

信念修正理论视角下的绿蓝悖论

杜晓柳

摘 要: AGM 信念修正理论是一种刻画主体信念状态及其变动的理论。通过适当的修正和扩充, 它也可以用来描述主体进行枚举归纳活动时的信念。绿蓝悖论这一问题中的归纳主体的信念变动情况也可以通过这种修改后的 AGM 方法进行刻画。尽管这种刻画实际上并没有消除绿蓝悖论, 却也提供了一种新的视角解析问题, 并清晰明了地呈现出了绿蓝悖论的根源所在。

关键词: AGM 理论; 绿蓝悖论; 枚举归纳

中图分类号: B81

文献标识码: A

1 前言

自从绿蓝悖论这一问题在上世纪中叶被提出之后, 学者们就不断提出各种方案试图来解决它。不同于其中大多数解决思路, 辛克 (A. Zinke) 在信念逻辑的框架内提出了一项解悖方案。他在经汉森 (S. O. Hansson) 修正后的 AGM 理论¹基础上, 提出了一套方法来刻画信念主体在做归纳时的信念变动情况, 并认为通过这种刻画可以消解绿蓝悖论。([7]) 本文作者认为辛克的方法并没有解决绿蓝悖论, 但是他提出的这种刻画方法可以清晰地显现出绿蓝悖论的症结所在, 也可以帮助我们更深入地讨论该问题。

遵循以上思路, 本文的第二部分将阐述原始的 AGM 理论和它的部分缺陷, 以及汉森是如何试图修正这些缺点的。第三部分将介绍辛克如何将修正过后的 AGM 理论应用到归纳法情景中, 以及他以此来解决绿蓝悖论的尝试。第四部分将分析揭示辛克的这种方法不能解决绿蓝悖论, 不过却从另外一方面表明了绿蓝悖论的产生原因。最后一部分将对绿蓝悖论可能的解决方案作进一步的讨论。

收稿日期: 2022-09-13

作者信息: 杜晓柳 浙江大学哲学学院
250818965@qq.com

¹AGM 理论最早由阿尔罗若 (C. E. Alchourron)、加德福斯 (P. Gardenfors) 和梅金森 (D. Markinson) 三位学者提出并发展, 故学界取他们三位的姓名缩写来指称该理论。([9])

2 AGM 理论及其修正

AGM 理论是一套用命题语言来刻画主体信念及其变动的理论。首先它假定人们的信念都是些命题语句,再用一个一致的命题语句集合 K_A 来表述主体 A 的信念²,并称之为信念集。例如, $K_A = \{\varphi, \psi, \omega\}$ 就表示主体 A 相信了命题 φ 、 ψ 和 ω 。

此外,信念集还是演绎闭包的。³即,如果 A 相信 φ 和 $\varphi \rightarrow \psi$,那么 K_A 就不仅包含了 φ 和 $\varphi \rightarrow \psi$,还必须要包含 ψ 。

在 AGM 理论中,如果想要修改一个信念集以表示主体的信念发生了变动,那么需要经过两个步骤。第一步称之为收缩,第二步称之为扩张。扩张 (+) 非常容易理解,它表示主体增加相信了一个命题。例如我们想把 ω 添加到信念集 K_A 中,只要令新的信念集 $K_{A'} = \text{Cn}(K_A \cup \{\omega\})$ 即可(当然,为了保持信念集的一致性, ω 必须是不与 K_A 中已有语句相矛盾的语句)。记此过程为 $K_A + \omega$ 。

收缩 ($\div$) 是一个相对复杂的概念,它表示主体不再相信原来曾相信过的某个命题。例如,如果主体 A 不再相信 ω ,那么 K_A 中不仅要去掉命题 ω ,还要去掉能推出 ω 的那些命题。但是,如果 $K_A = \{\varphi, \psi, \varphi \rightarrow \psi\}$,那么当 A 不再相信 ψ 时,在 K_A 中除了去掉 ψ 之外,我们还应该去掉 φ 还是 $\varphi \rightarrow \psi$ 呢?这就要求理论为此做出额外的规定。首先我们定义 $K_A \perp \omega$ 为由所有推不出 ω 的、 K_A 的极大子集所构成的集合。选择函数 δ 被定义如下:

1. 如果 $K_A \perp \omega = \emptyset$, 那么 $\delta(K_A \perp \omega) = \{K_A\}$ 。
2. 如果 $K_A \perp \omega \neq \emptyset$, 那么 $\delta(K_A \perp \omega) \neq \emptyset$ 且 $\delta(K_A \perp \omega) \subseteq (K_A \perp \omega)$ 。⁴

然后我们定义收缩, $K_A \div \omega = \bigcap \delta(K_A \perp \omega)$ 。例如对于 $\{\varphi, \psi, \varphi \rightarrow \psi\} \div \{\psi\}$,如果主体决定继续相信 φ 而放弃相信 $\varphi \rightarrow \psi$,那么就规定 $\delta(\{\varphi, \psi, \varphi \rightarrow \psi\} \perp \psi) = \{\varphi\}$,最终得 $\{\varphi, \psi, \varphi \rightarrow \psi\} \div \{\psi\} = \{\varphi\}$ 。

总的来说,在 AGM 理论中,如果主体信念发生了变动,例如,当 A 决定将命题 ω 纳入自己的信念集时,我们首先要用 $\neg\omega$ 对 K_A 进行收缩,再用 ω 对 K_A 进行扩张。将这一变动过程记为 $K_A * \omega$,即 $K_A * \omega = (K_A \div \neg\omega) + \omega$ 。

这就是标准的 AGM 理论。([9])当然,它也具有很多缺陷。其中之一可以用如下例子来加以说明。

假设有两个信念主体:宋和陈。其中,宋相信哈维和玛丽都会去参加聚会。而陈相信以下两点:玛丽会去参加聚会;并且哈维和玛丽两人要么都参加聚会,要么都不参加聚会。

²当然,这意味着该理论默认一个主体的信念不会是自相矛盾的。

³在本文中,我们用 $\text{Cn}(K)$ 来表示集合 K 的演绎闭包。

⁴注意,这里只是定义了 δ 的取值范围,并没有规定 δ 到底要取 $K_A \perp \omega$ 中的哪些元素。

用 K_A 代表宋的信念集, 用 K_C 代表陈的信念集, h 代表哈维参加聚会, m 代表玛丽参加聚会, 又考虑到 AGM 理论要求信念集都是演绎闭包的, 因此 $K_A = \text{Cn}\{h, m\}$, $K_C = \text{Cn}\{m, h \leftrightarrow m\}$ 。很显然, 这两个集合相等。

现在, 假设宋和陈都得知了玛丽不会去参加聚会, 即他们都需要根据 $\neg m$ 来修改自己的信念集。如果仅根据我们的直觉来判断, 那么显然, 宋此时应该仍然相信哈维会去聚会, 而陈则应该相信哈维不会去聚会。但是在 AGM 理论中, 这一事实并不是那么明显。由于 $K_A = K_C$, 所以 $K_A * \neg m$ 和 $K_C * \neg m$ 实际上是对同一个集合进行修正; 而为了达到 $K_A * \neg m = \{h, \neg m, \neg(h \leftrightarrow m)\}$ 且 $K_C * \neg m = \{\neg h, \neg m, h \leftrightarrow m\}$ 的目的, 我们必须要求

$$\delta_A(K_A \perp \neg m) = \{h\}, \delta_C(K_C \perp \neg m) = \{h \leftrightarrow m\}.$$

注意, AGM 理论本身并没有给出挑选选择函数的标准, 所以这里完全是根据外部经验来设定选择函数的。也就是说, 当我们用 AGM 理论刻画了某个主体的信念后, 当该主体新接纳了一个命题时, 常常需要根据我们的外部判断来设定选择函数, 才能继续刻画出动变后的主体信念。

虽然通过上述设定选择函数的方式可以达到“宋相信哈维会去聚会, 而陈则不相信”的目的, 但是这种方式因其过于累赘而不能令人感到非常满意。在这个例子中我们可以注意到, 宋本身直接的信念是“哈维和玛丽都会去参加聚会”, 之所以 $h \leftrightarrow m$ 被包含在 K_A 中, 完全是因为 $\{h, m\}$ 的演绎闭包包含了 $h \leftrightarrow m$ 。所以宋应该自然地将 h 和 m 置于更高的位置, 当他要在它们和 $h \leftrightarrow m$ 之间做选择时, 应当不需要选择函数就抛弃 $h \leftrightarrow m$ 而保留 h 和 m 。

由于辛克刻画绿蓝悖论的方法建立在汉森的 AGM 修正理论基础上, 所以我们这里要介绍汉森的修正理论。不同于标准的 AGM 理论, 它把原先的信念集划分成两层: 基础信念集 B 和信念集 K 。 B 由且只由那些并非完全根据其它信念命题推导才得来的信念命题组成; K 是 B 的演绎闭包, 而 B 本身则不要求是演绎闭包的。此外, 所有的信念修正活动(扩张和收缩)都发生在基础信念集中。([4])

回到刚刚的例子中, 由于基础信念集不再包含仅仅根据其它信念推导得来的命题, 而只包含了主体直接相信的命题, 所以我们可以说宋的基础信念集 $B_A = \{h, m\}$, 陈的基础信念集 $B_C = \{m, h \leftrightarrow m\}$ 。 h 不属于 B_C 是因为在陈的脑海中, “哈维会来参加聚会”是根据“玛丽会去参加聚会”和“哈维和玛丽两人要么都参加聚会, 要么都不参加聚会”这两条信念推导得来的, 而基础信念集又不需要进一步扩张以获得演绎闭包的特性, 所以 h 不属于 B_C 。现在, 当我们要根据 $\neg m$ 分别修正这两个基础信念集时, 就可以无异议地说, $B_A * \{\neg m\} = \{h, \neg m\}$, $B_C * \{\neg m\} = \{h \leftrightarrow m, \neg m\}$ 。

3 归纳场景中的信念修正理论

辛克将这一修正后的信念修正理论应用到枚举归纳法的使用场景中, 并认为这给绿蓝悖论带来了一种解决方案。然而实际上, 他的方法并没有解决绿蓝悖论——但也为我们提供了一种工具来深入分析绿蓝悖论的产生原因。

为了清楚地说明这一点, 让我们先来考察一下辛克是如何用修正后的 AGM 理论, 来描述主体在使用枚举归纳法时的信念变动。

首先仍然假设有两个信念主体宋和陈, 并且到当前为止, 一共已经举办过第 $1, 2, \dots, n$ 次聚会, 而玛丽和哈维参加了每一次聚会。

继续假设, 对于宋来说, 他脑海中由此而来所直接相信的——换言之, 就是他的基本信念是, 玛丽参加了第 $1, 2, \dots, n$ 次聚会, 哈维参加 $1, 2, \dots, n$ 次聚会; 同样假设, 对于陈来说, 他的基本信念是玛丽参加了第 $1, 2, \dots, n$ 次聚会, 并且对于第 $1, 2, \dots, n$ 次聚会来说, 哈维和玛丽两人要么都参加了这次聚会, 要么都没有参加这次聚会。基于文献 [7] 提出的修正理论, 他们的基础信念集可以分别表示成 $B_A = \{m_1, m_2, \dots, m_n, h_1, h_2, \dots, h_n\}$, $B_C = \{m_1, m_2, \dots, m_n, m_1 \leftrightarrow h_1, m_2 \leftrightarrow h_2, \dots, m_n \leftrightarrow h_n\}$ 。

辛克所做的事情是定义出一种方法, 使得我们能够根据基础信念集中的命题进行枚举归纳。⁵ 称命题 p_{n+1} 可以根据集合 B 枚举归纳得来, 当且仅当对任意 $i \in \{1, 2, \dots, n\}$ 都有 $p_i \in B$ 成立。而称一个集合 B^+ 是 B 的归纳闭包, 当且仅当 B^+ 中任意命题 p 都满足:

1. $p \in B$ 或者 p 可以根据 B 枚举归纳得来;
2. $\{p\} \cup B$ 是一致的。

当然, 集合的归纳闭包不一定是唯一的。例如对于 $B = \{p_1, p_2, q_1, q_2, \neg(p_3 \wedge q_3)\}$, 则既可以根据添加 p_3 而得到 B^+ , 也可以根据添加 q_3 而得到 B^+ 。可以通过定义选择函数的方式来选定归纳闭包集合。具体来说, 我们可以首先令 $M_B = \{B' \mid B' \text{ 是 } B \text{ 的一个归纳闭包}\}$, 即用 M_B 来表示由 B 的所有归纳闭包所组成的集合。再规定 $\delta(M_B) \in M_B$ 为 M_B 的选择函数。可以注意到 $\delta(M_B)$ 实际上就是选中了 M_B 的一个元素。我们再令 $I(B)_\delta = \delta(M_B)$, 显然 $I(B)_\delta$ 就表示了 $\delta(M_B)$ 所选中的那个归纳闭包。在本文中, 将使用 $I(B)$ 来指称集合 B 的某个确定的归纳闭包集。

现在回到刚刚的例子当中。很显然的是, $I(B_A) = \{m_1, m_2, \dots, m_n, m_{n+1}, h_1, h_2, \dots, h_n, h_{n+1}\}$, $I(B_C) = \{m_1, m_2, \dots, m_n, m_{n+1}, m_1 \leftrightarrow h_1, m_2 \leftrightarrow h_2, \dots, m_n \leftrightarrow h_n, m_{n+1} \leftrightarrow h_{n+1}\}$ 。也就是说根据宋的基础信念, 可以归纳得出结论“玛

⁵他同时也强调他这样做不是为了创建一套理论, 而只是为了提供一种绿蓝悖论的解决方法。

丽参加第 $n+1$ 次聚会”和“哈维参加第 $n+1$ 次聚会”，而根据陈的基础信念可以得到“玛丽参加第 $n+1$ 次聚会”和“哈维和玛丽两人要么都参加第 $n+1$ 次聚会，要么都不参加第 $n+1$ 次聚会”。

由于信念集根据其定义是一致的，所以作为子集的基础信念集也是一致的，它的归纳闭包集当然也是一致的。所以 AGM 理论中的扩张和收缩步骤对这些集合仍然是适用的。

当宋和陈在第 $n+1$ 次聚会上发现玛丽不参加时，可以用 $\neg m_{n+1}$ 分别修正 $I(B_A)$ 和 $I(B_C)$ 。注意到 $I(B_A) * \neg m_{n+1} = \{m_1, m_2, \dots, m_n, \neg m_{n+1}, h_1, h_2, \dots, h_n, h_{n+1}\}$ ，而 $I(B_C) * \neg m_{n+1} = \{m_1, m_2, \dots, m_n, \neg m_{n+1}, m_1 \leftrightarrow h_1, m_2 \leftrightarrow h_2, \dots, m_n \leftrightarrow h_n, m_{n+1} \leftrightarrow h_{n+1}\}$ 。并且我们的常识判断，最终也的确应该是宋继续相信哈维将参加第 $n+1$ 次聚会，但陈则相信哈维将不会参加第 $n+1$ 次聚会。这与该理论的刻画结果一致。

总的来说，我们可以看到在这种理论的刻画下，根据基本信念“玛丽参加了第 $1, 2, \dots, n$ 次聚会，哈维参加 $1, 2, \dots, n$ 次聚会”会归纳得出结论“玛丽参加第 $n+1$ 次聚会”和“哈维参加第 $n+1$ 次聚会”；而根据基本信念“玛丽参加了第 $1, 2, \dots, n$ 次聚会，并且对于第 $1, 2, \dots, n$ 次聚会来说，哈维和玛丽两人要么都参加了这次聚会，要么都没有参加这次聚会”则会归纳得出结论“玛丽参加第 $n+1$ 次聚会”和“哈维和玛丽两人要么都参加第 $n+1$ 次聚会，要么都不参加第 $n+1$ 次聚会”。并且在接受到新证据“玛丽不参加第 $n+1$ 次聚会”后，前者会继续认为哈维将参加第 $n+1$ 次聚会，但后者则不再相信这一点。

4 对绿蓝悖论的解析

绿蓝悖论这一问题中的谓词“绿蓝”通常定义为：

x 是绿蓝的，当且仅当 x 是绿的并且在时刻 t 之前第一次被观测，或者 x 是蓝的并在时刻 t 之前未被观测过。([3]) 除此之外还有一些其它的定义形式。([5]) 辛克在试图解决该问题时，使用的是如下这一较为少见的定义形式：

x 是绿蓝的，当且仅当， x 是绿的 $\leftrightarrow x$ 已被观测过。([7])

为了方便讨论，本文会继续沿用他的定义。

蓝绿悖论则被表述如下：假设我们发现了一种新的矿石，并且我们观测了这类矿石的第 $1, 2, \dots, n$ 块矿石，发现它们都是绿色的，于是根据枚举归纳法，似乎可以归纳得出结论说：“下一块这类矿石也是绿色的”。但是根据刚刚给出的定义，我们实际上也能说这类矿石的第 $1, 2, \dots, n$ 块矿石都是绿蓝的，那么根据同样的枚举归纳步骤，我们应该也能说：“下一块这类矿石也是绿蓝的”。由于下一块这类矿石还未被观测过，这句话实际上就是在说：“下一块该类矿石不是绿的”。

所以,根据同样的证据,我们既能归纳得出结论说下一块矿石是绿色的,也能归纳得出结论说下一块矿石不是绿色的。很显然这两个结论是互相矛盾的。

辛克认为,根据上一节所描述的那种应用于枚举归纳场景的 AGM 修正理论,可以很好的说明为什么“下一块这类矿石也是绿色的”类型的归纳结论是合理的,而“下一块这类矿石不是绿的”则是不合理的。

用上一节中介绍的方法,我们同样可以刻画出绿蓝悖论问题中这些进行枚举归纳活动的主体的信念变动状态。事实上,根据文献 [1] 和 [2],绿蓝悖论可以看成是上一节中例子的一个特例;绿蓝悖论中两条互相冲突的预测,实际上就是上一节中互相矛盾的 $I(B_A) * \neg m_{n+1}$ 和 $I(B_C) * \neg m_{n+1}$ 。下面我们将更清楚地看出这一点。

如果我们用 O_i 表示第 i 块矿石被观测过了, G_i 表示第 i 块矿石是绿色的,那么对于一个使用枚举归纳法的“正常”主体 A 来说,既然观测了第 $1, 2, \dots, n$ 块矿石且发现它们都是绿色的,那么他的基础信念集 B_A 应当为 $\{O_1, O_2, \dots, O_n, G_1, G_2, \dots, G_n\}$ 。自然的,有 $I(B_A) = \{O_1, O_2, \dots, O_n, O_{n+1}, G_1, G_2, \dots, G_n, G_{n+1}\}$ 。又因为第 $n+1$ 个矿石还没有被观测,即 $\neg O_{n+1}$,有 $I(B_A) * \neg O_{n+1} = \{O_1, O_2, \dots, O_n, \neg O_{n+1}, G_1, G_2, \dots, G_n, G_{n+1}\}$ 。所以在 A 看来,在尚未观测第 $n+1$ 块矿石时,归纳得出结论“第 $n+1$ 块矿石是绿色的”仍然是成立的,但是他并不相信“第 $n+1$ 块矿石是绿蓝的”,因为命题 $O_{n+1} \leftrightarrow G_{n+1}$ 和 $I(B_A) * \neg O_{n+1}$ 是不一致的,而根据辛克的定义,命题 $O_{n+1} \leftrightarrow G_{n+1}$ 其实就是在说第 $n+1$ 块矿石是绿蓝的,所以 $I(B_A) * \neg O_{n+1}$ 无法包含 $O_{n+1} \leftrightarrow G_{n+1}$,就表示在尚未观测第 $n+1$ 块矿石时,命题“第 $n+1$ 块矿石是绿蓝的”与 A 的基础信念集是不一致的。

辛克认为这里值得注意的是“ $\neg O_{n+1}$ ”([7]),如果忽视了这一点而直接用命题 $O_{n+1} \leftrightarrow G_{n+1}$ 和 $I(B_A)$ 相比对的话,会发现该命题和 $I(B_A)$ 并不矛盾,从而误以为“第 $n+1$ 块矿石是绿蓝的”也是一个可以接受的归纳结论。但是正确的对比方法应该是将归纳结论和 $I(B_A) * \neg O_{n+1}$ 相对比,这样就可以排除掉像“第 $n+1$ 块矿石是绿蓝的”这类有问题的归纳结论。因此,辛克认为绿蓝悖论是“隐藏在归纳法幌子下的信念修正问题”([7]),只要正确使用信念修正理论,就可以排除掉有问题的归纳结论。

本文不否认忽视“ $\neg O_{n+1}$ ”会产生上述问题,但认为辛克的这套方法并没有真正解决绿蓝悖论,而且绿蓝悖论最大的问题也不在于忽视“ $\neg O_{n+1}$ ”上。首先记“第 i 块矿石是绿蓝的”为 $Grue_i$,如果一个主体 B 使用这种绿蓝谓词来描述证据,那么他得到的是 $O_1, O_2, \dots, O_n, Grue_1, Grue_2, \dots, Grue_n$ 。可得 $I(B_B) = \{O_1, O_2, \dots, O_n, O_{n+1}, Grue_1, Grue_2, \dots, Grue_n, Grue_{n+1}\}$ 。根据定义, $Grue_{n+1}$ 等价于 $O_{n+1} \leftrightarrow G_{n+1}$,所以可知 $O_{n+1} \leftrightarrow G_{n+1}$ 和 $I(B_B) * \neg O_{n+1}$ 是无矛盾的,反而命题 G_{n+1} 和 $I(B_B) * \neg O_{n+1}$ 是有矛盾的。也就是说,在这种视角下“第 $n+1$

块矿石是绿色的”才是有问题的那个归纳结论。

辛克用了另一种方式来刻画用绿蓝谓词描述归纳证据的主体信念。([7])他认为如果主体 C 说绿蓝式谓词,那么他得到的证据应该是 $O_1, O_2, \dots, O_n, O_1 \leftrightarrow Grue_1, O_2 \leftrightarrow Grue_2, \dots, O_n \leftrightarrow Grue_n$ 。如此得到的 $I(B_C) = \{O_1, O_2, \dots, O_n, O_{n+1}, O_1 \leftrightarrow Grue_1, O_2 \leftrightarrow Grue_2, \dots, O_n \leftrightarrow Grue_n, O_{n+1} \leftrightarrow Grue_{n+1}\}$ 。这样, $Grue_{n+1}$ 就和 $I(B_B) * \neg O_{n+1}$ 有矛盾了。

但是注意, $O_1 \leftrightarrow Grue_1$ 即是在说“第1块矿石是绿色的”,尽管这句话是用绿蓝谓词说出来的,但它表达的其实是主体 A 眼中的证据。换言之,辛克的这种刻画方式只不过是描述一个使用绿蓝谓词的主体根据证据“第1, 2, $\dots$, n 块矿石是绿色的”做枚举归纳而已,得到的结论当然是“第 $n+1$ 块矿石是绿色的”。而绿蓝悖论的真正问题在于,证据“观测了第 i 块矿石,它是绿色的”,在使用绿蓝式谓词的人眼中,实际上是证据“第 i 块矿石是绿蓝的,它被观测过了”,也就是我们刚刚看到的主体 B 的情况。而且根据先前的分析, $Grue_{n+1}$ 和 $I(B_B) * \neg O_{n+1}$ 是无矛盾的。

因此,通过这种信念修正理论来刻画绿蓝悖论,并不能够解决问题。不过,这种刻画方式却也凸显了产生绿蓝悖论的根源所在。

事实上,根据这种信念理论的刻画,“第 $n+1$ 块矿石会是绿蓝色的”这种归纳结论不被我们的信念所接受,其原因在于我们直觉所选择的证据是“观测了第1, 2, $\dots$, n 块矿石,它们都是绿色的”(即信念集 B_A),而不是证据“第1, 2, $\dots$, n 块矿石是绿蓝的,它们都被观测过了”(即信念集 B_B)——尽管这两个命题是等价的,但是我们直接相信的,即纳入基础信念集的,是前一类命题而不是后一类命题,因为当我们事实上看到一块绿色矿石时,直觉会认为“这块矿石是绿色的”,而不是“这块矿石是绿色的 $\leftrightarrow$ 这块矿石被观测过”。所以当它们各自得到的归纳结论互相冲突时,我们才会相信前一类命题得到的归纳结论(因为它是根据我们直觉所选择的证据归纳得来的),并且因为后一类命题得到的归纳结论与前一类结论(即例子中的 $I(B_A) * \neg O_{n+1}$)相矛盾,而认为后一类结论是不合理的。

反过来,我们也可以看出绿蓝悖论的诞生,来源于互相逻辑等价的描述证据的命题,经过归纳法程序之后有可能会得到互相矛盾的归纳结论。套用这里的信念修正理论,就是说两个互相逻辑等价的基础信念集 B_A 和 B_B ,它们对应得到的 $I(B_A) * \neg O_{n+1}$ 和 $I(B_B) * \neg O_{n+1}$ 有可能是互相矛盾的,而我们的直觉又完全倾向于其中一方:比如说我们直觉完全倾向于采用 B_A 来描述看到的归纳证据,所以会觉得与 $I(B_A) * \neg O_{n+1}$ 相矛盾的归纳结论是不可取的;而当归纳结论 G_{n+1} 与 $I(B_B) * \neg O_{n+1}$ 矛盾时,又因为前者才是根据我们的直觉证据 B_A 得出的结论,而觉得 $I(B_B) * \neg O_{n+1}$ 才是有问题的那一方——尽管从逻辑形式上来说,我们完全无法区分 $I(B_A) * \neg O_{n+1}$ 和 $I(B_B) * \neg O_{n+1}$ 。

5 解决绿蓝悖论的进一步思考

既然绿蓝悖论的产生原因在于逻辑等价的集合,可能会导致互相冲突的归纳结论,那么只要我们能够找出一种方法,可以指示我们什么时候采用 B_A 作为信念集,什么时候又采用 B_B 来作为信念集开始归纳,绿蓝悖论自然就会被消解。

现在,绿蓝悖论的许多解决方案都围绕“绿蓝”这一谓词展开。([1, 6, 8]) 具体来说,就是在一开始就否决掉用 B_B 来描述归纳证据的合法性,规定只能采用 B_A 来描述这些证据。但是这种方法会极大破坏掉归纳法的灵活性,这可以从我们设计的下面这个思想实验看出这一点:

假设某个白天,一位探险家和一位土著遭遇海难漂流到某个荒岛上。他们发现岛上生活着一种蜥蜴。他们注意到眼前的所有蜥蜴都是绿色的,于是夜幕刚刚降临时,探险家得出归纳结论:“我们遇到的下一条这种蜥蜴,也会是绿色的”。

然而土著对此抱有不同看法。他注意到这种蜥蜴的形态、习性等等方面都非常类似于他家乡的变色龙。而那种变色龙白天是绿色的,晚上则会变成另外一种颜色。因此土著得出的结论是:“我们遇到的下一条这种蜥蜴,如果是白天,那么它是绿色的;如果不是在白天,那么它就不是绿色的。”

如果仿照“绿蓝”的定义,我们作为土著和探险家之外的第三方给出这样一个定义: x 是变色龙色的,当且仅当 x 是绿色的 $\leftrightarrow x$ 是在白天被观测的。那么我们可以看到,这里土著的归纳结论实际上就是在说:下一条蜥蜴会是变色龙色的。很显然,这里的“变色龙色”是和“绿蓝”属于同一类型的谓词,而探险家使用的“绿色”则是正常的谓词。也就是说,这个例子中的探险家使用 B_A 来描述证据,而土著使用 B_B 来描述证据(尽管土著说的话里,直接使用的谓词是“绿色的”,但是我们可以看到他实际上是在用 B_B 作为自己的信念集进行归纳)。如果我们规定只能使用 B_A 来做归纳,那么在这个例子里我们就不得不采用探险家的归纳结论而否决掉土著的归纳结论。但是比起对变色龙这种蜥蜴一无所知的探险家,我们直觉上显然会更愿意采用土著的归纳结论。

因此,有时候我们会期待归纳法接纳 B_B 推出的结论,而不是 B_A 推出的结论。所以如果直接粗暴地规定归纳法只能用正常的谓词,即用 B_A 来做归纳——这样得到的形式上的归纳法尽管有可能在形式上消除绿蓝悖论,但是这也会导致我们某些直觉中的归纳论证无法被接纳。

继续考虑上面的例子。我们之所以希望接纳土著的结论,是因为我们作为旁观者,此前已经知道存在变色龙这种会改变自身颜色的物种,而土著又说这种新的蜥蜴长得很像变色龙,所以我们才会认为根据“观测到的蜥蜴都是变色龙色的”这种命题做归纳是合理的。但是如果我们根本没见过、也没有听说过会变色的动物,就像那个探险家一样呢?那么很明显,我们会认为根据“观测到的蜥蜴都是

变色龙色的”这种命题做归纳是完全不合理的。

所以当我们在直觉中进行归纳时，究竟根据什么标准来选择 B_A 还是 B_B 来做归纳呢？

在某种程度上，我们的直觉遵循了简单性标准，即人们在建构信念集时会：尽量采用简单的、已经被大量使用过的谓词，并尽可能采用简单的命题结构。但是这其中有一个前提，就是这样归纳得出结论应该符合我们的预期——或者说，当这样的信念集得来的归纳结论符合我们的预期时，我们才会在遵循简单性标准的前提下采纳这样的信念集。

在先前的例子中，探险家预期下一条蜥蜴仍然是绿色的，所以他才会以“第一条蜥蜴是绿色的”“第二条蜥蜴是绿色的”等等，如此来构建自己的信念集；而土著预期下一条蜥蜴不是绿色的（因为下一条蜥蜴不是在白天被观测的），所以他以实质上等同于“第一条蜥蜴是变色龙色的”这样的命题来构建自己的信念集。而我们对于下一条蜥蜴的预期判断，和土著的一样，所以此时我们才会选择 B_B 作为信念集（因为 B_B 的结论与预期不冲突，而 B_A 的结论与预期相冲突）。但如果我们不知道有关变色龙的知识，并因此预期下一条蜥蜴是绿色时，我们就会以 B_A 作为信念集开始归纳，就像探险家一样。

综上，人们在建构信念集时会采用符合我们预期的信念集。而之所以说在此基础上按简单性标准挑选谓词，则是出于常识考虑。例如在说到苹果颜色时，我们会说：“苹果在秋天是红色的，在其它时候则不是红色的”，而不会造出一个苹果色的概念： x 是苹果色的，当且仅当 x 是红色的 $\leftrightarrow x$ 是在秋天被观测的。然后说：“苹果是苹果色的”。在描述事物时，我们往往会使用熟悉的、简单的谓词，尽管这有可能会使得语句变得冗长，也不愿意用一个陌生的新谓词来让语句变短些。

因此，我们对于信念集的建构是基于预期的，并在此基础上依据简单性标准建构信念集。而我们对于未来的预期究竟是如何产生的，过往的经验和知识积累无疑是预期产生的支持证据之一，但是更加深入的探讨或许是下一个值得研究的问题。

参考文献

- [1] W. Freitag, 2015, “I bet you’ll solve Goodman’s riddle”, *The Philosophical Quarterly*, **65(259)**: 254–267.
- [2] W. Freitag, 2016, “The disjunctive riddle and the grue-paradox”, *Dialectica*, **70(2)**: 185–200.
- [3] N. Goodman, 1983, *Fact, Aiction, and Forecast*, Cambridge: Harvard University Press.
- [4] S. O. Hansson, 1999, *A Textbook of Belief Dynamics: Theory Change and Database Updating*, Dordrecht: Kluwer.

- [5] R. Israel, 2004, “Two interpretations of ‘grue’—or how to misunderstand the new riddle of induction”, *Analysis*, **64(284)**: 335–339.
- [6] A. Schramm, 2014, “Evidence, hypothesis, and grue”, *Erkenntnis*, **79(3)**: 571–591.
- [7] A. Zinke, 2018, “Against grue mysteries”, *Erkenntnis*, **85(4)**: 1023–1033.
- [8] 顿新国, 归纳悖论研究, 2012 年, 北京: 人民出版社。
- [9] 熊立文, “信念修正的 AGM 理论”, 现代哲学, 2005 年第 1 期, 第 127–131 页。

(责任编辑: 袁之)

The Grue Paradox in the Perspective of Belief Revision Theory

Xiaoliu Du

Abstract

AGM belief revision theory is a theory that portrays the state of an agent's beliefs and their changes. With appropriate modifications and expansions, it can also be used to describe the beliefs of agents when they engage in enumerative induction activities. The shifting beliefs of the agent in the grue paradox can also be portrayed by this modified AGM theory. Although this portrayal does not directly eliminate the grue paradox as scholars had initially expected, it does provide a new perspective on the problem and presents a clear picture of the root causes of the grue paradox.