构涵义域。

(2) 设 S 是一个从 $\{a, b\}$ 到 $\wp(CT)$ 的映射。令 $\text{SEN}_{\varepsilon_i} = \{|X^\varepsilon| : X^\varepsilon \in \text{SEN}_{\varepsilon} \text{ 并且 } X \in S(i)\}$ ，称为在结构 \mathfrak{G} 上 i 的涵义域，也称为主体涵义域。

(3) 令 $\text{INT}_{\varepsilon_i} = \wp(\text{SEN}_{\varepsilon_i})$ ，称为结构 \mathfrak{G} 上 i 的内涵域。

该定义中映射 S 是从主体集到常项幂集的映射，通过 S 得到主体涵义域，所以 S 也可称为主体涵义域映射。

定义 5.3 一个 \mathcal{L}^* 结构 \mathfrak{G}^* 是一个 \mathcal{L} 结构的扩张， $\mathfrak{G}^* = \langle \mathfrak{F}, \varepsilon, S, C_a, C_b \rangle$ ，其中 $\mathfrak{F} = \langle W, D \rangle$ 是一个 \mathcal{L} 框架， $\mathfrak{G} = \langle \mathfrak{F}, \varepsilon \rangle$ 是一个 \mathcal{L} 结构，也称为 \mathfrak{G}^* 的基础结构。 S 是 \mathfrak{G} 上的主体涵义域映射。对任意的 $i \in \{a, b\}$ ， C_i 是从 $\text{SEN}_{\varepsilon_i}$ 到 $\text{INT}_{\varepsilon_i}^W$ 的映射，称为 i 的概念生成映射，满足对任意的 $|X^\varepsilon| \in \text{SEN}_{\varepsilon_i}$ 和任意的 $w \in W$ ， $|X^\varepsilon| \in X^\varepsilon C_i(w)$ ， $|X^\varepsilon|^{C_i}$ 即 $C_i(|X^\varepsilon|)$ 。¹⁶

基础结构 \mathfrak{G} 是语言的标准语义，体现了语言的公共性。在此之上的主体涵义域映射 S 体现了对主体在语言掌握方面的个人性。概念生成映射体现了主体在认知方面的个人性。

定义 5.4 一个 \mathcal{L}^* 模型是一个二元组 $\langle \mathfrak{G}^*, \sigma \rangle$ ，其中 \mathfrak{G}^* 是一个 \mathcal{L}^* 结构， σ 是 \mathfrak{G}^* 的基础结构上的指派。

\mathcal{L}^* 结构、模型分别是 \mathcal{L} 结构、模型增加概念生成映射得到的结构和模型。概念生成

¹⁵ “ $X =_{\text{syn}} Y$ ”表示 X 和 Y 是同一个符号，在语形上相同。

¹⁶ 前面给出的概念语义的结构 $\langle W, D, \varepsilon, A, S, \{C_a : a \in A\} \rangle$ 是一般形式，这里所给的结构是相对于语言 \mathcal{L}^* 的简化形式，可以视为该一般形式的特例。

映射是架构在 \mathcal{L} 结构、模型之上的映射，并不影响改变底层的设置，比如不改变常项的解释，也不改变变元的解释。这意味，如果把概念看作对事物的认识，那么我们的认识如何不改变事实和语言层面的东西。

定义 5.5 设 $\mathcal{M} = \langle W, D, \varepsilon, \mathcal{S}, C_a, C_b, \sigma \rangle$ 是任意的 \mathcal{L}^* 模型， $i \in \{a, b\}$ 。对任意的 $w \in W$,

(1) $\mathcal{M}, w \models K_i P t$ ，当且仅当， $|t^\varepsilon| \in \text{SEN}_{\varepsilon_i}$ 并且 $|P^\varepsilon| \in |t^\varepsilon|^{C_i(w)}$ 。

(2) $\mathcal{M}, w \models K_i (t = t')$ ，当且仅当， $t =_{\text{syn}} t'$ ，或者， $|t^\varepsilon|, |t'^\varepsilon| \in \text{SEN}_{\varepsilon_i}$ 并且 $|t'^\varepsilon| \in |t^\varepsilon|^{C_i(w)}$ 且 $|t^\varepsilon| \in |t'^\varepsilon|^{C_i(w)}$ 。

(3) $\mathcal{M}, w \models \alpha$ ，当且仅当， $\mathfrak{M}, w \models \alpha$ ，其中 α 是形如 $P t$ ， $t = t'$ 和 $t \equiv t'$ 的公式， $\mathfrak{M} = \langle W, D, \varepsilon, \sigma \rangle$ 是由 \mathcal{M} 中相应部分得到的 \mathcal{L} 模型。

$\neg \alpha$ ， $\alpha \rightarrow \beta$ ， $\forall x \alpha$ 的赋值由常规定义可得，从略。

模型有效 ($\mathcal{M} \models \alpha$)，有效 ($\models \alpha$) 也由常规定义可得，从略。

7. 在认知层面看弗雷格谜题

概念语义可以用于在认知层面对弗雷格理论给出形式刻画，表述要点 (1.3)，以在认知层面讨论弗雷格谜题。

在该语义中，在可能世界 w 上主体 a 知道晨星是昏星，被理解为，在可能世界 w 上， a 的“晨星”概念的内涵中有“(是)昏星”(即昏星 $^\varepsilon$)这一内涵项，相应的形式表示是， $|昏星^\varepsilon| \in |晨星^\varepsilon|^{C_i(w)}$ 。如果 a 还知道昏星是晨星，相应地有 $|晨星^\varepsilon| \in |昏星^\varepsilon|^{C_i(w)}$ 。

在可能世界 w 上主体 b 不知道晨星是昏星，有两种情况：(7.1) b 知道晨星，也知道昏星，但是不知道它们是同一个星（不知道“晨星”和“昏星”的指称是同一个对象）；(7.2) b 不知道晨星，或者不知道昏星，至少有其中之一，因此不知道晨星就是昏星。

对此首先需要做些说明。设在标准语义下语词“晨星”表示某个体 d 。“ b 知道晨星”有三种理解：(1) b 知道语词“晨星”（尽管这并不是正常的理解）；(2) b 不知道语词“晨星”，但是知道个体 d （通过其他语言表达知道 d ，甚至没有通过语言，只是见过 d ）；(3) b 通过“晨星”和 d 的表示关系知道 d ，这包括 b 知道语词“晨星”，也知道语词“晨星”所表示的对象 d 。理解 (3) 的要点是强调 b 必须通过“晨星”和 d 的表示关系知道 d 。在弗雷格谜题讨论中，“ b 知道晨星”所应该采取的理解是 (3)。既然如此，相应的否定句“ b 不知道晨星”的意思应该是 b 没有通过“晨星”和 d 的表示关系知道 d 。这包括几种情况：(1-3) b 不知道语词“晨星”；(2-3) b 知道语词“晨星”但是不知道 d ；(3-3) b 知道“晨星”也知道 d ，但是不知道“晨星”表示 d 。这三种情况都可以使理解 (3) 意义下的“ b 知道晨星”不成立。

在概念语义中，“ b 知道晨星”是形式表示是 $|晨星^\varepsilon| \in \text{SEN}_{\varepsilon_b}$ 。于是关于情况 (7.1) 有： $|晨星^\varepsilon|, |昏星^\varepsilon| \in \text{SEN}_{\varepsilon_b}$ ，但是 $|昏星^\varepsilon| \notin |晨星^\varepsilon|^{C_i(w)}$ 或 $|晨星^\varepsilon| \notin |昏星^\varepsilon|^{C_i(w)}$ ，根据定义，由此可得 $\mathcal{M}, w \not\models K_b(\text{晨星} = \text{昏星})$ ，即“ b 不知道晨星是昏星”。

关于情况 (7.2)，不妨考虑“ b 不知道晨星”。对此有三种情况，根据 $\text{SEN}_{\varepsilon_b}$ 的定义，对于 (1-3) 和 (2-3) 都有 $|晨星^\varepsilon| \notin \text{SEN}_{\varepsilon_b}$ ，(3-3) 可以视为 b 没有按标准语义掌握“晨星”，这

也使得 $|晨星^\varepsilon| \notin SEN_{\varepsilon_b}$ 。于是，由定义 5.5 (2)，有 $\mathcal{M}, w \not\models K_b(\text{晨星} = \text{昏星})$ 。

$a = a$ 是一个特殊情况。这里假定，对任意的认知主体，比如不论是否知道晨星，甚至之前是否听说过“晨星”这个词，总知道晨星是晨星（认为“晨星是晨星”为真），这由定义 5.5 (2) 中的真值条件“ $t = t'$ ”体现。

概念语义是涵义语义的扩张，底层是涵义语义部分，关注语言和事实的描述，是客观的、公共的部分，上层是概念、内涵等形式刻画部分，关注与认知主体个人认知有关的概念和相关知识的描述，上下层既有关联又相对分离。在概念语义中，可以描述晨星等于或不等于昏星的事实，也可以描述某一主体关于晨星和昏星的概念和知识，从而使得事实情况和认知情况可以分层描述，达到弗雷格谜题在认知层面的自然消解。根据上述定义，在概念语义中可以得到，存在模型 \mathcal{M} ，以及其中的可能世界 w ，使得

$$\mathcal{M}, w \models \text{晨星} = \text{昏星}, \mathcal{M}, w \models K_a(\text{晨星} = \text{昏星}), \mathcal{M}, w \not\models K_b(\text{晨星} = \text{昏星})$$

由此可得前面说的 (3a), (3b) 和 (4a)。其中 (4a) 也是与关于同一替换律问题所希望得到的结果。于是，弗雷格谜题的消解从认知层次自然过渡到逻辑层次。

从字面上看，弗雷格谜题（问题 (1)）可以理解为追问 $a = a$ 和 $a = b$ 为什么会有认知上差异的深层次哲学原因。这里只是通过严格的方法表达出从而也是说明“知道 $a = a$ ”和“知道 $a = b$ ”在认知方面究竟有什么差异。所得到的结果是：只需依赖语形，我们就可以知道 $a = a$ ，但是知道 $a = b$ 则要依赖更多的方面，要有关于事实和语言方面的相应知识等，这同时也是在说明“ $a = a$ ”和“ $a = b$ ”提供了什么，有什么不同的认知价值，对此概念语义给出了具体的细节。这就是“发展路线”在认知层面对弗雷格谜题的消解。

8. 同一替换律问题

在可能世界语义学式的语义学中有不同层次的同一替换律，由它们分别可以得到下面的表达式：

$$A \quad \models a = b \rightarrow K_a(a = b)$$

$$B \quad \mathcal{M} \models a = b \Rightarrow \mathcal{M} \models K_a(a = b)$$

$$C \quad \models a = b \Rightarrow \models K_a(a = b)$$

所谓在同一替换律问题上消解弗雷格谜题，就是给出使得这些表达式不成立的形式语义学。这些表达式从内容上看，按所列顺序，依次从强到弱。从消解难度上看，与强弱成反比。前面已说明，可能世界语义学可以消解 A，但是无法消解 B 和 C。

在概念语义上，(4a) 自然成立，所以 A 也已经自然消解。只需考虑 B 和 C。

考虑 B。要说明 B 在概念语义中不成立，只需构造模型，使得 (1) $\mathcal{M} \models a = b$ 且 (2) $\mathcal{M} \not\models K_a(a = b)$ ，其中 $a \neq_{\text{syn}} b$ 。设 $\mathcal{M} = \langle W, D, \varepsilon, S, C_a, C_b, \sigma \rangle$ ， $a^\varepsilon = b^\varepsilon$ ， $S(i) = \{a, b\}$ ， $|a^\varepsilon|^{C_a(w)} = \{|a^\varepsilon|\}$ ，其他任意。关于 (1)，由 $a^\varepsilon = b^\varepsilon$ ，可得 $\mathcal{M} \models a = b$ 。关于 (2)，由 $|a^\varepsilon|^{C_a(w)} = \{|a^\varepsilon|\}$ 和 $a \neq_{\text{syn}} b$ 可得 $|b^\varepsilon| \notin |a^\varepsilon|^{C_a(w)}$ 。由此可得 $\mathcal{M} \not\models K_a(a = b)$ 。

考虑 C。 $\models a = b$ ，当且仅当，对任意结构中的解释 ε 和任意可能世界 w ，都有

$a^{\epsilon}(w)=b^{\epsilon}(w)$, 即 a 和 b 无论怎么解释 $a=b$ 都真, 这可以视为 a 和 b 逻辑等同。从形式语义学的角度看, 只要是由不同初始符号给出的个体词, 就可以有不同的解释。因此不存在两个这样的个体词是逻辑等同的。在现实中通常也不存在这类专名。如果根本就没有这样的个体词, C 就是成立的。从这个意义上说, C 无法得到消解。当然, 这种成立只是空成立, C 也因此是一个无用因而也是无害的同一替换律。但是这并不是消解 C 的理由。从直观上看, 合理的解释是, 即使事实上逻辑等同的 a 和 b , 某一主体也可以不知道二者等同。在概念语义中可以表达的是, 即使 a 和 b 事实上是逻辑等同的, 只要它们本身是不同的符号, 在概念语义中就可以加以区分, 使得某主体不知道它们是等同的情况得以表达。对此克里普克的信念谜题提供了一个有趣的例子 (Kripke,1979), 从中可以得到:

法国人皮埃尔 (Pierre) 根据传说相信隆德 (法语伦敦 Londres 的音译) 是一座漂亮的城市, 而根据自己的亲身体验知道伦敦 (London) 不是漂亮的城市。皮埃尔不知道伦敦就是隆德。

关于这个例子, 在概念语义中可以构造模型使得 $|London^{\epsilon}| \not\subseteq |Londres^{\epsilon}|^{C_{\text{Pierre}}(w)}$, 由此可得 $\not\models K_{\text{Pierre}}(\text{London}=\text{Londres})$ 。不过此例可以用于 C 的消解还依赖于“London”和“Londres”必须逻辑等同。从直观上看, 这一点可以由翻译关系保证, 即英语“London”和法语“Londres”因翻译关系要求它们在“任何情况下”都被赋予同样的语义, 从而使得“London”和“Londres”逻辑等同。上面说“在现实中通常也不存在这类专名”是局限于某个自然语言语言 (如汉语、英语) 之内而言的。就语言 \mathcal{L}^* 的设定来说, 还没有考虑自然语言中的翻译因素, 所以目前的概念语义还不能完全用此例消解 C 。本文的重点是给出弗雷格理论基础上的概念语义, 这个问题可以在增加相关形式表达后得到解决。有关翻译的形式处理以及个体词的逻辑等同问题需要另文专述。

同一替换律是逻辑演算中常用的规则。这里的问题不能得到合理的解决, 也就表明相应的逻辑还没有建立。正式因为如此, 有关内涵问题的逻辑研究一直在进行, 同一替换律问题成了这些研究成功与否的“试金石”。在这方面重要的成果有 Church、卡尔纳普、蒙太古和克里普克等人的工作。内涵语义和内涵逻辑有不同的类型, 其中可能世界语义学和模态逻辑因为在理论和应用方面都硕果累累而最为显赫。但是, 同一替换律问题表明可能世界语义学不是成功的内涵语义学, 模态逻辑也不是成功的内涵逻辑。同一替换律问题由此而被凸显, 成为新一代内涵逻辑研究的中心问题, 这就是七十年代中期出现的超级内涵逻辑。三十余年, 这一分支走向繁荣。据文学锋 (2009; 45-46) 总结, 目前解决超级内涵问题的方案已有十几种之多, 可以分为两种策略和四个主要方向: 外延模拟策略——非经典可能世界语义学, 结构式语义; 内涵初始策略——高阶超内涵逻辑, 一阶超内涵逻辑。方案多或许不是件好事, 因为这说明人们还在探索之中, 关于同一替换律问题还没有找到公认的解决办法。

按历史的发展线索, 概念语义应属于超级内涵语义学, 甚至从理论背景和所用术语看, 概念语义与外延模拟策略下的非经典可能世界语义学最为接近, 但是有本质不同。概念语

义从根本上说不是外延模拟，而是在力图回答什么是概念、内涵以及涵义的基础上给出相应的形式表达，重点是关于概念的形式刻画，以此到达关于一些内涵现象和相应的推理规律的解释。概念语义也不以解决同一替换律问题为中心，这个问题的解决只是一个“副产品”。如果一定要按上述方式确定某个归属，概念语义可以看作某种认知背景下的形式语义学，或某种认知语义学。虽然其中用到术语“可能世界”，但这里的可能世界只是语境、认知环境等意义上的可能世界，没有可能世界的本体论意义，也没有可能世界之间的可及关系，因此概念语义与用于模态逻辑的可能世界语义学其实已有实质性的不同。

同一替换律问题是弗雷格谜题在逻辑层面表象，根子在于我们关于概念、内涵，甚至涵义等这些“对象”还没有充分的认识，还不完全清楚它们究竟是什么。外延的同一替换律在内涵语境中应该失效，但这应该是解决了这些问题之后的自然结果，反之，在此之前要解决同一替换律问题，是不可能完成的任务。

对弗雷格谜题消解来说，解决同一替换律问题的意义或许可以这样看：从产生过程看，弗雷格谜题是由同一替换律引起的，因此弗雷格谜题的消解最后还应该回到同一替换律。

9. 结语

弗雷格谜题的形成有多方原因，直接指称论对弗雷格理论批评是主要原因之一，以至于可以说，这是产生于直接指称论哲学立场的谜题。尽管如此，弗雷格理论确有不足，比如什么是涵义不够清楚等。最重要的是，弗雷格理论只有关于涵义与指称的理论，即只有语言层面的理论，而缺少认知层面的理论。既然要谈“新知识”，讨论“认知方面的意义”，没有关于认知方面的理论是其根本性的不足。这个不足限定了弗雷格理论在形式化方面前进的幅度。结果是，同一替换律问题的解决难有令人满意的结果，让“谜题”多添了几分“迷”的色彩。

概念语义以逻辑学和哲学关于概念的理论为思想基础，具体地说，是在弗雷格理论的基础上增加了认知层面的理论，主要有关于概念、内涵、涵义及其它它们之间关系的理论，包括相关形式描述的讨论等。本文的重点在于给出基于这个理论的形式语义。理论上的扩展关系在形式语义的构建中也有形象的体现。概念语义由两层结构组成，底层是基于语言公共性的涵义语义结构 $\langle W, D, \varepsilon \rangle$ ，在此基础上，通过加入个人对于语言的掌握因素 \mathcal{S} （主体涵义域映射）和认知因素 \mathcal{C} （概念生成映射）等得到了概念语义结构 $\langle W, D, \varepsilon, \mathcal{S}, \mathcal{C}_a, \mathcal{C}_b \rangle$ 。弗雷格谜题涉及四个层面的问题，主要是语言与认知层面的问题。理清语言与认知的交互作用是形式语义得以建立的关键。

弗雷格谜题产生于同一替换律，因此问题一开始就被集中在同一性等相关问题上。在哲学方面，出现了各种关于同一性问题的讨论，甚至甚至引出了分析性、必然性、本质主义问题讨论等。在逻辑学方面，出现各种围绕同一替换律问题的内涵语义学和内涵逻辑。问题的提出是源于同一替换律，但问题的解决不应只围绕同一替换律，而应该从什么是涵义、什么是概念、什么是内涵这些根本性问题入手。同一替换律问题的解决应该是随着这

些问题解决而自然解决。概念语义的提出正是基于这样的考虑，原来的一些解决问题的思路和方式似有误区。¹⁷

概念语义是基于四层语义对象和两层结构考虑下得到的语义学，在具体问题的处理中，可以根据所要讨论的问题考虑其相应的简单形式，也可以根据所要讨论的问题增加相应的内容。例如，为方便讨论问题，这里考虑的是两个主体的语言 \mathcal{L}^* 以及相应的六元组式的结构 $\langle W, D, \varepsilon, S, C_a, C_b \rangle$ ，还可以考虑更多的主体及一般的概念语义框架。再如，对主体涵义域的设定，也用到了为简略而给出的假定，使得 $SEN_{\varepsilon_a} \subseteq SEN_{\varepsilon}$ 。根据不同的谜题和所要解决的问题，在形式语义学方面，概念语义可以做相应的增改。

概念语义使得弗雷格谜题可以在不同层次上得到合直观的消解，但是也还存在需要进一步解决的问题。例如，在逻辑层面目前的概念语义还只是消除了一些不该成立的同一替换律，还没有给出该成立的同一替换律，即还没有给出适用于内涵语境的同一替换律。一个语义理论的最终成功，还有赖于正面结果的建立。

参考文献

- [1] 王路 (2006), 《弗雷格哲学论著选辑》, 弗雷格著, 王路译, 商务印书馆。
- [2] 文学锋 (2009), 面向自然语言内涵性的 LNL: 语境内涵逻辑 (第三章), 《面向知识表示与推理的自然语言逻辑》, 鞠实儿等, 经济科学出版社, 2009。
- [3] 叶闯 (2007), 信念之谜弗雷格式解决的有效性分析, 《西南民族大学学报》(人文社科版), 2007 年第 11 期, 123-128。
- [4] 周北海 (1997), 《模态逻辑导论》, 北京大学出版社。
- [5] 周北海 (2004), 概称句本质与概念, 《北京大学学报》, 2004 年第 4 期, 20-29。
- [6] 周北海 (2008), 涵义语义与关于概称句推理的词汇逻辑, 《逻辑学研究》, 2008 年第 1 期, 38-49。
- [7] 周北海 (2010), 从剧本的视角看名与指称, 《哲学研究》, 2010 年第 1 期。
- [8] Bealer, G. (1993), A Solution to Frege's Puzzle, *Philosophical Perspectives*, Vol.7, Language and Logic (1993), pp. 17-60.
- [9] Fine, K. (2007), *Semantic Relationism*, Blackwell Publishing.
- [10] Kripke, S. (1979), A puzzle about belief, *Meaning and Use*, Margalit (ed.), D. Reidel Publishing Company, 254-257.
- [11] Mao, Yi and Zhou, Beihai (2007). Interpreting Metaphors in a New Semantic Theory of Concepts [C], *Lecture Notes in Artificial Intelligence* 4384 (T. Washio et al. Eds), JSAI 2006, 177-190, Springer-Verlag Berlin Heidelberg.
- [12] Salmon, N. (1986), *Frege's Puzzle*, The MIT Press.
- [13] Zhou, Behai, and Mao, Yi (2010), Four semantic layers of common nouns, *Synthese*, Vol 175, Issue 1, 47-68.

¹⁷ 与同一替换律问题有些类似，在概念语义中认知全能问题也自然得到消解，而这在模态逻辑式的认知逻辑中亦是一个难以解决的问题。