

人们对亚里士多德的“四因说”很熟悉，它是科学解释理论的源泉。随着科学的发展，亚氏学说中的诸多目的因解释被抛弃，现代科学采纳的是更简洁的动力因的形而上学基础。但是，目的因错在哪儿，它一点儿作用都没有了吗？而且在自然科学中，动力因果

解释并不够。

应当如何更好地理解并理顺这些关系，中国科学院哲学研究所学术所长、复旦大学特聘教授刘闯和青年学者、复旦大学哲学学院博士研究生苏无忌对此作了严谨的分析和论证。

在科学解释中，为何不能抛弃目的因

■刘闯 苏无忌

在日常生活中，目的常常是用来解释现象的原因，比如“运动员努力训练是为了在比赛中获胜”“竖立标志牌是为了提醒行人”。但在科学解释中，目的似乎完全不在解释之列。科学解释中的因果解释总是在寻找发生在结果之前的原因，这在哲学上被称为动力因。

比如苹果落到了地上，科学的解释是它受到了重力的作用，不会是它有回到大地的目的。在生物演化过程的科普视频或者文章之中，即使作者在作出类似“为了能吃到蚂蚁，食蚁兽进化出了长舌头”的表述之后，往往也要补上一句“这只是一形象化的说法，演化本没有目标”云云。似乎目的因与科学规律是完全绝缘的。

这与自然科学呈现给我们的世界图景相关。动力因的形而上学基础可以这样理解，宇宙间的事物随着时间的推移不断变化，之前发现的变化总是会导致之后的新变化，因此，对于某变化引起的事件，便可以问在其之前发生了哪个事件导致了它。

1 目的因果被排除出科学的图景，隐藏着更大的问题

从科学史的角度看，亚里士多德的“四因说”是科学解释理论的源泉。但他的因果概念比现代的宽泛许多。在他那里，科学解释即为因果解释，因果解释也是语意和逻辑理论的重要组成部分。概括地说，所谓原因，就是“为什么某某会是这样的”这类问题的回答。

亚里士多德的“四因”有：质料因、形式因、动力因和目的因。以一尊铜像为例，它为什么坚硬又有金属光泽？那是“因为”它是由青铜铸造而成的，青铜即为其质料因。为什么铜像看上去像人？是因为它具有雕塑艺术的某种形式。这就是它的形式因。该铜像为什么会出现在那儿？那是因为它在某工匠或艺术家的作品。这是动力因。该铜像的目的因可以说是为了表彰人物。

“四因”可分为两组，质料因和形式因一组，动力因和目的因一组。根据德国哲学家莱布尼茨的形而上学原理，前一组的因果必定是同一的，属于同一个体；而后一组的因果则可以属于不同的事或物。青铜像的“青铜”与“像”必定属于同一个东西，而创作它的艺术家与形象所表彰的人物都与青铜像不是同一个东西。后一组因果的概念更具有“一物生一物”的含义。

也许就是这个原因，四因说中只有后一组因果概念为后人继承下来。经过近代科学对自然现象的“祛魅”，目的因逐渐从自然科学的解释模式中消失，动力因被近代科学改造成为解释自然现象的唯一合法路径，但目的因在人文科学领域仍然保持着它应有的地位。解释理性人类的行为似乎不可能彻底排除掉引行为的动机、意向和目的。

哲学家休谟和康德都对日常和科学的因果概念做了独到的分析。在休谟看来，即使是解释人的行为，因果解释的模式仍然是动力因：某人前一时刻的动机导致他后一时刻的行为，因此给出该行为的动力因果解释，只不过在“动机”中包含了诸如行动的目的等意向性的东西。

举个例子，说张三打了李四“是为了”惩罚他是目的因果解释。首先必须意识到，这个解释与亚里士多德说重物自然落向地面“是为了”回到地球中心的解释完全不同。张三作为有意识和意向的人，他行为的目的可以在他的意识中转化成他行动的意向，作为动机驱动他的行为。换句话说，“张三打李四是为了惩罚他”可以无遗漏地转述为“张三因为想惩罚李四而打了他”。后一句便是动力因果解释的例子。亚氏的那个解释则没有这样转述的可能，至少在当代科学的语境中不能说“重物因为‘想’回到地球中心而落向地面”。

目的因的概念比较复杂一些，它依赖于形而上学框架更加宏大。比如在亚里士多德那里，宇宙中必须存在有目的性结构的系统，整个宇宙就是这样一个系统。像水的目的就是向下流入大海，空气的目的就是升上天空。

然而，随着科学的发展，亚里士多德的宇宙观已经被摒弃，现代科学采纳的是更简洁的动力因的形而上学基础。

虽然目的因解释的地位很尴尬，但人们始终没能从哲学上弄清楚它为什么被抛弃，是因为其解释方法是错的，还是因为其提供的解释内容是错的？为什么恰恰相反，人类行为的目的因解释很容易被人们广泛地接纳呢？那些目的因解释成立的条件，为何在解释自然现象的时候不适用了呢？动物的认知行为是一种有目的的行为吗？可以或应该用目的因来解释认知行为吗？

这一系列的问题是本文讨论的起点。

而康德为目的因做了有力的辩护。在他看来，生命体与无生命体之区别在于体内各部分的联系是否具有“目的性”，以及个体是否为自身的原因（同类繁衍）。而无生命体既不能产生自身，也不具有统摄整体的目的性，因此只可能有动力因果解释。

康德的目的因果观在今天看来有一点值得借鉴。自然系统为繁衍（即自身因果）而做的付出，成为系统各部分相互合作的目标，因此适应度（fitness）便可以被用来解释器官及其功能的存在。正如英国经济学家、博弈论四君子之一宾默尔所引吭鸣唱，苦苦追寻其动力因是没有希望的。即便你知道了所有鸣禽在早春时节体内的分子排列、化学反应、环境影响……你也不会离回答该问题更近一寸。可是该现象的目的因果解释却非常简单，鸣禽为了在繁殖季节保护自己的地盘，用长期演化出来的“歌喉”向其他鸣禽“宣布”它的拥有权。



鸣禽。

从以上对目的因果的分析，可以得出以下结论：对于自然界中有意识意向的系统，解释其有目的的行为并不需要沿用传统目的因果的解释模式，因为意向的意向蕴含着行为的目的性，而意向就如同欲望，则是心理解释中常用的动力因果解释要素。

相反，不具有意识意向的系统，其目的性行为的解释却不可能用这样的手段来转换，不能直接说某生物体具有增强其繁殖机会的意向，因此通过一代代长期的努力成功地拥有了大脑、血液等。

如此看来，从现代科学的视角，目的因果的根源不就是“拟人化”的结果吗？在远古时代，人们普遍相信“万物有灵”，因此把万事万物都当作人一样去解释它们的行动，从而产生各种目的因果的解释。随着科学进步，科学家发现，它们并不是像人这样的，因此目的因果就被排除了科学的图景。如此一来，一切都能解释得通。

但是，这背后隐藏着更大的问题，人类行动的意向性或目的性是哪儿来的？人的行动为什么会有目的，为什么可以用意向或理由来解释行动？难道不是自然之中真的存在有目的因果的存在，而人类的行动遵循了它呢？

实际上，我们看到，自然科学之中真有这样的概念。

熵是热力学中的重要概念。不严格地说，它象征了系统的混乱程度。孤立系统的熵只增不减，直到达到熵最大的热平衡状态。爆炸将高楼化为废墟、高楼因年久失修而倒塌、人终究要死亡等都是热力学熵增原理的必然结果。

整个宇宙可被看作一个巨大的孤立系统，最终它也只能达到热平衡，也就是奥地利物理学家玻尔兹曼所说的“热寂”，宇宙的各部分都具有同样的温度，不复存在宏观的运动或变化。熵增过程是宏观系统无时无刻不在经历的过程，而其终点即系统的最大熵值或热平衡态。

那么可以说热平衡态是所有宏观系统运动变化的“目的”吗？人生的目的是死亡吗？宇宙的目的是热寂吗？这样说似乎非常奇怪，但热现象的这种目的性，实际上是其他种类的目的因果解释的背景。生物系统的目的几乎都是为了“反抗”热力学目的而存在和被认识的。这就是为什么说，如果没有死亡，为了生存而奋斗就不能被称为生物世界的目的了。

耗散、死亡或者热寂这样的“目的”似乎没有被科学家视为解释自然现象的目的因果解释机制。我们在热力学中找不到明显的相关例子。而其中一个很重要的原因是18世纪关于热的本质的争论以“运动说”战胜“热质说”告终——英国物理学家朗福德用实验证明了热量与物体内微观部分无序运动的关系。从此耗散和死亡等现象均得到了满意的“动力因果”解释。系统的熵之所以不减，是因为系统内微小粒子不停地、无序地运动导致的，而不是冷热不同的东西接触后会趋于一致，是“因为”热平衡是它们的共同目的。

上文说到，康德是近代哲学中维护目的因果解释的第一人。但他并不是以“回到亚里士多德”这样的口号来捍卫的，他深知牛顿力学的胜利对目的因果解释的毁灭性打击。因此，康德认为目的因果解释作为科学的解释进路，只有在生命现象中才有一席之地。

康德对生命现象的目的因果论把系统的繁衍作为必要条件之一，另一个条件是系统各部分相互

3 生命体的各种演化现象也遵循极值原理吗

熟悉美籍华裔科幻作家特德·姜《你一生的故事》（其改编成电影《降临》）的读者应该记得，其中的外星人七支桶用费马定理作为它们物理学最基本的原理。费马定理说，光实际走过的路径一定是所有连接起始点和终点之间路径最短的。从费马定理中可以推导出反射、折射、透镜成像等几何光学的所有定理。

第一次读到此处的读者一般都会震惊，因为这个表述似乎光线一开始就知道它最后要去往哪里一样。如同使用费马定理构建科学理论也并非仅仅是科幻作家的构想。一切都可以归结到极值原理（extremum principles）的各种不同形式与作用。在物理学中，几乎一切理论都可以从“最小作用量原理”中推出：与牛顿力学等价的哈密顿力学中，物体（物理系统）的作用量（action）是与物体的动能与势能有关的拉格朗日函数在相空间上的积分。当它取最小值时，可以得到与牛顿方程等价的欧拉-拉格朗日方程。经典力学的运动方程是以动力因果来解释力学现象的基本自然规律，而最小作用量原理“解释”了物体为什么会遵循这个方程而不是其他方程运动。经典力学、场论、狭义和广义相对论、量子力学，以及经典或量子流体力学都可以从最小作用量原理中推出。

那么是否可以整个物理学都可以纳入最小作用量原理呢？情况并非如此。最小作用量原理适用的系统，在物理学中叫作保守系统（conservative systems），也是能量守恒的系统，或者是系统内没有“耗散”现象发生。那么耗散现象遵循任何极值原理吗？上文提到的热力学第二定律，或熵增定律，也即耗散现象的极值定律。孤立系统的任何变化，要么是熵不变（可逆过程），要么是熵增过程（不可逆过程）都可以算极值原理。

上一节已经谈到演化论中的目的因果解释，那么生命体的各种演化现象也遵循极值原理吗？从最新的理论——自由能原理的角

2 大自然中的运动完全无目的性可言吗

合作的目的是“为了”维持系统的生存，而此目的反过来为部分为何存在提供了解释。问题来了，虽然生存和繁衍的确可以科学地当作生命体的终极目的，但为此目的生命体内各部分为什么必须具备它们特定的形态与功能，部分之间又为什么必须如此这般相互联系呢？比如热系统变化的目的是达到热平衡，而系统内各部分必须具有交换热量以求均温的倾向或功能，这样的目的因果解释似乎简单清晰、没有问题。

可是生命体的目的因果解释却没有这么简单。首先，什么是生命体的生存状态（或有繁衍能力）？系统需要具备什么条件才能生存繁衍？再有就是，怎么知道生命体为了生存繁衍的目的必须有器官与器官间的合作？不能回答这样的问题，康德式的目的因果解释就无法自圆其说。当然，我们现在知道，达尔文的演化论可以为康德的目的因果论提供自然主义的诠释。

达尔文演化论的原始隐喻是家畜育种，《物种起源》便是以这个隐喻开篇的。大自然像一位高明的育种师，利用环境中不同物种的生存竞争来自然选择最适应环境的物种（准确地说是种群内的竞争）。因为每个物种世代同类繁衍，生存竞争与自然选择在时间的长河中造就了地球上的物种。如果得冠的赛马是人工选种培育的结果，那么野马或者任何其他动物植物便是自然选择培育的结果。育种师有目的和方法，比如培育赛马的目的是让它跑得快，大自然同样有目的和方法，但自然选择的目的只有一个，选择最适应环境的物种。

演化博弈论是成功引进了经济学的博弈思想、用“适应度”替换“效用”而得到的理论。它拓展了达尔文演化论，把仅能考虑物种与环境的契合的演化论，变成了不但能考虑于环境还能考虑于种群中其他特性或策略的博弈竞争理论。这里的隐喻就不再是家畜育种员，而是有理想和目的的博弈者。显然，演化博弈论现象的目的因果解释中的“原因”都是“为了最大化适应度”。

度来看，极值原理仍然是比生命现象更为基本的自组织系统行为的基本原理。

以英国理论神经科学家卡尔·弗里斯顿为代表的神经心智科学家新近发明的理论表明，大自然中的自组织系统之所以能够抗拒耗散、维持自身远离热平衡态，就是因为这种系统遵循最小自由能原理。这个原理在理念层面上与上文讨论的最小作用量原理一脉相承。

演化现象是“复制子”生存繁衍的现象，复制子可以是基因，也可以是个体或群体，甚至可以是文化单元，如模因。当代演化论的“运动方程”更像是分子动力学方程，方程的解中有所谓的演化稳定策略状态。具备某些特征或策略的物种一旦达到这个状态，既可以保持稳定且不被其他特征或策略侵入。这个特殊状态的原理是经济学中的纳什平衡态，而它们都有热平衡态的影子。

没有复制子的多元化机制（如基因突变）和自然选择就没有演化现象，而这一切都离不开更为原始的“自组织系统”的存在。根据以弗里斯顿为代表的能动推理学派的理论，维持自组织系统的条件是马尔可夫链，原理是“最小自由能原理”。

自组织系统的存在和发展既需要与环境隔离以防耗散，又需要准确“认识”环境以便在生存竞争中取胜。前者为主要需求，由马尔可夫链满足，后者是在前者的前提下演化出来的“对策”机制，而系统的自由能最小化则是系统对策略机制演化的约束原理。

可以从理论上推导出，在满足生物系统的其他生存条件的情况下，系统在马尔可夫链内能够最小化自由能的认知系统是一套遵循贝叶斯推理原理、最小化预测误差的多层次机制。从这个意义上说，动物心智不但本身是一部遵循极值原理的机器，而且它的诞生也是遵循极值原理，通过演化博弈的漫长过程演化而来的。



生存竞争。



电影《降临》剧照。

作者供图

那么要问，为什么生物学和热力学不同，目的因果解释似乎仍然占据了一席之地呢？当然，生物学中有很多现象都是用动力因果解释的，比如说肢体创伤的发生和治疗。只有演化论这片领域，目的因果仍有地位。那么要再问，这种目的因果解释最终可能被动力因果解释取代吗？有可能把血液存在是为所属肌体供氧的目的转变为动力因吗？又有可能说，血液器官“知道”自己的功能，因此为此功能而存在和行动，就像消防队员知道自己对社会的功能，因此履行职责一样吗？前者似乎不行，而后者显然不行。

现在的问题是首先要解释为什么大自然中存在以供氧为功能的血液。这样的问题正好是演化博弈论可以圆满回答的。即便我们找到了一份存在过的血液生成的动力因果解释——哪一组分子发生了怎样的生物化学反应后产生的血液，我们还是没有回答“血液这样的器官为什么会存在”这个问题。

回到亚里士多德解释重物下落的例子，我们可以看出，“重物的目的是回到地心”与“血液的目的是为所属肌体供氧”，其实不是同样类型的目的因果命题。因为把地球中心的目的因果解释换为万有引力的动力因果解释，重物下落的目的性便被取缔了。但是把血液行为的解释换成动力因果解释，血液服务肌体的目的依然存在。而这一点对于生物现象中同类的目的因果解释均成立。这也是康德的目的因果论辩护的重点，那就是生物体的各个部分为了维持整体的生存而各自具有合作功能。

上文已经说过，达尔文演化论弥补了康德的目的因果论的局限，为部分与整体之间的目的论关系提供了理论基础。同样的结论也适用于鸣禽为何在早春引吭的目的因果解释：即找到这种行为的具体动力因果细节，鸣禽早春引吭的目的——相互告诫不要侵犯对方的支配领土仍然存在，并且用它来解释行为似乎仍是合理的解释选择。

总结来说，大自然中原子、分子在离散状态下的运动似乎完全无目的性可言。对这类运动的解释似乎只能用动力因果解释。但即便有目的，用它而不是用动力因果解释似乎还是有问题。

热平衡是所有（孤立）热系统变化的终极目的，但用它来解释热系统的行为似乎不合适，部分原因是，在许多情况下，热力学需要解释的恰恰是系统如何由于某种原因，即动力因克服了熵增的趋势。

然而，在生物学领域中，康德坚信存在着自然的目的因果，目的是解释生物体生存繁衍现象的合理模式。达尔文的演化论及其后续发展似乎为康德的目的论提供了应有的补充。

从上述论证可以看到，在自然科学之中，虽然都是统一用动力因果规律解释的自然，但是有了自然规律统一的目的因果解释还不够，还可以再问为什么。比如，为什么重物运动的目的因果是以引力定律统一决定的？最小作用量原理的确最终回答了这个问题。因为它不仅告诉我们物体必须怎样运动，这是一阶自然规律的范畴，它还告诉我们为什么它们必须这样运动，这就是二阶或元原理的作用。重物自然下落而不是往上飞是因为地球中心是被地球弯曲了的空间的最低谷（爱因斯坦引力场论的最小作用量原理如是说）。难道说这不是“目的因果解释”的典范吗？

由此，我们可以知道目的因果解释绝不仅仅在生物现象领域中才有立足之地，最基本的力学现象也需要目的因果的解释。弄清了自然规律与最小作用量原理之间的关系，也就弄清了动力因果与目的因果之间的关系，我们看到目的因果是统一或奠基动力因果的元层次因果律。

历史上诸多目的因果解释，如上面提到的亚氏的解释被正当抛弃，不是因为它们是目的因果解释，而是因为它们是错误的解释。这个情况与历史上不少动力因果解释之被抛弃是同样的道理。正如上面说到的，时空弯曲解释重物下落是亚氏目的因果解释重物下落的当代版本，两者都是目的因果解释的范例，不过前者对后者错而已。

目的因果解释的成功与否往往取决于解释现象的形而上学理论框架是否成功。但即便是相信实证主义，极值原理的功能仍然可以理解。自然规律不过是观察实验数据的简单归类，而它们之所以各自具有不同的形式但都是自然规律，则是因为它们都具有最简单省力的形式，这可以说是对极值原理的实证主义诠释。

同样，依靠极值原理，我们也可以知道生物感觉和器官的目的性从何而来。为何哺乳动物的眼球可以转动？简单的回答可以是，它是为了观察时眼睛有一定的“行动”能力来帮助消除预测误差。如果眼球不能动，生物体往往用复眼的结构来达到同样目的。