

山西大学
2008 届硕士学位论文

怪人假设与可设想性论证

——查尔默斯二维语义学探析

作者姓名 陈敬坤
指导教师 魏屹东教授
学科专业 外国哲学
研究方向 认知哲学
培养单位 哲学社会学学院
学习年限 2005 年 9 月—2008 年 6 月

二〇〇八年六月

Shanxi University

The dissertation for master degree of philosophy

**The Assumption of Zombie and the Conceivability
Argument: an Inquiry into Chalmers's
Two-Dimensional Semantics**

June. 2008

目 录

引言：	1
一 怪人的可设想性论证及其困难	4
1.1 怪人假设与物理主义	4
1.2 怪人假设和反物理主义论证	6
1.3 传统可设想性论证及其困难	9
二 基于二维语义学的可设想论证	15
2.1 二维主义的发端	15
2.2 早期二维主义	16
2.3 二维语义的核心观点	18
2.4 二维语义方法与可设想性论证	20
三 对查尔默斯论证的质疑和评述	26
3.1 怪人是否可想象	26
3.2 修正的 CP 论题是否成立	27
3.3 二维语义方法是否适用	29
结论	35
参考文献	36
致谢	40
附录	41

CONTENTS

Preface :	1
Chapter 1 Conceivability Argument of Zombie and its Difficulties.....	4
1.1 Assumption of Zombie and Physicalism.....	4
1.2 Assumption of Zombie and Arguments against Physicalism.....	6
1.3 Traditional Conceivability Argument and its Difficulties	9
Chapter 2 Conceivability Argument Based on Two-Dimensional Semantics	15
2.1 The Origin of Two-Dimensionalism	15
2.2 Early Two-Dimensionalism	16
2.3 The Core Thesis of Two-Dimensionalism	18
2.4 Two-Dimensional Approach and Conceivability Argument	20
Chapter 3 Question and Comment on the Argument of Chalmers.....	26
3.1 Is Zombie Conceivable ?	26
3.2 Is the Refined CP Thesis Tenable ?.....	27
3.3 Is Two-Dimensional Approach Applicable ?.....	29
Conclusion	35
Reference	36
Acknowledgements	40
Appendix	41

中文摘要

怪人假设的提出和物理主义的盛行密不可分，所谓怪人指的是这样一种系统，它们与普通人看起来完全没有差别，只是没有心灵或意识，或者说没有现象性质。显然，这样一种怪人在现实生活中是不存在的，但问题在于，哲学家们所设想的这种怪人是否具有某种程度的可能性呢？一些人认为，哪怕是最低限度的可能性，也会导致物理主义的破产从而不得不接受某种形式的二元论。很多哲学家认为这是官人概念最重要的意义。用怪人假设来反对物理主义将导致了具有普遍性的问题，比如可设想性和可能性以及先天性和必然性之间的关系。同时，怪人问题将导致认知上的困难：它重新提出了他心问题。

查尔默斯将怪人假设作为一种精致的可设想性论证重新提出，引起了热烈讨论。查尔默斯所反对的物理主义观点蕴含了怪人的不可能性，但直觉上我们感到怪人是可设想的，似乎也是可能的。在当代心灵哲学中，主要的反物理主义论证，如知识论证、解释鸿沟论证、模态论证等，和可设想性论证一样，都是从物理事实与现象事实之间存在着认知鸿沟这一前提出发，得出二者之间存在本体论鸿沟的结论。而物理主义者则坚持，认识论的前提并不能导出本体论的结论。查尔默斯对传统的可想象性论证进行修正，认为从认识论前提合理地推出本体论结论是可行的，其步骤大体是从认识论前提推出模态结论，进而达到本体论结论。这一步骤的关键是在认知领域和模态领域之间建立关联，查尔默斯认为二维语义方法能够很好地实现这一目的。如果查尔默斯是对的，那这就是对物理主义的一个严重挑战。

二维语义方法源于直接指称主义和描述主义之间的争论，可以追溯到克里普克对后天必然性的说明。二维主义区分了两种内涵，先天性和后天性、可能性和必然性都可以定义为两种内涵在可能世界的赋值。查尔默斯发展了认知的二维主义，用认知术语定义模态概念，从而为可设想性论证提供支持。近几年，二维主义的影响渐大，索姆斯对二维主义的描述主义倾向进行了细致批判，查尔默斯则认为他所主张的认知的二维主义能够很好地兼顾直接指称主义和描述主义。

本文着重分析查尔默斯的可想象论证，认为这种基于二维语义的可想象性论证仍然存在明显的缺陷，并不能真正驳倒物理主义。

关键词：怪人；物理主义；可设想性；可能性；二维主义

The Assumption of Zombie and the Conceivability Argument: an

Inquiry into Chalmers's Two-Dimensional Semantics

Graduate Name: Chen Jingkun Major: Foreign Philosophy

Directed by Prof. Wei Yidong

ABSTRACT

The proposal of the assumption of zombies is closely related to the prevalence of physicalism. A zombie is a system exactly resembles us in all physical respects but has no mind or conscious experience or qualia. It is obvious that such a system cannot exist actually, but the problem is whether or not the zombies conceived by the philosophers can possess some kind of possibility. It is argued that if zombies are so much as a bare possibility, then physicalism is false and some kind of dualism must be accepted. For many philosophers that is the chief significance of the zombie idea. Opposing the physicalism with the use of zombie also leads to a more universal questions about relations between conceivability and possibility. And zombies also leads to epistemological difficulties: it reinstate the 'other minds' problem.

Chalmers renews the concept of zombie when he take it as a refined conceivability argument. The kind of physicalism which opposed by Chalmers entails the impossibility of zombies, while intuitively we feel that zombies are conceivable and thus possible. In the contemporary philosophy of mind, a number of popular arguments against physicalism, such as the knowledge argument, the explanatory-gap argument, and the modal argument, conclude that there is an ontological gap between physical process and consciousness as the epistemic gap between physical truths and phenomenal truths. However, physicalists hold that the epistemic premises do not entail ontological conclusion. Chalmers revises the traditional conceivability argument and argues that the inference from epistemic premises to ontological conclusion is reasonable. His process is to reason first from

epistemic premises to modal conclusions (about necessity and possibility), and from there to ontological conclusions. The point of the process is the link between the epistemic domain and modal domain. Chalmers assumes that two-dimensional semantics can rebuild the bridge between conceivability and possibility. This is a real challenge for physicalism if Chalmers is right.

Two-dimensional approach directly originates from Kripke's interpretation of the necessity a posteriori which causes the controversy between direct referentialism and descriptivism. Two-dimensionalism distinguishes two kinds of intension so that the terms such as a priori, a posteriori, possibility and necessity can be defined as the different evaluation in possible worlds. To support the conceivability argument, Chalmers develops a version of epistemic two-dimensionalism and defined the modal concept in epistemic terms. In recent years, new debates begin with the rise of two-dimensionalism and the use of new semantic methods and modal concepts. Soames detailedly attacks the descriptive tendency of two-dimensionalism. While Chalmers argues that his epistemic two-dimensionalism can maintain both referentialism and descriptivism.

The paper focuses on the analysis of Chalmers's conceivability argument against physicalism and concludes that the refined conceivability argument cannot confute physicalism virtually because of some deficiencies.

Key words: Zombie; Physicalism; Conceivability; Possibility; Two-Dimensional semantics

引 言

怪人^①可以典型地描述为在物理上与人类相同但却没有意识的某种系统。但这种描述显得有些笼统，首先，对于“物理的”这个词至少可以有三种不同的理解：行为的、功能的和微观物理的^②，这里通常取狭义的理解，即在细胞甚至更基本的微观物理层次上的。从行为怪人到功能怪人，再到物理怪人，其与意识动物的相似程度递增，物理怪人蕴含功能怪人和行为怪人，与意识动物在物理层面完全一致——甚至是分子对分子、原子对原子的严格吻合；其次，“没有意识”这种说法比较狭窄，事实上以下系统都可以视为怪人：缺失的现象性质（absent qualia）、颠倒的现象性质（invert qualia）、异质的现象性质（alien qualia）、变动的现象性质（dancing qualia）等。但我们满足于这样一种描述，虽然有些笼统，但最为典型。

“怪人”这一思想在哲学思考中反复出现，只是表现形式不尽相同，他心的认知问题就是一种表现形式。怪人观念至少可以追溯到笛卡尔的自动机（automata）。

按照笛卡尔的二元论，心灵和物质是两种完全不同的实体，前者是没有广延的纯粹思维，后者具有广延而不能思维。笛卡尔进而认为，人体和动物都是自动的机器，按照类似“刺激-反应”的模式行动，其行为完全可用物理语言来解释。“我首先曾把我看成是有脸、手、胳膊，以及由骨头和肉组合成的这么一架整套机器，就像从一具尸体上看到的那样，这架机器，我曾称之为身体。”^③而灵魂或心灵寓居于身体中，根据自然法则而作用，如同舵手控制船只^④。赖尔将笛卡尔的二元论观点称为“机器中的幽灵的教条”^⑤（将这种观点完全归于笛卡尔本人是不公正的），根据这一教条，真正的人既不能被定义为身体，也不能被定义为心身结合体，而只能是

^① 怪人（Zombie）一词源自安哥拉的班图语，原意为“鬼魂”或者“逝去的灵魂”，最早伴随着海地的伏都教（voodoo or vodou）一起出现在威廉·B·希尔布鲁克的小说《魔幻岛》中，在伏都教的传统观念中，巫师可以通过有魔力的符咒或者药剂使死人复活，或者使普通人丧失自由意志或灵魂，从而能够像使用奴隶一样随意支配他。根据希尔布鲁克的描述，这种人与常人一般但面无表情，目光呆滞，没有记忆没有意识，只是机械地服从控制者的命令。因此 zombie 通常译为“丧尸”、“僵尸人”、“行尸走肉”等等，令人顿生寒意，好莱坞甚至以类似题材演绎出了大量恐怖电影。

但无论是海地的 zombie 还是好莱坞的 zombie，都与哲学上的 zombie 有着根本区别。哲学上的 zombie 指的是一种在行为、功能或物理方面与人类完全一样，但却没有意识或感受特性的生物。上述通常译法过于毛骨悚然，并不适合哲学 zombie。虽然哲学上一般称之为“怪人”，但这一翻译并不确切，因为实际上哲学 zombie 表面上看来并不怪，他们的言谈举止与我们并无二致，换言之，即使放宽图灵测试中的限制条件，甚至完全去除这些条件，我们也无法得到我们想要的结果，——怪人并没有任何外在的可识别特征，他只是没有感觉经验的质。也有研究者根据发音直接译为“宗比人”或根据意义称为“失心人”，后者也不确切，因为“失心”有失魂落魄的意思，仍然可以识别出来，而“宗比人”则显得拗口，因此本文仍沿袭“怪人”这一译法。

^② 这三种同一方式大致对应于三种理论，即行为主义、功能主义和同一论，参见 Guzeldere, G, Varieties of zombiehood[J]. *Journal of Consciousness Studies*, 2(1995): 326-333.

^③ [法]笛卡尔. 第一哲学沉思集[M]. 庞景仁译. 北京: 商务印书馆. 1986. 第 24 页。

^④ 参见[法]笛卡尔. 第一哲学沉思集[M]. 庞景仁译. 北京: 商务印书馆. 1986. 第 85 页。笛卡尔后期进一步提出了更为精致的松果腺理论，尽管这个理论被科学的发展证伪。

^⑤ [英]赖尔. 心的概念[M]. 徐大建译. 北京: 商务印书馆. 1992. 第 10 页

单个的心灵。总之，人体不过是一架机器，理性的的心灵可以以一种神秘方式熟练操作它，而它也能够神秘地影响心灵的思维。并且，笛卡尔认为“即使里边没有精神，也并不妨碍它跟现在完全一样的方式来动作，这时它不是由意志指导，因而也不是由精神协助，而仅仅是由它的各个器官的安排来动作”^①笛卡尔的这些观点被认为是完全机械主义的。人虽然因心灵而与动物有着根本的不同，与人的机械行为相伴随的是他的意识和经验感受，但是否存在一种外表和行为看起来都和人类一模一样的机器？或者说人是否可以没有心灵而只以肉体的方式存在？笛卡尔认为可以从两个方面揭示自动机与人的区别：机器只能僵化地做出反应而不能创造性地运用语言，也不能在不同的条件下产生相应的非动词性的行为（non-verbal behavior）。因此在笛卡尔看来，没有什么机器能够像人那样行动，他的结论是：解释人的行为必须诉诸超物理的非物质的的心灵。

可见，笛卡尔的自动机还不是哲学怪人，因为这种自动机仍然具有外在的可识别性，怪人的可能性在笛卡尔那里也没有成为真正的哲学问题，因为在笛卡尔看来像真正的人那样举止言谈的自动机是不存在的。笛卡尔所提出的两个区分自动机和人的方法对于怪人而言是完全不适用的，毋宁说，怪人和正常人之间根本不存在这样一种可识别的问题，因此，笛卡尔并不仅仅是“只差拼出zombie这个词”^②。

笛卡尔的二元论解释很难解决心身如何发生作用这一问题。但笛卡尔所提出这种实体二元论贴近常识，并且可以通过各种方式加以改造。马勒伯朗士、莱布尼茨等人分别提出偶因论和平行论，试图通过否定心灵与身体之间的因果关系来克服笛卡尔所遇到的困难，但这些理论除了诉诸上帝并没有什么高明之处，赖尔的批判使人们逐渐抛弃了作为实体的心灵概念，现代科学技术则使人们越来越接受物理主义的一元论。

经典物理学认为物理世界是因果闭合的，每一个物理结果都有一个物理原因。当这一原则用于解释人的行为的时候便导致了神经生理学的诞生。但如果人类行为在物理上就可以获得完全充分的解释，那么意识将置于何地？一种最方便不过的解释是把意识也当成一种物理过程，比如，疼痛就是C纤维肿胀。然而意识现象毕竟很难完全用物理学术语来解释，一些思想家断定其中包含非物理的因素。并且物理主义坚信，如果两个事物在物理方面相同一，则它们在一切方面同一。如果在物理方面与我们严格同一的怪人是可能的，那么物理主义的这一信条将面临危机。一种妥

^① [法]笛卡尔. 第一哲学沉思集[M]. 庞景仁译. 北京: 商务印书馆. 1986. 第 88-89 页

^② 参见斯坦福百科全书“zombies”词条，该词条由 Kirk, R 撰写，在谈到笛卡尔的自动机的时候他认为笛卡尔就差拼出 zombie 这个词了，但笛卡尔的自动机和 zombie 存在着重要的区别。

协的观点是副现象论，这一观点承认世界的因果闭合性，也承认物理事件可以产生心理结果，但是心理事件是惰性的，不能产生任何物理上的结果，也不能干预物理世界的因果秩序，这样，意识现象就其因果性来说只是物理现象的“副现象”或“副产品”，因此允许怪人的可能性，斯多特（Stout, G. F.）等人就认为，如果副现象论是正确的，那么怪人就是可信的。^①

上世纪 70 年代很多哲学家提出类似的设想以反对物理主义，较早作为心理物理同一论的反例提出来的有坎贝尔（Campbell, K.）的“仿真人”（imitation man），其脑状态在物理化学属性上与我们的脑完全一样但却感觉不到疼痛，也看不到颜色。但最早大篇幅专门讨论怪人的应当是科克（Kirk, R.），1974 年他连续发表了《知觉与行为》（*Sentience and Behaviour*）以及《怪人 vs 物质主义》（*Zombies vs. Materialists*）两篇文章，明确提出了哲学上的怪人假设，认为怪人对于所有物理主义都是一个反例。科克设计两种怪人，一个他称为“丹”，他的现象性质正被剥夺，感觉属性正一个接一个地消失。尽管如此，只要它保留了完整的意识，大多数时候他仍然作出合理的行为。另一个是“格列佛”，一队小人侵入其大脑，切断了输入输出神经之间的联系，监视从输入神经进来的信息，并输送信息给输出神经以产生和原先一样的行为。这时候格列佛在行为特征方面与先前无异，只是没有感觉和其它经验。布洛克（Block, N.）和肖梅克（Shoemaker, S.）等人也设计了一些其他类型的怪人，如行为或者功能上与人类无异但没有感受特性，与科克的设想基本一致。不过科克的文章之后的很长一段时间里鲜有哲学文献详细地讨论哲学怪人，相关的争论主要是围绕着“现象性质”展开。

1994 年在第一届图克森会议上，查尔默斯提出区分意识的困难问题和容易问题，引起重大争论。意识的容易问题“可直接接受认知科学的标准方法的处理”^②，可以用计算或神经机制的术语予以解释，而对于困难问题，这些方法无济于事，因为它涉及的是经验或者现象性质。这一区分的意义在于反对用纯粹的物理科学方法对待意识问题，避免心灵哲学丧失其研究对象。1996 年查尔默斯出版著作《意识的心灵》，在该书的第三章，他提出了反对物理主义的五个论证，第一个就是怪人的可能性问题^③，引起哲学家们的强烈兴趣，成为了心灵哲学中的一个热点话题，围绕怪人的可

^① 斯多特并没有直接谈到怪人概念，它只是描述了副现象论者的世界的不可思议性：如果从不曾有经验的个体出现，那么世界的结构和过程将是另一回事，人类的身体只是在机械地制造和使用桥梁、电话和电报，装模作样地著述和阅读、在会议上长篇大论、争论物理主义……参见斯坦福百科全书“zombies”词条。

^② 查尔默斯. 勇敢地面对意识难题. 见高新民等编《心灵哲学》. 北京：商务印书馆. 第 361 页。

^③ Chalmers, D.J., *The Conscious Mind: In Search of a Fundamental Theory*[M], Oxford University Press, 1996, pp94-106.

设想性和可能性发表了大量论文。钱德勒 (Gendler, T.S.) 和霍桑 (Hawthorne, J) 2002 年编的论文集《可设想性与可能性》反映了关于这一问题的争论。查尔默斯借助二维语义学发展并完善了传统的可设想性论证, 由于二维语义学涉及可能世界问题, 并且在指称问题上具有描述主义倾向, 因此在语言哲学中也引起了激烈争论。加西亚卡宾特洛 (Garcia-Carpintero, M.) 和马西亚 (Macia, J) 2006 年编著的论文集《二维语义学》收录了近年这方面的重要论文。

一 传统可设想性论证及其困难

怪人是作为一种思想实验提出来的，以反对心灵哲学中的流行的物理主义，怪人问题在 70 年代并没有引起广泛关注，但 90 年代查尔默斯重新提出后迅速成为争论的热点，原因在于，科克等人在讨论怪人的可能性的时候诉诸人的直观，仅仅强调其逻辑可能性，而查尔默斯的贡献在于他将怪人假设作为一种精致的可设想性论证提出来，攻击物理主义尤其是后天物理主义。

1.1 怪人假设和物理主义

怪人在物理方面与我们完全一致但没有意识经验，他们像我们那样举手投足，这一的想法虽然令人不安但却能帮助我们弄清楚现象意识的问题，尤其是物理主义的问题。

目前物理主义流派众多，观点各有差异，比如同一论、还原论、取消论、副现象论等等，总的说来，除了少数极端的二元论者，当代哲学家几乎都相信世界是物理的，如果从这种极宽泛的意义上理解物理主义，那么关于物理主义和反物理主义的争论实际上大多都是物理主义内部的争论，具体来说，有非还原的物理主义和还原的物理主义之间的争论，也有物理主义和属性二元论的争论。因此很难给物理主义下一个定义，但是物理主义者基本上都接受这样的观点，即对这个世界的物理描述是对它的完整的描述，精神事实按照某种方式依赖于物理事实，并且从物理事实中必然地产生出来的。根据对这种必然性的不同理解，可以把物理主义分为先天物理主义和后天物理主义。^①

先天物理主义主张意识现象先天蕴含于物理事实，认为精神事件在逻辑上随附于物理事件，同一论和副现象论，甚至取消论均可归入这一类物理主义，查尔默斯称之为A型物理主义。先天物理主义认为意识研究中只存在容易问题没有所谓的困难问题，否认现象事实和物理事实之间存在认知上的鸿沟，否认怪人的可设想性。这种物理主义有时表现为分析的功能主义或逻辑行为主义，它并不否认意识现象的存在，但只能通过功能或行为术语来定义，也就是说，我们有意识只是因为我们具有某种功能性的能力，对于意识，除了功能，没有什么需要解释。^②

^① 在康德那里先天性和必然性是相等同的，先天性是必然性的来源、根据和保证。但克里普克论证了后天必然性和先天偶然性的存在。物理主义一般坚持物理事实必然蕴含现象事实，但这种必然蕴含是逻辑上的先天必然性还是后天经验的必然性则存在争议。

^② 这类物理主义者论证其观点的策略无非两点，一是类比，即在其他所有领域中解释功能便理解了全部，意识同样如此；二是间接的否认，认为拒绝物理主义将导致二元论。参见 Chalmers, D.J., *Consciousness and its Place in Nature*[C], in Stich, S.P. & Warfield, T.A. (eds), *The Blackwell Guide to Philosophy of Mind*, Blackwell Publishing Ltd. 2003, pp.118-111.

先天物理主义因为与我们的直觉相去甚远因而很难得到辩护，克里普克对后天必然性的论证似乎使他们看到了希望，物理主义者也试图用后天必然性来支持他们的观点，也就是说，心身关系在某种程度上类似于水和 H₂O 或者基因和 DNA、启明星和长庚星之间的关系，名称不同但是指称相同，这种同一并非来自概念分析，而是来自经验发现。因此意识概念虽然不同于物理的和功能的概念，但是我们将会发现它们指称的其实是同一个东西。因此，这种后天物理主义者承认物理事实和现象事实之间存在认知上的鸿沟，但否认二者之间存在本体上鸿沟；怪人虽然可设想，但形而上不可能，困难问题和容易问题并不意味着本体论的区分。

查尔默斯承认这一观点很具有吸引力，但困难在于，这种类比并不恰当，意识现象和相应的物理现象之间的认知鸿沟并不存在于其它领域，克里普克本人也反对将心身关系的同一性（如疼痛和C纤维肿胀）与通常类型的科学同一性（如热和分子运动）作类比^①，因为“在热和分子运动的事例中，有某种东西，即热的感觉，它是外部现象和观察者之间的一种媒介物。在心理-躯体的事例中，不可能存在这样的媒介物”^②，换言之，“热”这个词的指称是由其指称对象的一种偶然性质确定的，而疼痛直接就是疼痛本身，它并不由某种偶然性质标示出来，而是由它的直接的现象性质来标示。因此C纤维肿胀和被感觉为疼痛的C纤维肿胀显然是不同的，“在大脑状态与心理状态之间的对应性具有一种明显的偶然的因素”^③，而“同一性不是一种可以偶然地存在于对象之间的关系”。这就意味着如果一个人感觉不到疼痛那就可以说疼痛对他而言不存在，而我们完全可以设想某人体内发生了C纤维肿胀但是没有感觉到疼痛。查尔默斯认为，这就是为什么我们不能无矛盾的设想一个与我们的世界在物理上同一但却没有水或基因的世界，但我们却可以设想一个与我们的世界在物理上同一但却没有意识的世界。因此，除了物理事实所提供的，意识或者现象性质还需要额外的东西才能产生，而如果没有这种额外的东西，那就意味着意识需要非物理的属性，而这种属性在纯物理的世界中并不存在。^④但是要接受某种形式的二元论也面临很大的困难，毕竟显微镜和望远镜没有给意识留下任何藏身之所。正如克里

^① 克里普克区分了三种同一性论题：人与其躯体的同一；特定感觉与特定大脑状态的同一；心理状态的类型与相应物理状态的类型的同一。克里普克最为关注的是第三种同一性，即类型与类型之间的同一性，例如“疼痛”和“C纤维肿胀”的同一。克里普克认为“不能把这种情况解释为和通常类型的科学同一性相类似，例如和热与分子运动的同一性相类似”。参见《命名与必然性》第121页。

^② [美]克里普克. 命名与必然性[M]. 梅文译. 上海：上海译文出版社. 2001. 第128页。

^③ 同上，第131页。

^④ 为了让生物“把中枢神经的刺激感觉成疼痛，而不是呵痒、或者温暖、或者木然无知”^④，上帝除了单纯地创造中枢神经的刺激之外还需要要做些“额外的工作”吗？克里普克在讨论同一性的时候提出了这样的问题并且作出肯定的回答。但克里普克同时强调“对同一性论题的否定并不暗示我接受了笛卡尔的二元论”。参见《命名与必然性》第132页。

普克所承认的那样“心身问题上一个悬而未决而且及其令人困惑的问题”。^①

值得注意的一个问题是，怪人假设与副现象论的关系。一种观点认为，从副现象论可以很自然地得出怪人概念，因为副现象论认为意识并不起因果作用，而如果意识其因果作用我们就不会认为怪人是可设想的。但事实上怪人和副现象论之间并没有必然联系，很多后天物理主义者既承人怪人是可设想的，也承认意识其因果作用。另外，在笛卡尔的二元论中意识也其因果作用，这种观点也和怪人的可设想性设甚至可能性相容。

1.2 怪人假设和反物理主义论证

主要的反物理主义论证包括：解释空缺论证、知识论证、模态论证、可想象性论证等^②。这些论证彼此区别却又相互联系。解释空缺论证体现了反物理主义者的基本信念，即，意识的现象性质和大脑的物理过程之间存在解释空缺，其他几种反物理主义论证都可以看作是解释空缺论证的扩展和延伸。在这些论证中可设想性论证比较特殊，因为知识论证和模态论证都可以转化为某种形式的可设想性论证。

1.2.1 解释空缺论证和可设想性论证

“解释空缺论证”又译为“解释鸿沟论证”，这个论证可以追溯到莱布尼茨的“磨坊（或工厂）论证”（mill argument）。“我们必须承认，不能机械地解释知觉和依赖于它的东西，也就是说，不能用数字和运动来解释它。假设有一个机器的结构适合于产生思想、情感和知觉，而且我们设想它在体积被增大的同时仍然保持着原来的比例，这样一来，我们可以像进入一个工厂那样进入它里面。那么在进入它之后，我们将发现一些相互作用零件，却不能发现任何可以用来解释一个感觉的东西。”^③ 莱布尼茨要求我们设想的这架机器就是大脑，当科学家们研究大脑的工作机制的时候他们所发现的只是一些物理事件，这对于解释经验的性质毫无帮助。这一论证基于以下认识，即，无论物理学对人体和大脑的结构功能的了解多么透彻，也无法解释何以某事物在我们眼里看来是红色而非绿色、何以C纤维肿胀便产生了疼痛。这就表明主观感受特性并没有在物理学上得到解释，甚至被排除在解释之外，因此，在

^① [美]克里普克. 命名与必然性[M]. 梅文译. 上海译文出版社. 1998. 第132页。

^② 查尔默斯将属性二元论的观点也作为一种主要的论证形式，但属性二元论更是一种解决方案，而不是论证方法。查尔默斯人为属性二元论论证的要点在于，如果心理状态和物理状态是偶然同一的，或者说，如果心理状态和物理状态的同一并非先天的，那么它们必定是由不同的属性标示出来。其结构可以表示为：

(1) 对于所有物理术语p和现象术语q，‘p=q’并非先天的。

(2) 如果‘p=q’并非先天的，则p和q经由不同的属性标示出其指称。

(3) 结论：对于所有物理术语p，q用来标示出其指称的属性不同于p用来标示出其指称的属性。

显然，前提(2)和可设想性论证的CP论题有着密切联系，因为‘p=q’并非先天的意味着 $\sim(p=q)$ 是可设想的。

^③ 转引自：[美]约翰·海尔. 心灵哲学导论[M]. 高新民等译. 北京：中国人民大学出版社. 2005. 第80页。

意识感受特性和大脑物理过程之间存在着解释空缺，这种空缺在自然事物的解释中是不存在的。列文强调，或许物理主义在本体论上是正确的，但是在认知上却留下了解释空缺。这种将本体论上的必然性和认知上的先天性区分开来的做法显然仍是受了克里普克的影响。列文认为怪人的可设想性论证是“解释空缺的主要表现形式”，不过并没有什么有效的办法来填补这种解释空缺，即使承认怪人是可能的，这一问题依然存在。

怪人论证是最具代表性的可设想性论证，以至于可设想性论证成了怪人论证代名词。这一论证的思路大致如下：既然物理主义坚持 P 蕴含 Q 是必然的，这就意味着 $P \wedge \sim Q$ 形而上不可能，如果能证明 $P \wedge \sim Q$ 形而上可能，那么物理主义的观点显然就是错误的。这里 $P \wedge \sim Q$ 意味着只具有物理属性而没有感知特性，正是我们所谈论的是怪人，我们可以设想它们的存在，如果能够从怪人的可设想性推出其形而上可能性就能顺利驳倒物理主义，因而，这种可设想性论证又称为怪人论证，其基本形式如下：

- (1) $P \wedge \sim Q$ (怪人) 可设想；
- (2) 如果 $P \wedge \sim Q$ (怪人) 可设想，则 $P \wedge \sim Q$ (怪人) 形而上可能；
- (3) 结论：如果 $P \wedge \sim Q$ (怪人) 形而上可能，则物理主义错误。

1.2.2 知识论证

知识论证是杰克逊 (Frank Jackson) 提出来的著名的思想实验，熟知所有物理知识的天才科学家玛丽生来就被关在只有黑白二色的屋子里，她了解不同的颜色是由于不同波长组合的光线刺激视网膜所造成的，但是当玛丽走出黑白屋子或者得到一台彩电时会发生什么呢？她是否学到了她之前所不曾知道的东西呢？直觉告诉我们，玛丽显然获得了关于视觉经验的新知，因而现象事实有别于物理事实，物理主义的观点是错误的。这个论证引起哲学家们的极大兴趣和热烈讨论，查尔默斯认为这个简单的论证并不足以推翻物理主义，因为“物理事实”这一术语还比较含糊，它既可以只限于微观物理学、化学、生物学等基础的、低层次的领域，称为“窄物理事实”；也可以在宽泛的意义上包含较高层次的属性和事实，称为“宽物理事实”，宽物理事实本身并不属于但形而上地随附于窄物理事实，也就是说后者蕴含前者。这样，知识论证就可以稍作修正：

- (1) 玛丽知道所有的事实都可以由窄物理事实推导出来；
- (2) 玛丽并不知道所有的现象事实

(3) 如果现象事实并非从窄物理事实中推导出来,那么它显然并不蕴涵于窄物理事实。

(4) 并非所有的现象事实蕴涵于窄物理事实。

如果以 P 代替物理事实或物理真理,以 Q 代替现象事实或现象真理,将推导理解为先天蕴含,把前提(1)和前提(2)合并,这样得到的论证在结构上类似于可设想性论证:

(1) P 蕴含 Q 并非先天的

(2) 如果 P 蕴含 Q 并非先天的,则 P 蕴含 Q 并非必然的

(3) 如果 P 蕴含 Q 并非必然的,则物理主义错误

(4) 结论:物理主义错误

前提(1)相当于怪人可设想,前提(2)相当于怪人的可设想性蕴含其可能性。

这两个论证有着细微但重要的差别,P 蕴含 Q 并非先天这一前提比玛丽知道所有的事实都可以由窄物理事实推导出来这一前提更强,原因在于,后者之为真可以在以下情况仍得到保证,即,玛丽在黑白屋子里并不能获得涉及 Q 的概念,比如,作为现象概念的“红色”,这样的话玛丽当然就不能推出他所不知道的概念。

1.2.3 模态论证

克里普克反对同一论的论证就是他的模态论证,查尔默斯认为这一论证与可设想性论证最为接近,设‘p’表示一种特殊的疼痛状态,设‘c’表示一种相对应的大脑状态,或者是某个同一性论者所希望的与 p 相同一的大脑状态,如 C 纤维肿胀,这一论证的形式大致如下:

(1) ‘p=c’明显是偶然的;

(2) 如果‘p=c’明显是偶然的,那么在认知情境中就会存在这样一个世界,其中有一个人和我在质上是同一的,但与‘p=c’相应的陈述却是错误的。

(3) 如果在认知情境中就会存在这样一个世界,其中有一个人和我在质上是同一的,但与‘p=c’相应的陈述却是错误的,那么在这个世界中‘p=c’就是错的。

(4) 如果在某个世界中‘p=c’是错的,那么‘p=c’错误

(5) 结论:‘p=c’错误。

这个论证同样可以转化成可设想性论证:

(1) $\sim(p=c)$ 是可设想的

(2) 如果 $\sim(p=c)$ 是可设想的,则 $\sim(p=c)$ 是可能的

(3) 如果 $\sim(p=c)$ 是可能的, 那么‘ $p=c$ ’就是错误的

(4) 结论: ‘ $p=c$ ’错误。

区别在于克里普克的论证仅仅反对同一论, 而不是一般的反对物理主义, 但对这一论证仅进行适当扩展显然并不困难。

总的看来, 可设想性论证在反物理主义论证中具有特殊的重要性, 其它反物理主义论证都在某种程度上与之相联系, 因此, 澄清可设想性论证也有助于理解其他论证。并且可设想性论证在逻辑上属于否定后件式, 对物理主义的威胁更为直接。

1.3 传统可设想性论证及其困难

上文已经给出了一个可设想论证的简单形式:

- (1) 怪人是可设想的
- (2) 凡可想象的都是可能的
- (3) 所以, 怪人是可能的

显然这个论证在形式上并没有什么问题, 但是, 它的两个前提都是有问题的, 它们并没有清晰地表述出来, 即使得到澄清也是存在争议。关键问题是如何理解“可设想的”在这一语境中的含义。

许多哲学家乐于承认怪人在某种意义上是可设想的。但这个“某种意义”有时候显得过于宽泛。比如, 说“疼痛与C纤维肿胀这两个概念并没有本质上的、先天的联系”, 这一主张通过以下观点得到支持: “完整地掌握其中一个概念而对另一个一无所知, 这在原则上是可能的”^①。不过按照这种标准, 说圆周率是有理数也是可设想的了。因此, 前提1和2中所引用的可设想性须严格限定。更重要的是这一论证还依赖于一条形而上学的假设: 可设想的即可能的, 或者说可设想性蕴含可能性(以下将这一假设简称为CP论题)。CP论题长期以来几乎被当作公理直接拿来运用, 显然, 如果CP论题不成立, 那么可设想论证就立刻失去了支撑。

1.3.1 传统的可设想性论证

心灵哲学中的可设想性论证可追溯至笛卡尔, 他在论证其观点的时候常常使用这一技巧^②, 在第六沉思中, 他提出了一个完整的本体论二元论的可设想性论证:

^① Hill, C.S., Imaginability, conceivability, possibility and the Mind-Body Problem[J], *Philosophical Study*, 87(1997), p76.

^② 笛卡尔在“我思故我在”的论证中实际上就已经潜在地涉及了CP论题, 因为在笛卡尔看来, 设想一个没有思想者的思想过程是极其荒谬的, 换言之, 没有主体“我”的“思”是不可设想的。笛卡尔同样采用可设想性论证来证明上帝的存在, 这个证明真正来说应该可追溯到安瑟尔谟关于上帝存在的本体论论证。参见: 李麒麟. 休谟在其因果论证当中的‘可设想性原则’[J]. 外国哲学. 第十八辑(2005). 第186页。

“首先，因为我知道凡是我清楚、分明地领会的东西都能就像我所领会的那样是上帝产生的，所以只要我能清楚、分明地领会一个东西而不牵涉到别的东西，就足以确定这一个东西是跟那一个东西有分别或不同的，因为它们可以分开放置，至少由上帝的全能把它们分开放置；至于什么力量把它们分开，使我把它们断定为不同的东西，这倒没有关系。从而，就是因为我确实认识到我存在，同时除了我是一个在思维的东西之外，我又看不出有什么别的东西必然属于我的本性或属于我的本质，所以我确实有把握断言我的本质就在于我是一个在思维的东西，或者就在于我是一个实体，这个实体的全部本质或本性就是思维。而且，虽然也许（或者不如说的确，像我将要说的那样）我有一个肉体，我和它非常紧密地结合在一起；不过，因为一方面我对我自己有一个清楚、分明的观念，即我只是一个在思维的东西而没有广延，而另一方面，我对于肉体有一个分明的观念，即它只是一个有广延的东西而不能思维，所以肯定的是：这个我，也就是说我的灵魂，也就是说我之所以为我的那个东西，是完全、真正跟我的肉体有分别的，灵魂可以没有肉体而存在。”^①

笛卡尔这一论证可以概括为以下形式：

- 1、凡是我清楚、分明地领会的东西都能就像我所领会的那样是上帝产生的；
- 2、如果我能够清楚、分明的领会对象 A 的存在与对象 B 的存在是完全分离的，那么，全能的上帝能够使 A 和 B 的存在彼此分离（由 1 得出）；
- 3、如果上帝使 A 和 B 的存在彼此分离，则 A 和 B 事实上肯定是完全不同的东西；
- 4、我能够清楚、分明地知道我的思维（心灵）和身体的存在彼此分离；
- 5、上帝是我的心灵和身体的存在彼此分离（由 4，2 得出）

结论：我的身体和我的心灵事实上是相分离的（由 5，3 得出）

对于前提 4，笛卡尔认为心灵和身体的分离是可设想的，也就是说，我们可以设想没有肉体的灵魂（disembodiment），也可以设想没有心灵的肉体（automata），因此，二者之间的差别是“实在的”^②。笛卡儿给出了一个相对独立的论证：

- 1、我知道我的心灵存在，并且它本质上是一个思维的东西；
- 2、如果我的身体存在，我就知道它是一种广延的东西，因为归于身体而言广延是它的本质；
- 3、我能够设想某个无广延而能思维的存在，反之亦然。

^① [法]笛卡尔. 第一哲学沉思集[M]. 庞景仁译. 北京：商务印书馆. 1986. 第 82 页。

^② 从经院哲学派生而来的用语，即具有完全不同甚至相互排斥的本质的两个实体之间的差别。

结论：我能够清楚、分明地知道我的心灵的存在与我的身体的存在是相分离的。

显然，这个论证把CP论题作为隐含前提直接运用了，休谟在反对“原因的必然性”的时候同样把CP论题作为一条公设毫无保留的接受下来并加以运用^①，他明确指出“形而上学中有一条确立的公理，那就是：凡心灵能够清楚地设想的任何东西，都包含可能存在的观念，换句话说，凡我们所设想到的东西都不是绝对不可能的。我们能够形成一座黄金色的山的观念，由此就可断言，这样的一座山可能真正存在。我们不能形成一座没有山谷的山的观念，因此就认为这样的山是不可能的。”^②与休谟同时代的哲学家大多接受这一“公理”，克拉克（Clarke, S）、普莱斯（Price, R）以及（Wolff, C）等人甚至提出了一个更强的命题：反不可设想的都是不可能的^③。的确，当我们被问及某种事物是否可能时，我们总是试图去设想它，如果能够设想则认为可能的，反之则不可能。

但这一直觉显然并不像“两点之间直线最短”那样令人信服，因为可设想性是一个认知概念，而可能性属于模态概念，前提（2）存在着从认知领域到模态领域的跨越。可设想性的门槛越低，接受前提（1）就越容易，而接受前提（2）就越难，因此对可设想性论证的“常规诘难（standard objection）”就是反驳CP论题。对CP论题最具影响力的批评来自克里普克和普特南，他们在意义理论方面的工作直接切断了可设想性和可能性之间的联系。

尽管克里普克反对物理主义者将心身关系和“水是H₂O”进行类比，但他严格区分了认知领域的模态领域，切断了二者之间的联系。可设想性是一个认知概念，事物对于某主体是否可设想取决于他知道什么相信什么，或者是他使用了什么样的概念或表征方式。按照这种认知的理解，可设想性并不直接蕴含可能性，因为这时候可设想性取决于我们所能提供的概念和背景信息，一个命题即使在现实中是不可能的，我们也可以设想它，换言之，认知设想中可能的未必形而上可能。比如，我们说哥德巴赫猜想无论正确还是错误都是可设想的，但这两种情况只有一种是可能的，而“水不是H₂O”虽然可设想，但是不可能。普特南也指出，虽然就某种特殊的意义而言，“水不是H₂O”是可设想的，但逻辑上是不可能的，可设想性并不是可能性的

^① 原因的必然性可表述为“任何开始存在的事物都必须有其存在的原因”，休谟认为这个命题既不具有直观的确定性，也没有论证的确定性。在说明后者时休谟指出：“……很容易设想任何一个对象在此时刻是不存在的，而下一个（时刻）就是存在的，而不用将该对象与其相区别的原因的观念或者生成原则的观念相连接。因此，对想象而言，这种在‘原因’的观念和‘开始存在’的观念之间的分离性是显然可能的；因而这些对象之间的现实的分离也就是可能的，其中并不包含任何矛盾和荒谬……” [英]休谟. 人性论[M]. 关文运译. 北京：商务印书馆. 1980. 第45页。参见：李麒麟. 休谟在其因果原则论证当中的“可设想性原则”. 外国哲学. p.181-183.

^② [英]休谟. 人性论[M]. 关文运译. 北京：商务印书馆. 1980. 第45页。

^③ Woudenberg, R., *Conceivability and Modal Knowledge*[J], *Metaphilosophy*, 2006, Vol.37, No.2, p.210.

依据^①。

1.3.2 可设想性与可能性

我们的感知能力向我们揭示了什么是现实的。为什么会是这样呢？一个被广泛接受的解释是：我们的感知机制对现实世界很敏感，后者作用于前者，产生刺激和反应的系统模式。同样，我们的感知能力向我们揭示了什么是可能的，但是对于为什么以及在何种程度上是这样却没有一个普遍接受的解释。而可设想性论证的关键就在于CP论题，也就是“可设想性—可能性”（或者“不可设想性—不可能性”）推理，这一论题包含四个方面的问题：第一、存在何种可能性，使得一种情形的可设想性可以被认为是那种情形的可能性的一个指示。第二、设想某事物是什么意思？第三、什么情况下可设想性能够可靠地导出可能性，第四、如何在特定实例中运用“可设想性—可能性”推理，以获得关于现实世界的主张。^②前面两个问题要求对可设想性和可能性两个概念进行澄清，在此基础上才能确定CP论题成立的条件，最后则是如何具体运用CP论题。

可能性这一概念在笛卡尔那里意味着无矛盾性，正如他在第二组答辩里说的：“一切不可能性，或者，如果我可以在这里用经院哲学的话来说，一切矛盾性仅仅在于我们的概念或思想里，因为它不能把互相矛盾着的观念结合到一起，而并不在于在理智之外的任何东西里，因为，就是由于它在理智之外，所以显然它是没有矛盾的，而是有可能的。”^③这就是说，当我们错误地把互不相容的观念混在一起的时候，所有这些自相矛盾的或者不可能的东西都只能存在于我们的思维中，而不可能发生在我们的理智之外的任何地方；因为正是有事物存在于我们的理智之外这一事实清楚地表明了它并非自我矛盾的，而是可能的。这种逻辑上的可能性意义最为宽泛，无法与可设想性建立内在的联系。这就需要对可能性概念进行进一步的分析，现在通常将可能性区分为认知可能性和非认知可能性。

认知可能性的定义与主体或主体的集合有关，具体而言就是提供给所涉及的主体的知识或证据，比如，某人 S 可以对认知可能性提供一种宽容的（permissive）解释，根据这种解释，P 对于主体 S 是认知上可能的，仅当 S 并不知道非 P，或者根据一种更严格的解释，P 对于主体 S 是认知上可能的，仅当 P 与 S 所知道的东西是形而上相容的（metaphysically compossible）。据此，我们还能够给出一种中间解释，P 对

^① Putnam, H., Is Water Necessarily H₂O?[A], in Conant, J.(ed), *Realism With A Human Face*[C], Harvard University Press, 1990, pp.55-57.

^② Gender, T., and Hawthorne, J., Introduction: Conceivability and Possibility[A], in Gender, T., and Hawthorne, J. ed. *Conceivability and Possibility*[C], Oxford University Press, 2002, p.2.

^③ [法]笛卡尔. 第一哲学沉思集[M]. 庞景仁译. 北京: 商务印书馆. 1986. 第 153 页。

于主体 S 是认知上可能的，仅当 S 的证据并不保证相信非 P，或者说根据 S 所知道的我们不能合理地期望他能确定非 P。甚至还可以在一种更为宽泛的意义上理解认知可能性，即：P 在认知上是可能的仅当非 P 不是先天的。

这些描述之间存在着重要的区别：按照严格的解释，认知可能性蕴含形而上可能性，按照宽容的解释，这种蕴含则不能成立。但是，不管这些描述的细节，无论在何种意义上，可设想性显然并不一般地导出认知可能性。如果我知道那只猫在垫子上，那么对于我来说那只猫不在垫子上在认知上是不可能的（即使在宽容的意义上）。而且，我能够很容易地设想猫不在垫子上的条件，因此我能够很容易地设想认知上不可能的事物。

非认知可能性一般认为有三种：逻辑可能性、形而上可能性、规则可能性（如物理或生物学可能性）。按照通常的描述，P 是逻辑上可能的仅当不能根据标准的演绎推理规则从 P 推出矛盾；P 对于相关的规则而言是规则上可能的，仅当 P 与那些规则所表达的真理是一致的；形而上可能性被认为是最重要的，可以用可能世界的术语来描述现实性和可能性：P 是现实的仅当 P 处于现实世界，P 是（形而上）可能的仅当 P 处于某可能世界。按照这样一种描述，形而上可能性的范围比规则可能性宽，比逻辑可能性窄：无论哪种意义上都不能说某物既是红色的又不是红色的，某物比光还跑得快形而上可能但物理上不可能。

“设想（conceive）”大约是这样一种活动，即通过语词、观念来表征包含现实或非现实特征的情形，根据布里坦（Brittain, C.）的考证，conceive 这个词和 concept 事实上是同源词，前者可追溯至拉丁语中的动词 concipere，后者则是其过去分词 conceptus，古代常用动词 concipere，conceptus 的名词性运用直到三或者四世纪的时候才出现，在此之前一直使用的是 notio（即英语中的 notion）。这样看来，对“conceive”一词的所做的较宽泛的使用是合理的，比如想象、预想、构想、构思、思议、描绘、映像等等。但休谟对可设想性的定义是纯粹逻辑上的：P 是可设想的，当且仅当 P 本身并不包含任何逻辑上的矛盾。这一定义显然过于宽泛，而且这一定义直接和逻辑可能性等同起来，于是可设想性和可能性之间就形成了循环定义：可设想的就是逻辑上无矛盾（可能）的，可能的就是可设想的，这也是笛卡尔和休谟将 CP 论题直接作为公理来运用的原因。

但是，当我们说某事物或者某情形可设想或者不可设想的时候究竟是什么意思呢？尽管存在各种不同的解释，但是有一点是确定的，即，从根本上来说这是一个认知概念，这就是说，一种事物或情形对于某主体是否可设想取决于他知道什么或者相信什么。具体而言，P 对于主体 S 是可设想的，当且仅当 S 认为 P 可能为真；或者 S

不认为P必然假；或者S不知道P是假的；或者P（作为一种事态而非命题或句子）被认为是可能的。^①按照这样一种认知的解读，可设想性并不直接蕴含可能性，因为这时候可设想性取决于我们所能提供的概念和背景信息，明显的例子是我们可以设想现实中并不可能的命题。因此，单纯的概念分析并不能在可设想性和可能性之间建立牢固的联系。

^① Yablo, S., Is conceivability a guide to possibility?[J], *Philosophical and Phenomenological Research*, 53(1993),p26.

二 基于二维语义学的可设想性论证

查尔默斯发展了一种认知的语义学，并用来恢复先天性和必然性之间以及认知领域和模态领域之间的联系，进而修正了传统形而上学中的 CP 论题，提出了一个精致的可设想性论证。

2.1 二维主义的发端

宽泛地说，二维主义就是将表达式与两种不同的语义值联系在一起，在这种意义上弗雷格关于含义和指称的区分也可以看成是一种二维主义^①。弗雷格以来的描述主义传统认为，名称的意义不仅在于其指称，而且包含由描述语提供的认知信息，正是后者决定了名称的指称。而认知意义和先天性是紧密联系在一起的，如果一个命题在认知上是琐碎的，它就是先天的，因而也是必然的。

克里普克针锋相对地对描述主义传统提出了批判，首先，他认为名称是严格的固定指示词，而描述语只是貌似固定指示，一旦事实并非所预想的那样，其指称的对象可能会发生改变；其次，不知道相关的认知信息并不妨碍我们对名称的正确使用；第三，我们所拥有的认知信息很可能是错误的，不能作为识别指称的唯一标准。这三个论证通常分别被称为模态论证、认知论证和语义论证，再加上克里普克对后天必然命题的说明，所有这些使描述主义遭遇到严重的困难。但克里普克的直接指称或因果指称理论似乎并未解决弗雷格的有新知的同一性陈述的难题，而描述主义也针对三大论证提出了各种反驳^②，至于后天必然性，这个概念看起来似乎是自相矛盾的：假定 S 为后天必然命题，既然是后天的，那么其 $\sim S$ 就是可设想的，如果没有经验证据来排除 $\sim S$ ，那就说明 S 是假的；而如果 S 是必然的，S 就不可能为假，也就是说，如果 S 是必然的，我们就不需要经验证据以知道 S，因此，如果 S 是后天的它就不可能是必然的。

关于后天必然性讨论一直以来颇受关注，毕竟克里普克的说明线条粗犷，相比之下他举的那些经典例子显得更有说服力。索姆斯认为二维主义产生的一个重要原因就是为了解释后天必然性。粗略来说，其主要思路是把表达式与两个相关命题联系在一起，后天必然的句子除了表达必然命题之外还与一个偶然命题相联系。这一思路其实并不陌生，因为克里普克对后天必然性的说明实际上已经包含了二维方法的萌芽：

^① Chalmers, D.J., Two-Dimensional Semantics[A], in E. Lepore & B. Smith (eds), *The Oxford Handbook of Philosophy of language*[C], Oxford University Press, 2006, p574.

^② Soames, S., Reference and Description: The Case Against Two-Dimensionalism[M]. Princeton University Press. 2005, pp35-37.

克里普克承认“任何必然真理，无论是先验的还是后验的——都不可能作相反的证明”，^①但这并不意味着后天性和必然性这二者是矛盾的。“就某些必然的后验真理而言，我们可以说，在具有适当的定性同一性证据的情况下，一个适当的、相应的定性陈述可能是伪的。”^②这句话是什么意思呢？克里普克解释道：“设‘ R_1 ’和‘ R_2 ’为等号两端的两个严格指示词。那么‘ $R_1=R_2$ ’这个式子如果是真的，它也就是必然的。‘ R_1 ’和‘ R_2 ’的指称可能分别由非严格的指示词‘ D_1 ’和‘ D_2 ’所确定。……虽然‘ $R_1=R_2$ ’是必然的，但是‘ $D_1=D_2$ ’这个式子却很可能是偶然的，我们之所以常会错误地认为‘ $R_1=R_2$ ’这个式子可能被证明为伪的，其根源即在于此。”^③

这里，克里普克已经提供了一种二维解释的雏形，其核心观念就是“一个适当的、相应的定性的陈述”（即‘ $D_1=D_2$ ’），与原先的必然陈述“ $R_1=R_2$ ”不同，它可以是伪的，因为“如果这个世界可能被证明是另外的样子，那么它也许本来就是那种样子。否认这个事实就是否认自明的模态原理：由某种可能性所蕴含的东西本身必然地是可能的”^④。也就是说，尽管‘ $D_1 \neq D_2$ ’不具有形而上的可能性，但在我们的认知设想中是可能的，我们不能先天的排除这种可能性，因而“ $R_1=R_2$ ”这一必然陈述是后天的，这就澄清了后天必然性这一概念的困惑。这种解释模式同样可用于对先天偶然性的说明^⑤。显然，克里普克在这里区分了两种可能性：认知可能性和形而上可能性，尽管后天必然命题包含无法先天排除的偶然命题，但这种偶然命题仅仅是认知上可能的，并不具有形而上的可能性。

2.2 早期二维主义

早期的二维主义代表人物有卡普兰（Kaplan, D.）、埃文斯（Evans, G.）等人，具有明显的语境论倾向。卡普兰主要用二维方法对直指词（indexicals）的特性和内容所作的分析，以“我现在饿了”这个句子为例，按卡普兰的分析，如果这个句子是张三在 12 点的时候说的，那么当且仅当张三在 12 点时饿了该句子所表达的命题为真，命题表达了言语行为的内容；换一个语境，如果是李四在 18 点说了同样一句话，则表达的内容完全不一样。因此，表达的特性（character）就在于它是从语境到内容的一种函项，即在言语行为的语境中标示出该语境中的那个表达的内容。如果把这里的内容理解为内涵，那么表达的特性显然也是一种二维内涵。因此，我们可以说，

^① [美]克里普克. 命名与必然性[M]. 梅文译. 上海译文出版社. 1998. 第 119—120 页。

^② 同上，第 120 页。

^③ 同上，第 120—121 页。

^④ 同上，第 118 页。

^⑤ Garcia-Caipintero, M., Two-dimensionalism: A Neo-Fregean Interpretation.[C]. in Garcia-Caipintero, M. & Macia, J. (eds), Two-Dimensional Semantics: Foundations and Applications, Oxford University Press, 2006, p183.

“我现在饿了”这个句子的特性在第一语境中标示出的命题是“张三 12 点饿了”，在第二语境中标示出的命题是“李四 18 点饿了”，同样的方式，我们可以说，该句子中索引词“我”和“现在”在第一语境中分别标示出张三和 12 点，在第二语境中分别标示出李四和 18 点。可见，“我现在饿了”这个句子只是在某些世界其命题内容为真，但其特性所产生的命题在所有说出它的语境中都为真。

卡普兰认为，索引词和指示代词的特性反映了它们的认知意义。因此，如果“这个就是那个”在认知上有意义，则“这个”和“那个”虽然指称同一对象，但它们的特性不同。

不过，卡普兰的二维方法的运用基本上只限于直指词和指示代词，他认为名称（names）无论在何种语境中都指称同一对象，这样的话，共指称的名称将具有相同的特性，这显然和他们的认知角色不一致，和他之前对直指词和指示代词的分析也不一致。另外，对于直指词，其语境和认知意义本来就是紧密联系在一起，而对于名称和自然种类词等许多其它表达，这二者之间并没什么联系，因此，卡普兰的方案更适合于分析表达的语境，而不是表达的认知意义。

埃文斯的分析关注的是摹状名称（descriptive names），即由摹状词确定指称的名称，他举的例子是“朱丽叶斯”，并规定用它来称呼发明“零”的那个人——如果有人独立的发明了它的话，因此，就有下面的句子：

a) 如果有人发明了“零”，则朱丽叶斯发明了零。

按克里普克，这个句子表达的是一个偶然命题，“朱丽叶斯”在所有世界标示出实际的发明者，但由于在这些世界中是其他人（威廉·G·华兹华斯）而非朱丽叶斯发明了零，因此该命题在所有世界为假。然而在埃文斯看来，这只是表面上的偶然性，可以定义为：S 表面上是必然的，当且仅当“必然 S”为真。然而在更深层的意义上，句子 a) 是必然的，他认为，一个句子是必然的，那是因为它表达了必然的内容，他还认为存在一条连接“朱丽叶斯”和零的发明者这二者之间的一条先天的语义规则，正是这条语义规则使得句子 a) 的内容必然为真，这种必然性埃文斯称之为深层必然性（deep necessity）。进而，埃文斯将句子的内涵区分为表层内涵（superficial intension，即句子所表达的命题）和深层内涵（deep intension，即句子的内容）。具体来说，对于任意句子 S，它的表层内涵是 S 所表达的命题为真的世界的集，它的深层内涵是 S 的内容为真的世界的集。这样，对于句子 a)，其表层内涵在某些世界为假，但深层内涵在所有世界为真。

可见，对于“朱丽叶斯”这样的摹状名称，深层必然性是与先天性联系在一起，

深层内涵则与表达的认知角色联系在一起，把这种联系扩展到其他表达上将会是一件引人入胜的工作，但埃文斯没有讨论其他表达，而且如果进行这种扩展也还不清楚。

另一方面，埃文斯认为对于一般专名，存在一条联系名称及其指称的语义规则，因而，指称也是内容的一部分，如此，则包含一般专名的同一（如前面的“马克·吐温是乔治·布什”）具有必然的内容，不仅是表面必然，而且具有深度必然性，因此，它们的深层内涵也就不存在什么区别。可见，在专名中，深层必然性和深层内涵并不像在摹状名称中那样紧密地和先天性或认知意义联系在一起。

可以看到，早期二维主义和语境有着密切联系，一些表达式的指称会随着语境的改变而变化，因此，表达式的指称似乎以某种方式依赖于外在世界的存在方式。这一点对于索引词尤其明显，但是否所有表达式都具有这样的性质？二维主义的支持者倾向于沿着这种思路进行扩展。

2.3 二维语义的核心观点

上文指出，语义学的二维方法基于以下观点，即，表达式的外延甚至内涵都是以某种方式依赖于外在世界的性质。以“水”为例，如果水指称 H_2O 的世界成为现实的，则水在所有世界指称 H_2O ，但如果其它世界（如，普特南的孪生地球）成为现实的，则“水”所指示的可能是完全不同的东西而不是 H_2O 。

如果一个表达式的内涵取决于世界的性质，那么就可以用从世界到内涵的函项来表示这种关系，而内涵本身又是从世界到外延的函项，于是这就揭示了一种二维结构。

以“水”为例，除我们所处的现实世界 W_1 ，我们设想仍有两个可能世界：

W_2 孪生地球，透明可饮用的水状液体其实是 XYZ；

W_3 透明可饮用的水状液体由 95%的 H_2O 和 5%的 XYZ 混合而成（仍称水为 H_2O ）；

如果以横轴表示现实世界的呈现方式，纵轴表示世界可能的反事实的呈现方式（每列表示一个可能世界），那么我们就可以用下图来表示“水”的外延对世界的依赖关系：

第二维：虚拟世界 →

第一维：现实世界 ↓

	W ₁	W ₂	W ₃
@W ₁	H ₂ O	H ₂ O	H ₂ O
@W ₂	XYZ	XYZ	XYZ
@W ₃	H ₂ O	H ₂ O	H ₂ O

这里@表示“现实的”或者“被视为现实的”，这样从@W₁到@W₃构成了这个矩阵的第一个维度。不同的二维主义对第一维度的具体规定有所不同，有时将它理解为表达的语境（比如卡普兰），有时理解为认知可能性（比如查尔默斯）。这里每一横行表示，当该行的世界成为现实的，或者被视为现实的时候“水”的内涵，称为第二内涵，它是从可能世界到外延的函项，可以表示为 $F(W \rightarrow E)$ 。

最上面一行从 W₁ 到 W₃ 表示的是世界可能的方式或者说反事实的可能世界，它们构成了这个矩阵的第二个维度，卡普兰将其理解为命题的赋值条件，查尔默斯将其理解为形而上的可能性。第一内涵是从视为现实的可能世界到外延的函项，可以表示为 $F(@W \rightarrow E)$ ，这样当 W₁ 或 W₃ 成为现实的，“水”的第一内涵给出 H₂O，当 W₂ 成为现实的，“水”的第一内涵给出 XYZ。这就是说，无论现实世界以何种方式呈现出来，“水”的第一内涵都只给出任何具有水这种物质的所有属性的物质。因此，第一内涵是先天的，相反，第二内涵并不是先天决定的，因为它本身取决于现实世界的具体特征。

整个这个二维结构直观地表示了一种二维内涵，即，世界的有序偶对 $(W \times W, W$ 表示可能世界) 到外延的函项，可以表示为 $F(@W \rightarrow (W \rightarrow E))$ 。对于句子“水是 H₂O”同样能够在二维结构中给出其真值：

水是 H ₂ O	W ₁	W ₂	W ₃
@W ₁	T	T	T
@W ₂	F	F	F
@W ₃	T	T	T

这样，“水是 H₂O”的后天必然性就可以解释为命题的第二内涵在所有虚拟世界为真而第一内涵在某些视为现实的可能世界为假。相应地，先天偶然性可以解释为

命题的第一内涵在所有视为现实的可能世界为真而第二内涵在虚拟世界为假。也就是说，先天性和后天性、可能性和必然性都可以定义为两种内涵在可能世界的赋值。这就很好地解释了具有相同外延或指称的表达式为何具有不同的二维内涵和第一内涵，同时也能在理论上很好地解释克里普克的后天必然性和先天偶然性。

2.4 二维语义方法与可设想性论证

如果把卡普兰的语义内容理解为表达式的内涵，那么他所说的“特性”显然就是一种二维内涵。不过，卡普兰对二维方法的运用基本上只限于直指词和索引词，他认为名称无论在何种语境中都指称同一对象，这一点和克里普克一致，况且卡普兰的原意就是捍卫直接指称，所以索姆斯将卡普兰的观点称为温和的二维主义，甚至承认在这种温和的意义上我们都是二维主义者。^①但是可以预见，如果将卡普兰的观点稍加扩展，便直接危及直接指称主义。查尔默斯 1996 年在其颇有影响的著作《意识的心灵》中就采纳了二维方法并用于反对心灵哲学中的物理主义，但查尔莫斯当时观点明显带有很强的卡普兰的烙印，之后十年里他致力于摆脱卡普兰语境因素的影响，发展一种认知的二维语义学，并试图用认识术语来定义模态概念，从而为可设想性论证提供支持。

2.4.1 认知的二维语义

二维主义的关键在于如何定义与先天性相联系的第一内涵。前面已经提到，对二维主义有两种不同的理解，一种是卡普兰式的语境的理解，一种是查尔默斯所主张的认知的理解。按照语境的理解，第一内涵是从语境到外延的函项，按照认知的理解，第一内涵是从认知可能的“情形”（scenario）到外延的函项。查尔默斯认为语境的理解具有局限性，因为不同类型的表达式对语境的依赖程度不同，对于索引词和指示词，它们本身就和表达语境密不可分，对于其它表达式，如名称和自然种类词，这种分析就不大合适。而且，查尔默斯的目的是要恢复认知意义和先天性之间的联系，语境的理解并不胜任^②。从认知的角度来看，对于现实世界可能发生的事实，我们在认知设想中可以有这样或那样的假设，查尔默斯称之为“情形”，我们不能先天地排除它们，而如果这些情形实现出来，那么相关的表达式可能具有完全不同的真值或意义。

基于这样一种观念，查尔默斯用认知术语对先天性、后天性以及必然性、偶然

^① Soames, S., Reference and Description: The Case Against Two-Dimensionalism[M], Princeton University Press, 2005, p44.

^② Chalmers, D.J., The Foundations of Two-Dimensional Semantics[A], In Garcia-Carpintero, M. & Macia, J.(eds), *Two-Dimensional Semantics: Foundations and Applications*[C], Oxford University Press, 2006, pp65-75.

性等概念进行了定义^①：

- (1) 句子 S 是形而上必然的，当且仅当 S 的第二内涵在所有世界为真；
- (2) 句子 S 是后天必然的，当且仅当 S 的第二内涵在所有世界为真而第一内涵在有些情形为假。
- (3) 句子 S 是先天偶然的，当且仅当 S 的第一内涵在所有情形真而第二内涵在有些世界为假。
- (4) 句子 S 是先天的（认知上必然的），当且仅当 S 的第一内涵在所有情形为真。
- (5) “ $A \equiv B$ ”是形而上必然的，当且仅当 A 和 B 的第二内涵相同。
- (6) “ $A \equiv B$ ”是先天的（认知上必然的），当且仅当 A 和 B 的第一内涵相同。

在（5）和（6）中，A和B是任意同类型的表达，“ $A \equiv B$ ”是真句子当且仅当A和B有共指称，如果A和B是个体词，则 $A \equiv B$ 相当于“ $A = B$ ”，如果A和B是句子，则 $A \equiv B$ 相当于“A当且仅当B”。（6）很容易让人想起弗雷格的观点：“ $A = B$ ”不具有认知意义当且仅当A和B含义相同^②。显然，和弗雷格一样，查尔默斯也主张，在意义和认知领域之间存在某种联系。

（1）和（4）定义了两种不同的必然性，而（1）描述的实际上就是克里普克的观点，即，如果 S 在所有可能世界为真，则 S 是形而上必然的。这是对形而上必然性的传统理解。（4）则是（1）在认知领域的反映，它是查尔默斯认知二维语义的核心观点，因为在其它的二维框架中，先天性和第一内涵（即认知内涵，在不同的二维框架中具体名称不同，如斯道纳克称之为对角内涵，杰克逊称之为 A 内涵）之间的联系是受约束的，而查尔默斯使这种联系具有普遍性，在查尔默斯看来，二维语义的核心立场就在于，涉及专名和自然种类词的后天同一，其第一内涵必定在某些情形中为假。

根据这两种必然性，我们可以得到两种可能性：如果 S 的第二内涵在世界 W 为真，或者说世界 W 被视为反事实的，从而令 S 为真，则我们说 W 满足 S，即 S 形而上可能，查尔默斯把这种可能性称为第二可能性；如果 S 的第一内涵在情形 W 为真，则我们说 W 确证 S，即 S 在认知上可能，这种可能性称为第一可能性。

如果查尔默斯的定义方式是可行的，那就在认知领域和模态领域之间搭建了一

^① Chalmers, D.J., Two-Dimensional Semantics[A], in E. Lepore & B. Smith (eds), *The Oxford Handbook of Philosophy of language*[C], Oxford University Press, 2006, pp586-588.

^② Chalmers, D.J., The Foundations of Two-Dimensional Semantics[A], In Garcia-Carpintero, M. & Macia, J. (eds), *Two-Dimensional Semantics: Foundations and Applications*[C], Oxford University Press, 2006, p.64.

座桥梁，我们可以说 S 是可想象的，如果它在认知上是可能的，反之亦然，因为 S 并没有被先天排除，虽然还不能直接得出可设想的就是可能的，但查尔默斯认为在可设想性和可能性之间仍然存在某种关联，通过对可设想性概念和可能性的分析，我们能够从有关可设想性的前提得出关于形而上可能世界的结论。

2.4.2 二维语义方法中的 CP 论题

查尔默斯对 CP 论题的反例进行了分析，他的这种分析实际上可以分为两类，因为这些反例要么出于概念上的混淆，要么出于模态上的混淆。对于前者，很显然，并非所有可设想的都是可能的并不意味着所有可设想的都是不可能的，也就是说，可设想性和可能性并非完全没有联系，这需要我们对可设想性这一概念进行辨析。首先，查尔默斯区分了当下的可设想性（*prima facie conceivability*）和观念的可设想性（*ideal conceivability*），前者受到当下主体认知能力的限制，而后者依赖于观念上的理性反思。举例来说，“ $2+2=5$ ”显然在这两种意义上都是不可设想的，设想一只会飞的猪却不存在什么困难，而复杂的数学真理的错误概念上不可设想，当下可设想（如果主体对相关知识一无所知）。其次，查尔默斯又区分了否定的可设想性（*negative conceivability*）和肯定的可设想性（*positive conceivability*）。当 S 不能通过先天推理来排除，我们就说 S 在否定意义上是可设想的，如果主体能够毫无矛盾地设想 S 如此这般的条件，我们就说 S 在肯定意义上是可设想的。

查尔默斯认为，相比当下的可设想性，观念的可设想性更适合用来与可能性建立联系，因为前者显得过于宽泛。比如，当我们说哥德巴赫猜想和这一猜想的反面都是可设想的，这其中的可设想性就是指当下否定的可设想性，不可能两种情形同时都是概念上可设想的，也没有理由认为观念的可设想性蕴含可能性存在反例。

对于模态上的混淆，查尔默斯认为可以用二维方法来澄清。以“水不是 H_2O ”为例，根据二维语义，当我们设想水不是 H_2O 时，我们设想了这样一个条件，即 XYZ 而不是 H_2O 充满海洋湖泊的孪生地球，没有理由怀疑这样一个 XYZ 世界（孪生地球）是形而上可能的，而如果 XYZ 世界成为现实的，我们就必须承认水是 XYZ。因此，尽管 XYZ 世界并未满足“水不是 H_2O ”（因为 XYZ 世界不是现实的），但 XYZ 世界确证了“水不是 H_2O ”（因为 XYZ 世界是形而上可能的），也就是说，“水不是 H_2O ”虽然不具有第二可能性或者说形而上可能性（因为其第二内涵在所有世界为假），但在某些情形中是可能的，因而具有第一可能性。

其次，查尔默斯认为，就“可设想性”的某种意义来说，水不是 H_2O 是不可设想的一一如果水在现实世界毫无疑问地就是 H_2O ，在这种意义上，任何水不是 H_2O

的可设想性条件（如孪生地球）更应被描述为水仍是 H_2O 的可设想性条件，只是其中的水状物质不是 H_2O 。因此“水不是 H_2O ”看似可设想，其实并非真的可设想，查尔默斯把这种可设想性称为第二可设想性。第二可设想性往往是后天的，因为水是 H_2O 这一点我们只是通过后天经验才知道的。与此相对应的是涉及先天领域的第一可设想性。S 是否具有第一可设想性取决于先天推理，否定的可设想性也是一种第一可设想性，因为它的定义方式是看什么东西能够被先天地排除，而“水是 H_2O ”并不是先天建立起来的，在这种意义上我们可以正确地说“水不是 H_2O ”是可设想的。

因此，当我们说“水不是 H_2O ”这类句子可设想但不可能的时候，我们是在第一可设想性的意义上使用“可设想”，而在第二可能性的意义上使用“可能”，第一可设想性当然不蕴含第二可能性。但我们可以说“水不是 H_2O ”具有第一设想性和第一可能性，这是否意味着第一可设想性蕴含第一可能性？查尔默斯并没有直接给出这种蕴含关系成立的依据，而只是指出我们找不到这种蕴含关系的反例^①，因而第一可设想性蕴含第一可能性是无法拒斥的。尽管这样的论证并不充分，但已经足够在认知领域和模态领域建立起一座桥梁。结合上面对可设想性概念的分析，传统的CP论题就有了两个精确的阐释^②：

（CP+）观念上肯定的第一可设想性蕴含第一可能性；

（CP-）观念上否定的第一可设想性蕴含第一可能性。

由于否定的可设想性蕴含肯定的可设想性，因此 CP-蕴含 CP+，但反过来是否成立似乎就很难确定了，因此 CP-或许比 CP 更强。但这种区别对可设想论证并没有太大影响，因为在查尔默斯看来，怪人具有概念上的第一可设想性（无论是否定的还是肯定的）是显而易见的，按照修正的 CP 论题，怪人也就具备第一可能性。不过这时候怪人仍然不具有形而上的第二可能性，但解决这一点并不困难，因为，根据二维语义，如果表达式 S 的第一内涵和第二内涵相同，则一个世界确证 S 当且仅当它满足 S，也就是说，只要 S 具备第一可能性，它同时也就具备了第二可能性。解决了这些关键问题，查尔默斯对传统的可设想性论证进行了修正。

^① 查默斯承认可能存在潜在的反例，但是那会导致一种“强必然性”。根据二维语义的分析，像“水是 H_2O ”这样后天必然性的例子都具有必然的第二内涵和偶然的第一内涵，也就是说，这些陈述具有形而上的必然性而不具有认知上的必然性，这种必然性可以成为“弱的后天必然性”；而如果这种后天必然命题的第一内涵也是必然的，那么这种必然性就是一种“强的后天必然性”，可以简称为“强必然性”。如果否认第一可设想性蕴含第一可能性，势必会导致这样一种强必然性，而这种强必然性很难得到辩护，详细讨论可参见本文第三部分。

^② Chalmers, D.J., Does Conceivability Entail Possibility?[A], In Gendler, T. & Hawthorne, J. (eds), *Conceivability and Possibility*[C], Oxford University Press, 2002, p.171.

2.4.3 修正的可设想性论证

通过对可设想性和可能性的二维语义分析，查尔默斯解决了CP论题的合理性，进而对传统的可设想论证进行了改进，论证过程为^①：

- (1) 怪人具有观念上肯定的第一可设想性；
- (2) 观念上肯定的第一可设想性蕴含第一可能性；
- (3) 怪人具有第一可能性（由1和2得出）；
- (4) 物理术语的第一内涵和第二内涵要么一致要么不一致；
- (5) 现象术语^②的第一内涵和第二内涵是一致的；
- (6) 如果物理术语的第一、第二内涵一致，则怪人具有第二可能性（由3和5得出）
- (7) 如果物理术语的第一、第二内涵不一致，则怪人不具有第二可能性，因而现象事实被更深层的实在所蕴含。
- (8) 要么怪人形而上可能，要么现象事实被更深层的实在蕴含（由4、6和7得出）
- (9) 如果怪人形而上可能，则物理主义错误。
- (10) 如果现象事实被更深层的实在蕴含，则物理主义错误。

结论是物理主义错误（由8、9、10得出）

这个论证的前提（2）即修正的CP论题，前提（4）是一个分析命题，至于前提（5），涉及克里普克对有关感知特性的概念的理解。克里普克认为“疼痛”等感知特性是我们意识中最简单最直接的，我们无法区分疼痛本身和对疼痛的感觉，而对于物理概念我们能够区分其属性与实在。查尔默斯采纳了这一观点，认为现象术语的第一内涵和第二内涵是一致的，这一点并不会引起什么疑义。^③因而关键在于分析当物理术语的第一内涵和第二内涵一致或不一致时将分别导致何种结论。

假定它们相同，根据二维语义，如果句子S的第一内涵和第二内涵相同，则一个世界只有在满足S的情况下才能确证S，也就是说S只有具备了第二可能性才能

^① See. Chalmers, D.J., *The Conscious Mind: in Search of a Fundamental Theory*[M], Oxford University Press, 1996, pp131-136; Also in his 'Does Conceivability Entail Possibility?'[A], In Gendler, T. & Hawthorne, J. (eds), *Conceivability and Possibility*[C], Oxford University Press, 2002. pp196-200.

^② Chalmers, D.J., *Consciousness and its Place in Nature*[A], In S. Stich & F. Warfield (eds), *Blackwell Guide to Philosophy of Mind*[C], Blackwell, 2003, p.117.

^③ 查尔默斯指出，前提（5）对于这一论证并不是必须的，即使现象术语的第一内涵和第二内涵并不一致，也能结论说物理主义错误。论证过程大致如下：首先假定物理术语的第一内涵和第二内涵一致，根据物理主义， $P \rightarrow Q$ 是后天必然的，按照二维语义的分析，这意味着 $P \rightarrow Q$ 的第一内涵是偶然的；令S'的第二内涵和S的第一内涵相同，则 $P' \rightarrow Q'$ 的第二内涵，即 $P \rightarrow Q'$ 的第二内涵是偶然的，因而 $P \rightarrow Q'$ 表达的命题也是偶然的；Q'当然为真，但是按物理主义，对于所有真句子S，“ $P \rightarrow S$ ”是必然的，因而物理主义错误。而如果物理术语的第一内涵和第二内涵不一致同样会导致元泛心论，因此物理主义错误。

具备第一可能性，既然 P 和 Q 的第一第二内涵都一致，那么“ $P \wedge \sim Q$ ”具备了第一可能性就必定具有第二可能性，因而怪人是形而上可能的，物理主义是错误的。

假定它们不同，这时候我们无法推出怪人是形而上可能的，而只能满足于怪人的第一可能性，但同样也会产生物理主义者无法接受的结果。物理主义坚持世界在微观层次上由基本粒子构成，如果物理术语尤其是微观物理学术语的第一第二内涵并不一致，那就意味着微观物理概念的指称通过特定的“理论角色”（即结构特征上的规定性）得以确定，但却指向这些角色背后的内在本质，也就是说，微观物理概念的第一内涵指向任意充当该角色的属性（不涉及内在本质），而第二内涵指向实际担当该角色的属性（也就是得到例示的属性，而不涉及其角色）。^①但是既然怪人只具有第一可能性，那就意味着“ $P \wedge \sim Q$ ”的第一内涵在某情形为真而第二内涵在所有世界为假，也就是说必定有某世界 W（如怪人世界）确证 P 且确证 $\sim Q$ 而没有世界满足 P 且满足 $\sim Q$ 。这样的话，世界 W 与我们的世界显然就是同构的（理论角色或者说微观物理层次上的结构相同），区别在于世界 W 中的“理论角色”由不同的内在属性来担任。这样问题就来了，我们的世界蕴含 Q，而物理上与我们的世界同构的世界 W 却不蕴含 Q，那么 Q 就不是由结构特征决定的，只能由不同的东西（即内在的属性）来蕴含，也就是说意识或感知特性萌生于内在的本质属性，这些属性的本质并不是通过感知或科学揭示给我们，而又与意识紧密地联系在一起，它们本身可能就是感知特性或元感知特性（即通过一定方式结合在一起从而构成感知特性）。这种观点查尔默斯称之为泛元心论（panprotopsychism），如果接受这一观点，那么，当我们想象怪人的时候，我们在现实世界的物理系统中所确定的只是结构属性，而非内在的元感知特性。显然，这种具有二元论倾向的观点很难被物理主义接受，而如果把把这些内在属性也视作物理属性又显得十分勉强^②，毕竟物理学对于这些内在属性究竟是什么没有任何有效的说明。

可见，不论物理术语的第一第二内涵是否一致，其产生的后果都是物理主义是无法接受的。如果一致，那么怪人就是形而上可能的，如果不一致，那就会导致元泛心论，两者对于物理主义都无法接受。就论证过程来看，这种基于二维语义的可想象性论证非常严密，关键在于查尔默斯通过二维方法使得 CP 论题更加精致，更加容易辩护。但这并不意味着这种修正的可想象性论证解决了所有难题。

^① Chalmers, D.J., Does Conceivability Entail Possibility?[A], In Gendler, T. & Hawthorne, J. (eds), *Conceivability and Possibility*[C], Oxford University Press, 2002, p.197.

^② 斯图加就要求扩展物理主义的内涵，他认为怪人是否可想象取决于我们如何理解物理属性，如果将这些内在属性也视为物理的，则怪人既不可想象也不可能。参见, Stojar, D., The conceivability Argument and Two Conceptions of the Physical[J], *Philosophical Perspectives*, 15(2001):393-413.

三 对查尔默斯论证的质疑和评述

3.1 怪人是否可想象

关于怪人的可想象性，巴罗格（Katalin Balog）提出了一个有趣的反对意见，其策略大致如下：由于怪人世界是我们世界的精确的物理摹本，那么物理主义在怪人世界显然是绝对正确的，而怪人哲学家同样能够在他们的世界提出一个查尔默斯式的可想象性论证以证明物理主义在他们的世界是错误的，如果查尔默斯的论证过程明确无误，那么这一矛盾就说明怪人论证的前提出了问题。^①查尔默斯认为原来的论证有效并不代表怪人哲学家的论证也有效，原先的论证前提中的Q指的是某人具有意识现象，但怪人哲学家的论证中的Q相应的也是指“某人具有意识现象”吗？而如果怪人世界的两位哲学家就意识现象或感知特性是否存在这一问题发生争论，哪一种观点应该得到支持呢？正如普特南所指出的：“可能世界中的人们虽然能够想和‘说’和我们所能想、所能说的任何话语，但……他们不能指称我们所指称的东西。”^②因此原来论证中正确的前提在怪人哲学家的论证中是错误的。

查尔默斯的辩护是有道理的，虽然巴罗格的反对意见颇有诡辩的意味，但却给我们很多重要的启示，虽然怪人世界作为一种认知“情形”在某种意义上（先天无法排除）是可想象的，或者说它具有观念的否定的可想象性，但是否也具有观念的肯定的可想象性呢？毕竟后者要求我们必须在观念上毫无矛盾地构想使怪人世界如此这般的条件，也就是说我们还必须为条件填充细节。这其中就有两个问题：首先，我们需要在多大程度上填充这样的细节？丹尼特认为，那些接受怪人可能性的人并没有充分地设想它们：“当哲学家们宣称怪人可想象时，他们总是低估了概念（或想象）的艰巨性，所想象的东西最终与他们原先的规定相违背。”^③而当我们充分地设想怪人的时候，立刻就会出现第二个问题——这种观念上的无矛盾的构想是否果真是完全无矛盾的？先天性并不等于观念或概念上的一致性，比如，“猫是动物”在概念上是协调一致的，但却并非先天的，“一个特定的类别是不是动物的一个种这样的问题，是一个要求经验研究的问题”。^④概念分析不可能完全摆脱经验知识的背景，当我们运用部分概念构想怪人以及怪人世界的时候，这些概念之间或许是协调的，但在更广泛的概念背景中，这些概念能否获得其一致性就很值得怀疑。我们知道，

^① Balog, K., *Conceivability, Possibility, and the Mind-Body Problem*[J], *Philosophical Review*, 108(1999), pp.510-514.

^② [美]普特南. 理性、真理和历史[M], 童世骏、李光程译. 上海译文出版社 2005. 第9页。

^③ Dennett, D., *The Unimagined Preposterousness of Zombies*[J], *Journal of Consciousness Studies*, 2(1995), p.322.

^④ [美]克里普克. 命名与必然性[M]. 梅文译. 上海译文出版社. 1998. 第101页。

在能量守恒和转化定律创立之前，人们认为机械永动机在概念上是协调的，而在热力学第二定律创立之前，人们又认为热能永动机在概念上是协调的。因此，概念上的一致性并不能保证怪人在肯定的意义上是可想象的。当然，诚如克里普克所指出的，当我们谈论可能世界时，我们“只能部分地描述可能世界”，^①完整地描述可能世界的全部细节既不必要也不可能，这一点即使在现实世界也不可能做到。但是不能完整的描述可能世界并不意味着我们不需要充分设想它们，提出一种假设比让我们相信它可能为真容易得多。巴罗格的反对意见实际上体现了我们对怪人的无知，事实上我们无法想象当一个物理层次上与我们一模一样但没有意识经验的孪生人站在我们面前时将是何等的情形，我们也无法预知这位孪生人的任何行为、动作、表情等等，可以说怪人的可想象性源于我们的直觉，但最终与我们的直觉相悖。

3.2 修正 CP 论题是否成立

查尔默斯通过二维语义修正了传统的 CP 论题，但他的具体论证仍然存在明显的缺陷和不足，主要在于他并没有对第一可设想性蕴含第一可能性的论证提出直接的证据，并且强必然性是否存在也引起了极大的争议。

3.2.1. 第一可设想性是否蕴含第一可能性

关于 CP 论题，我们发现查尔默斯对第一可想象性和第二可想象性的区分其实很取巧，他们分别与第一可能性与第二可能性密切相关，查尔默斯强调关于可想象性的讨论只能在陈述的层面进行^②，因此，一个陈述的可想象性实际上涉及两个方面，一是相关可能世界的可想象性，二是陈述在该世界中的真值。因此，当查尔默斯说陈述 S 具有第一可想象性，实际上是说，S 在想象的情形 W 中为真，而情形 W 本身是形而上可能的，或者说可想象世界与可能世界至少是共外延的，于是从可想象性推出可能性就是顺理成章的事情了。这里有两点值得注意，首先，与可想象性概念的前两种区分不同，第一可想象性和第二可想象性的区分一开始就和两种可能性的区分联系在一起，它们的定义方式甚至是一样的，我们说 S 具有第一可想象性，如果它在某情形中得到确证，也就是说 S 的第一内涵在该情形中为真，按照第一可能性的定义 S 当然也具有第一可能性。显然，第一可想象性蕴含第一可能性是预先规定好的，而不是通过证明给出的。而事实上，查尔默斯对这一关键问题也从来没有给出正面的论证，而只是反复强调反对这一蕴含将导致一种无法辩护的强必然性。其次，查尔默斯在这里有一个巧妙的转换，陈述 S 的可想象性变成了可能世界本身的可想象

^① [美]克里普克. 命名与必然性[M]. 梅文译. 上海译文出版社. 1998. 第 23 页。

^② Chalmers, D.J., *The Conscious Mind: in Search of a Fundamental Theory*[M], Oxford University Press, 1996, p.68.

性，而后者是可能世界语义学的根本，因而，查尔默斯说，即使克里普克也不否认“水不是 H_2O ”在认知上是可想象的。但普特南认为，我们对《命名与必然性》存在一种误读，克里普克实际上并不赞成说“水不是 H_2O ”是可想象的，而只是说可以想象“水不是 H_2O ”是可想象的，也就是说，如果用“ \diamond ”表示模态算子“可想象的”，则“ $\diamond\diamond p \rightarrow \diamond p$ ”是不能接受的，^①虽然它没有违反模态逻辑通常的规则。这里普特南强调的是，即使我们能够想象水不是 H_2O 的孪生地球，但这不等于说水不是 H_2O 直接就是可想象的，更不能说水不是 H_2O 是可能的。

这里更进一步的分歧实际上涉及克里普克和查尔默斯对可能世界的不同理解。克里普克和普特南坚持认为可能世界是“由我们赋予的描述条件给出的”，^②并不具有本体论的实在性。按照这一观点，像查尔默斯那样说“当某情形或可能世界成为现实的（或被视为现实）”是不恰当的，是对可能世界的误解，因此，克里普克的语义学考察本质上只能是一维的。很难说对于可能世界的理解哪一种更正确，但是取消现实世界的特殊地位更便于一般地讨论表达式在任意可能世界中的意义，因此查尔默斯的二维方法对于解决语义学中长期以来的一些争论会很有帮助，并且在二维框架中我们确实很难拒绝修正的CP论题，尽管我们注意到查尔默斯并没有给出CP论题的直接的肯定性的证明。

3.2.2. 强必然性是否存在

根据二维分析，像“水是 H_2O ”以及“启明星是长庚星”这样后天必然性的例子具有必然的第二内涵和偶然的第一内涵，这就是说，这些陈述具有第二必然性和第一偶然性：存在着确证其为错误的中心可能世界（孪生地球，或者启明星长庚星并不相同的世界）。如果S是这种具有偶然的第一内涵的后天必然命题，我们就可以说S是弱后天必然命题。相反，当且仅当S具有必然的第一内涵，我们就可以说它是强后天必然命题，强必然性指的就是在所有中心形而上可能世界都得到确证的后天必然性。

查尔默斯认为，很难给出这种强必然性的例子，所有克里普克的后天必然性的例子都是弱后天必然的^③。首先，很明显CP一意味着不存在强必然性，如果S是否定地可想象的同时又是第一可能的，则 $\sim S$ 将是强必然的。如果S是强必然的，则 $\sim S$

^① Putnam, H., Is Water Necessarily H_2O ?[A], in his *Realism With A Human Face*[C], Harvard University Press, 1990, p55.

^② [美]克里普克. 命名与必然性[M]. 梅文译. 上海译文出版社. 1998. 第23页。

^③ 在《意识的心灵》一书中，查尔默斯给出了更多说明：1、强必然性不能通过与其它后天必然性的类比而获得支持；2、它是一种比克里普克的例子更为极端的后天必然性，要求在世界层次上区分逻辑的和形而上的可能性；3、它导致模态概念的过度增生；4、对一致性提出强烈质疑；5、强必然性是粗暴的和无法解释的；6、在心身问题中保留强必然性的唯一动机是拯救物理主义。

就是否定地可想象的但第一不可能的。由于CP+较CP-更弱，因此CP+是否意味着不存在强必然性并不清楚。当然，任何CP+的反例都会产生强必然性，但反过来就未必了。为了说明这一点，我们可以定义两类强必然性，看它是产生CP+的反例还是只产生CP-的反例：

否定意义的强必然性：S 具有第一必然性和第二必然性但 $\sim S$ 在否定意义上可设想；

肯定意义的强必然性：S 具有第一必然性和第二必然性但 $\sim S$ 在肯定意义上可设想。

这样，所有肯定意义的强必然的都是否定意义的强必然的，但反过来却未必，CP-和CP+分别相应于不存在否定意义的强必然性和不存在肯定意义的强必然性这两个论题。

反对者提出一个强必然性的例证。上帝的存在是形而上必然的但并非先天的。根据这一观点，“上帝存在”（设为命题G）是后天必然的，G 具有第二必然性， $\sim G$ 在否定意义上是可想象的（当然在肯定的意义上也是可想象的）。如果G 是普通的后天必然命题，则G 具有第一偶然性：存在确证 $\sim G$ 的形而上可能世界。但如果不存在上帝，那就没有这样的世界。“上帝存在”的第一内涵和第二内涵看来并没有什么区别，因此，如果一个世界满足G，那么它也确证G。而如果G 具有第二必然性，那G 也具有第一必然性，这就得出如果上帝存在这一哲学观点如果正确，那它就是强必然的：至少是否定意义的强必然性，而如果加上上面肯定意义的可想象性的主张，它也是肯定意义上的强必然性。

查尔默斯认为，对于任何后天必然命题的S，存在一个确证 $\sim S$ 的情形。而如果S 是强必然的，虽然也存在一个确证 $\sim S$ 的情形^①，但这一情形并没有相对应的形而上可能世界。在上帝的例子中，虽然存在一个（否定意义和肯定意义上的）情形确证“不存在上帝”，但根据这一观点，不存在于此情形相对应的中心世界，也不存在本身就确证“不存在上帝”的中心世界。我们可以在直觉上将这一点表述为，按照这一观点，中心的形而上可能世界的空间比认知的可能情形的空间小，也就是说，存在着没有相应世界的情形。显然，这争论的关键落脚在认知可能性与形而上可能性之间的关系。

3.3 二维主义是否适用

^① 如果S 是肯定意义上的强必然性，则存在确证 $\sim S$ 的肯定意义的情形，如果S 是否定意义的强必然性，则存在确证 $\sim S$ 的否定意义的情形。

3.3.1 二维主义是否是描述主义

名称在所有可能世界指称同一对象而相应的描述语却未必，因此名称不同于描述语。这是克里普克反对描述主义的模态论证的主要内容，索姆斯认为，二维主义本身就是描述主义的一种表现形式，为了避免克里普克的模态论证，它把名称和自然种类词当成是伪装的固定描述语。按照这种理解，二维主义产生于确定指称的过程中，非固定性描述产生第一内涵，固定描述产生第二内涵。

但查尔默斯认为二维主义事实上并不依赖于这些描述主义的观点。尽管名称或自然种类词的第一内涵有时候似乎很接近描述语，比如，把“水”的第一内涵说成是“可饮用的透明液体”，凡是满足该描述的物质就是水的指称。但这仅仅是一种近似的用法，并不是说这种描述就是第一内涵本身。第一内涵严格来说只是一个函项，它体现的是语词中我们先天可把握的那部分意义，尽管这部分意义大抵只能通过描述语来表征。而且，一般认为名称的指称是名称的语言学意义的一部分，而相关描述语的指称却不是那些描述语的语言学意义的一部分，所以二者不是同义词。

因此，查尔默斯强调克里普克反对描述主义的论证并不适用于二维主义，但他并不否认二维主义和描述主义之间存在密切关系。和弗雷格一样，他认为指称并不是意义的全部，人们直觉上仍强烈地感觉到像“长庚星”和“启明星”这样共指称的名称，其意义在某些方面是有区别的，而且，即使它们不共指称在认知上也是可能的。表达式在认知上的区别与这些可能性联系在一起，因此，通过可能性和必然性的分析把握这些认知区别就是一条值得探索的途径。查尔默斯认为这应该成为二维主义的一个指导思想。索姆斯也承认当我们用这些词来表达思想的时候通常都会包含一些描述性的信息，但它们是而非语义的。查尔默斯则指出，虽然名称和自然种类词是直接指称，但这依赖于现实世界的呈现方式。

虽然在二维主义和描述主义之间直接划上等号并不合适，但说二维主义是一种弱的描述主义并不会引起什么争议。查尔默斯本人也承认这一点，同时他强调二维主义在继承描述主义的长处的同时又避免它的很多缺陷，比如，克里普克反对描述主义的模态论证和知识论证在二维框架内都可以得到很好的解释。

3.3.2 名称和自然种类词是否是索引性的

卡普兰之所以区分名称和索引词，原因在于前者的指称确定条件是决定它们意义的前语义因素，因而其指称的改变也就意味着意义的改变；而索引词的指称确定条件是语义的，因而指称在语境之间的改变和意义无关。索姆斯认为查尔默斯故意忽略了这一点，至少在《意识的心灵》中他确实把名称和自然种类词视为索引性的

固定描述语，而且这种描述语本身并不包含任何名称和自然种类词，它是纯粹定性的，因而“水”和“我”、“现在”这些索引词一样，其通常意义运用于不同的语境会产生不同的内容。索姆斯坚持克里普克的观点，认为“水”在孪生地球的同音异义表达式只是有着完全不同意义的不同表达式，名称和自然种类词都是固定指称。而卡普兰本人也是一开始就明确区分索引词和名称，反对把索引性扩展到名称和自然种类词。

从语境依赖的角度理解，索姆斯的批评并没有什么问题。因为名称和自然种类词毕竟与索引词有着明显的区别。在任何一种语言中名称和自然种类词的使用都具有特殊的重要性，这类表达式的指称不能由于语境而变得无法捉摸，否则人们将无法把握它，所谓知识也就无从谈起。如果仅仅是把卡普兰的语境原则简单地从索引词扩展到名称和自然种类词，那么这样一种二维主义将很难应对指称论的诘难，——尽管这样做能解释后天性与必然性之间的“矛盾”。但问题在于二维主义是否必然蕴含了名称和自然种类词具有索引性这一观点，如果二维主义能够避免这一结论，那么索姆斯的批评便失去目标。

首先，查尔默斯承认名称和自然种类词不是索引性的，其特性和语义指称是不变。同时，查尔默斯反复强调，他在《意识的心灵》一书中所描述的二维语义是不完全的，他本人也在《二维语义学基础》这篇文章中对语境的二维主义进行了细致的分析和批评，并明确倡导认知的二维主义。查尔默斯认为按照认知的理解，“水”这样的表达式在某特定语境中指称什么与第一内涵没有关系，因为第一内涵并不是卡普兰的特性或者语境内涵，重要的是孪生地球的某个特定描述在认知上蕴含了“水不是 H_2O ”。这样一种认知主张与“水”没有语境依赖和“水”不是索引词是相容的，这一点我们将在下面继续讨论。

3.3.3 认知内涵是否是私人的和非语义

索姆斯认为，在《意识的心灵》一书中二维主义是作为一个语义系统出现的，它把弗雷格的含义一分为二，区分出第一内涵和第二内涵，前者是从视为现实的世界到外延，后者是从视为反现实的世界到外延，这种区分实际上与卡普兰对表达式的表达语境和赋值条件的区分相一致。因而第一内涵与卡普兰的特性关系密切，而第二内涵则相应于卡普兰的语义内容。但实际上查尔默斯已经放弃了对第一内涵的语境式理解，转而采用认知概念来定义第一内涵。

索姆斯认为这样一种转变实际上是从公开的意义转到私人的思想，其思路是：尽管我们都使用公共语言 L ，但我们各自都有自己的内涵以决定我们在思想和交谈中

对 L 的运用。因此，尽管句子表达思想，但思想又不必和它们在语义上用 L 表达出来的东西混为一谈。索姆斯认为这种产生思想的系统实际上是一种私人语言，而认知的第一内涵就是把这种私人语言的特性运用于认知上可能的情形中而得到的。因此，所谓“水”和“长庚星”在不同的情形中指称不同物质，这并不是说它们在 L 中的语义指称改变了，而是说它们在认知的私人语言中的指称改变了，也就是说，名称和自然种类词具有一贯的特性和不变的语义指称，但它们还具有随情形而改变的认知指称。

于是这就一方面允许名称和自然种类词是非描述的、非索引性的，另一方面通过区分句子（包含 L 的语义性质）和句子的运用（在我们私人思想产生系统中）来解决关于认知意义以及先天性和后天性的问题，比如，在 L 中，“长庚星即启明星”在语义上可能表达一个先天可知的命题，即使它在我的私人的思想产生系统中不是先天的，因为我们平常使用这一思想来表达需要诉诸经验判断。

索姆斯把语境内涵到认知内涵的转变说成是从公开的意义到私人的思想、从语义的到非语义的转变。这种解读是成问题，首先，认知是通往知识的状态和过程，虽然包含一定的心理因素，但并非完全是任意的私人的思想。其次，意义的某个部分和认知领域联系在一起，这并不意味着这部分意义是非语义的，除非像索姆斯那样预先规定语义内容是不变的。正如查尔默斯强调的那样，这些问题涉及我们应该在何种意义上使用“语义的”这个词，而很难说对这个词的哪种运用更正确。

另外，查尔默斯显然仍是受了克里普克的启发，因为克里普克在解释“水是 H_2O ”的后天必然性时承认，虽然“水不是 H_2O ”形而上不可能，但在认知想象中是可能的。因此，从认知的角度入手的确能够避免把名称和自然种类词索引化，但由此而来的另一个问题是，如果“水不是 H_2O ”仅仅在认知想象中是可能的，那么二维主义就名不副实了。虽然关于名称不是索引词的论证不能驳倒二维主义，但查尔默斯需要证明凡先天可想象的都是形而上可能的。

3.3.4 认知上可能的是否形而上可能

克里普克认为我们只能从现实世界发现的对象和性质来理解可能性，对于某特定对象，其缺少偶然属性是形而上可能的，其缺少必然属性则形而上不可能，这两种情形之间并不存在直觉上和认知上的区别。形而上可能世界和认知可能世界都是宇宙可能具有的性质，并且只能通过关于现实个体和属性的约定来说明，也就是说，我们关于可能性的全部知识都源自现实世界中的个体和属性。索姆斯赞同克里普克的模态主张，在他看来，克里普克的后天必然性正是由于认识到某些认知可能状态

的形而上不可能性。也就是说，尽管我们需要关于本质属性的经验证据以排除某些认知可能世界，但对于其形而上不可能性我们却常常毫无意识，认识到这一点，二维主义就没有什么必要了。

索姆斯之所以像克里普克那样拒绝讨论仅仅认知上可能但形而上不可能的世界状态，是因为他关心的是在真正可能的各种条件下表达出来的命题，而不是由于错误使用而表达出来的命题。换言之，作为语境的只能是形而上可能世界，认知可能的世界状态是命题的赋值条件，后天必然命题在所有形而上可能世界为真，而在某些认知可能世界为假；先天偶然命题不需要现实世界的经验证据也能获知，即使它们在某些形而上可能世界中为假。索姆斯认为，二维主义的错误就在于把认知可能世界和形而上可能世界混为一谈，并由此解释后天必然性和先天偶然性，甚至把索姆斯对后天必然命题的解释也视为一种二维主义。

查尔默斯认为没有明显的证据表明认知可能性和形而上可能性之间具有不可通达性，也没有什么能妨碍我们把先天无法排除的认知可能情形视为现实的。而实际上我们往往很难分辨一种情形是否仅仅在认知上可能而不具有形而上可能性，至少克里普克和索姆斯贬低了认知可能性的意义。黄益民博士认为二维语义理论仍存在较大争议，解决这一问题只需区分“语义必然性”和“特性必然性”即可，^①但这一区分仍然于事无补，因为克里普克持本质主义立场，在“水是H₂O”这个句子中，H₂O是作为“水”这一物质的本质属性出现的，在克里普克看来，一事物缺乏偶然属性是形而上可能的，但缺乏本质属性是形而上不可能的。正如索姆斯强调的那样，理解这种仅仅认知上可能但形而上不可能的情形是理解后天必然性的关键。“特性必然性”似乎没有注意到克里普克对本质属性和偶然属性这一区分的强调，尽管对于具体的区分标准还存在争议。

3.3.5 先天性和认知可能性之间是否存在必然联系

索姆斯同意存在先天偶然命题，他给出的例子是包含模态算子“实际上”（用@表示，意即在现实世界或当下实际情形中）的句子，^②如“P当且仅当@P”，这个句子表达命题S：“P当且仅当P在@中为真”。S是偶然的，因此其第二内涵在某形而上可能世界W中为假，而W同时也是认知上可能的，因此S的第一内涵在W中也为假，可见先天性并不要求第一内涵在所有认知可能的情形中为真，因而二维主义的观点错误。

^① 参见：黄益民. 二维语义学及其认知内涵[J]. 哲学动态. 2007 (3)

^② Soames, S., *Philosophical Analysis in The Twentieth Century. Vol.2: The Age of Meaning [M]*. Princeton University Press.2003, pp417-422.

查尔默斯认为这一反驳是无效的，首先，索姆斯在解释后天必然性的时候说，如果命题 P 是偶然的和后天的，则后天必然命题 $@P$ 在某些认知可能的状态中为假。显然，索姆斯的观点前后矛盾。

其次，割裂认知可能性与先天性之间的联系也就是割裂了世界状态的认知可能性与命题的认知可能性之间的联系。因为如果某人得知命题 P ，他就排除了所有 $\sim P$ 的世界状态。同样，如果某人先天地知道了命题 P ，他就会先天排除所有 $\sim P$ 的世界状态，也就是说，如果 P 先天可知且在某世界 W 中 P 为假，那么它就会先天地排除 W 这一认知假设，换言之， W 就不是认知可能的，所以，只要 P 是先天的，它就在所有认知可能的世界为真。

索姆斯承认 T1 基本上来说是正确的，但是不能不加限制的使用，尤其是涉及到包含“@”算子的命题。我们能够从先天命题“ P 当且仅当 P ”推出命题“ P 当且仅当 $@P$ ”是先天偶然的，因为是现实世界 $@$ 中的主体而不是其它可能世界中的主体能够从 $@$ 中的真理 P 得出命题“ P 在 $@$ 中为真”，反之亦然，因此，在非现实世界中命题“ P 当且仅当 $@P$ ”之为假与其在 $@$ 中的先天可知性无关。查尔默斯认为这一策略同样无效，其他世界的主体能不能知道所谈论的问题并没有什么关系，重要的是在我们的世界中先天地知道“ P 当且仅当 $@P$ ”的主体知道什么并排除了什么。

另外，对于认知可能性这一概念，双方存在的分歧也是根本性的。索姆斯认为认知可能状态是一个世界可能具有的最大完全属性（*maximally complete properties*），这些属性取决于现实世界中实际存在的对象和属性，而查尔默斯的认知可能性是最大完全句子或命题，这一区别显然涉及从物模态和从言模态的争论。如果按照索姆斯的理解，认知可能性的基础是现实世界中的对象和属性，那么“长庚星不是启明星”的认知可能状态是不存在，而只要金星存在，“长庚星是启明星”就是先天的。查尔默斯认为这明显与我们的直觉相悖，毕竟“长庚星是长庚星”和“长庚星是启明星”在认知属性上是有区别的。

结 论

怪人假设其实是哲学上的一种思想实验，它与科学中的思想实验有着显著的区别。科学家们总是努力把思想实验变成一种使辩论双方都能接受的判决性实验，从而推动科学理论的进步。由于一种明显的非经验性特征，这种进步在哲学中是找不到的。通常，哲学中的思想实验非但不会消除争论达成共识，反而会使争论更加激烈，原因在于，哲学争论最终归结于本体论上根本信念和根本立场的差异，而这些信念是不可通约的，完全没有共同的基础。这一点在索姆斯和查尔默斯的争论中已经明显的表现出来。因此，思想实验的提出就已经包含了根本信念，并且其目的也是为了巩固这种根本信念，这样当然只能使争论升级。但哲学如果没有争论也就不成为哲学，哲学中的思想实验固然不可能彻底解决争论和分歧，但能够有力的推动哲学争论向更深的层面展开，以怪人假设为例，它能够帮助我们更加深入地思考意识的本质以及现象性质和物理属性性质的关系等问题，同时也对我们理解物理主义和反物理主义的争论非常有帮助。因此，哲学中的思想实验并不像阿特金森(Atkinson, D.)和佩吉伦伯格(Peijnenberg, J.)说的那样是“贫乏的”(poor)，只具有破坏性的后果^①。

围绕怪人的可设想性论证以及 CP 论题展开的争论涉及的内容非常广，已经远远超出心灵哲学的范围，查尔默斯引入的二维方法便涉及语义学、模态逻辑等，不同领域的理论和概念相互交叉融合，使得争论空前复杂。意识问题的研究也必定是沿着这种多领域多学科不断整合的趋势发展。本文涉及内容仅仅局限在查尔默斯的二维语义方法和可设想性论证的 CP 论题，更多的内容只能留待以后继续研究。

^① See Peijnenburg, J. & Atkinson, D., When are Thought Experiments Poor Ones?[J], *Journal for General Philosophy of Science*, 34(2003): 305-322.

参考文献

中文参考文献：

- [1] [法]笛卡尔. 第一哲学沉思集[M]. 庞景仁译. 北京：商务印书馆. 1986.
- [2] [英]赖尔. 心的概念[M]. 徐大建译. 北京：商务印书馆. 1992.
- [3] [美]哈特费尔德. 笛卡尔与《第一哲学的沉思》[M]. 尚建新译. 桂林：广西师范大学出版社. 2007
- [4] [美]克里普克. 命名与必然性[M]. 梅文译. 上海：上海译文出版社. 1998
- [5] [美]普特南. 理性、真理和历史[M], 童世骏、李光程译, 上海：上海译文出版社 2005
- [6] [英]休谟. 人性论. 关文运译[M]. 北京：商务印书馆 1980
- [7] [美]约翰海尔. 当代心灵哲学导论[M]. 高新民等译. 北京：中国人民大学出版社
- [8] 曹向阳. 当代意识科学导论[M]. 南京：东南大学出版社. 2003
- [9] 高新民、刘占锋等. 心灵的解构[M]. 北京：中国社会科学出版社. 2005
- [10] 高新民、储昭华主编. 心灵哲学[C]. 北京：商务印书馆. 2002.
- [11] 宫睿. 论休谟“想象”概念的多重含义[J]. 外国哲学. 2004 第十六辑. 34-53
- [12] 黄益民. 可想象性论证与后天必然性[J]. 云南大学学报 2007 (2): 40-47
- [13] 黄益民. 二维语义学及其认知内涵[J]. 哲学动态 2007 (3): 52-58
- [14] 李麒麟. 休谟在其因果原则论证当中的“可设想性原则”[J], 外国哲学 2005 第十八辑 180-190

英文参考文献：

- [1] Armstrong, D.M., The Mind-Body Problem: an Opinionated Introduction[M], Westview Press, 1999
- [2] Block, N. Are Absent Qualia Impossible? [J], *Philosophical review*, 89 (1980):257-274
- [3] Balog, K., Conceivability, Possibility, and The Mind-Body Problem[J], *Philosophical Review*, 108(1999): 497-528.
- [4] Brueckner, A., Chalmes's Conceivability Argument for Dualism[J], *Analysis*, 61(2001): 187-193.

-
- [5] Chalmers, D.J., *The Conscious Mind: In Search of a Fundamental Theory*[M], Oxford University Press, 1996
- [6] Chalmers, D.J., Materialism and The Metaphysics of Modality[J], *Philosophy and Phenomenological Research*, 59(1999):473-496.
- [7] Chalmers, D.J., Consciousness and its Place in Nature[A], In Stich, S. & Warfield, F. (eds), *Blackwell Guide to Philosophy of Mind*[C], Blackwell, 2003. Also in Chalmers, D.J. (ed), *Philosophy of Mind: Classical and Contemporary Reading*[C], Oxford University Press, 2002.
- [8] Chalmers, D.J., Does Conceivability Entail Possibility?[A], In Gendler, T. & Hawthorne, J. (eds), *Conceivability and Possibility*[C], Oxford University Press, 2002.
- [9] Chalmers, D.J., The Foundations of Two-Dimensional Semantics[A], In Garcia-Carpintero, M. & Macia, J. (eds), *Two-Dimensional Semantics: Foundations and Applications*[C], Oxford University Press, 2006.
- [10] Chalmers, D.J., Two-Dimensional Semantics[A], in Lepore, E. & Smith, B. (eds), *The Oxford Handbook of Philosophy of language*[C], Oxford University Press, 2006.
- [11] Chalmers, D.J., & Jackson, F., Conceptual Analysis and Reductive Explanation[J], *Philosophical Review*. 110(2001):315-61.
- [12] Cottrell, A., Sniffing the Camember: on The Conceivability of Zmbies[J], *Journal of Consciousness Studies*, 6(1999): 4-12.
- [13] Davis, M., Reference, Contingency, and The Two-dimensional Framework[J], *Philosophical studies*, 118(2004):83-131.
- [14] Dennett, D. C., The Unimagined Preposterousness of Zombies[J], *Journal of Consciousness Studies*, 2(1995): 322-326.
- [15] Dietrich, E., It on Seems as if Zombies Are Logically Possible, or How Consciousness Hides the Truth of Materialism: A Critical Review of The Conscious Mind [J], *Mind and Machines*, 8(1998): 441-461.
- [16] Evans, G., Comment on ‘Two Notions of Necessity’[J], *Philosophical studies* 118(2004):11-16.
- [17] Flanagan, O. & Polger, T., Zombies and The Function of Consciousness[J], *Journal of Consciousness Studies*, 2(1995): 313-321.

- [18] Gillett, G. and Loewer, B.(eds), *Physicalism and Its Discontent*[C], Cambridge University Press, 2001.
- [19] Guzeldere, G., Varieties of Zombiehood[J], *Journal of Consciousness Studies*, 2(1995): 326-333.
- [20] Harnad, S., Why and How We Are not Zombies[J], *Journal of consciousness studies*, 1(1994):164-167.
- [21] Hill, C.S., There Are Fewer Things in Reality than are Dreamt of in Chalmers's Philosophy[J], *Philosophical and phenomenological research*, 59(1999), 445-454.
- [22] Hill, C.S., Chalmers on the Apriority of Modal Knowledge[J], *Analysis*, 58 (1998):20-26.
- [23] Jackson, F., What Mary Didn't Know[J], *Journal of philosophy*, 83(1986):291-295.
- [24] Kirk, R., Zombies vs. Materialists[J]. *Aristotelian Society Proceedings*, 48(1974):135-152.
- [25] Kirk, R., From Physical Explicability to Full-blooded Materialism[J], *Philosophical Quarterly* 29(1979): 229-237.
- [26] Kirk, R., Physicalism, Identity and Strict Implication[J], *Ratio*, 24(1982): 131-141.
- [27] Kirk, R., Strict Implication, Supervenience, and Physicalism[J], *Australasian Journal of Philosophy*, 74(1996): 244-256.
- [28] Kirk, R., How Physicalists Can Avoid Reductionism[J], *Synthese*, 108(1996): 157-170.
- [29] Kirk, R., Nonreductive Physicalism and Strict Implication[J], *Australasian Journal of Philosophy*, 79(2001):545-553.
- [30] Levine, J., Conceivability and The Metaphysics of Mind[J], *Noûs* 32(1998):449-480.
- [31] Loar, B., David Chalmers's the Conscious Mind[J], *Philosophical and phenomenological research*, 59(1999):465-472.
- [32] Lycan, W. G., Stalnaker on Zombies[J], *Philosophical studies*, 133(2007):473-479.
- [33] Marcus E., Why Zombies are Conceivable[J], *Australasian Journal of Philosophy*, 82(2004):477-490.
- [34] Maslin, K.T., *An Introduction to the Philosophy of Mind*[M], Polity Press, 2007.
- [35] McGinn, C., *Minds and Bodies: Philosophers and Their Ideas*[M], Oxford University Press. 1997.

- [36]Moody, T., Conversation with Zombies[J], *Journal of consciousness studies*, 1(1994):196-200.
- [37]Peijnenburg, J & Atkinson, D., When are Thought Experiments Poor Ones?[J], *Journal for General Philosophy of Science*, 34(2003): 305-322.
- [38]Putnam, H., Is Water Necessarily H₂O?[A], Conant, J.(ed), *Realism With A Human Face*[C], Harvard University Press, 1990.
- [39]Stojar, D., The conceivability Argument and Two Conceptions of the Physical[J], *Philosophical Perspectives*, 15(2001):393-413.
- [40]Stalnaker, R., What is it Like to be a Zombie?[A], In Gendler, T. & Hawthorne, J. (eds), *Conceivability and Possibility*[C], Oxford University Press, 2002.
- [41]Shoemaker, S., On David Chalmers's the Conscious Mind[J], *Philosophical and phenomenological research*, 59(1999): 439-444.
- [42]Soames, S., *Philosophical Analysis in The Twentieth Century*, Vol.2: The Age of Meaning[M], Princeton University Press, 2003.
- [43]Soames, S., *Reference and Description: The Case Against Two-Dimensionalism*[M], Princeton University Press, 2005.
- [44]Worley, S., Conceivability, Possibility and Physicalism, *Analysis*, 63(2003): 15-23.
- [45]Wong, Kai-yi, Two-dimensionalism and Kripkean A Posteriori Necessity[A]. Gendler, T. & Hawthorne, J. (eds), *Conceivability and Possibility*[C], Oxford University Press, 2002.
- [46]Woudenberg, R., Conceivability and Model Knowledge[J], *Metaphilosophy*, 37(2006):210-221.
- [47]Yablo, S., Is Conceivability a Guide to Possibility?[J], *Philosophical and Phenomenological Research*, 53(1993):1-42.
- [48]Yablo, S., Concepts and Consciousness[J], *Philosophical and Phenomenological Research*, 59(1999), 455-463.
- [49]Yablo, S., Coulda, Woulda, Shoulda[A]. Gendler, T. & Hawthorne, J. (eds), *Conceivability and Possibility*[C], Oxford University Press, 2002.

致 谢

论文的完成首先要感谢导师魏屹东教授，学术和论文方面的悉心指导自不待言，他严谨的治学态度和雷厉风行的作风更深深影响了我，又给我充分的时间和自由研究自己感兴趣的问题。山西大学哲学社会学学院拥有一批学识渊博的教师，三年学习期间我从他们那里受益良多。宋炳延老师儒雅谦逊，使我明白哲学同样是一种生活方式；安希孟教授文字激扬，时时促我前行；逻辑方面遇到的困难则常常聆听毕富生教授的教导，胡瑞娜副教授也对我的论文提出宝贵的修改意见。社科院哲学所的黄益民研究员和北京大学的哲学硕士仲海霞也对我的论文提出了中肯的意见，在此一并向他们表示深深的感谢。

感谢我的父母，无论是工作还是考研，他们一直都尊重并支持我的选择；同时要感谢我的哥哥陈敬全，三次在北京国家图书馆复印资料的时候都得到他的细心照料，还有武汉大学的同学刘涛，为我在武大图书馆查找资料提供了诸多便利。最后感谢我周围的同学：柯进华、侯志忠、周以嵘、翟鹏、赵晓聃、李瑞艳等等，无法一一列举，三年期间得到他们很多帮助，也建立了深厚的友谊。

附 录

发表文章：

直指和描述之间：评索姆斯和查尔默斯关于二维语义的争论. 洛阳师范学院学报.
2008（1）：26-31

个人简介：

姓名：陈敬坤

性别：男

籍贯：湖北省黄石市大冶市

1998.9—2002.6 就读于华中师范大学中文系

2002.6—2004.6 就职于湖北多佳股份有限公司

联系方式：

电话：13994276672

Email: jingkun_chen@yahoo.com.cn