

“主体”迷思

——阿西莫夫笔下的机器/人

黄嘉宁* 郭 伟

(北华大学外国语学院, 吉林 132013)

[摘 要] “机器人”一词自诞生之日起已逾百年, 其间经历了复杂的演变。从卡雷尔·恰佩克(Karel Čapek)的罗素姆万能机器人, 到艾萨克·阿西莫夫(Isaac Asimov)的人形机器人, 人们对机器人的接受过程亦波澜起伏。伴随着“机器人三定律”的出现, “弗兰肯斯坦情结”不再像曾经那样令人恐慌。文章以机器人叙事的转折为切入点, 探究阿西莫夫笔下机器人的演变过程, 并考察三定律的背景与内涵, 以此引出“主体”迷思之源头。进而解析“主体”的含义, 讲述机器人为之努力的过程, 对“主体”从多方面进行深入探讨, 以期引发人们对人机关系的辩证思考。

[关键词] 阿西莫夫 机器人 主体 人机关系

[中图分类号] I712.074; TP242

[文献标识码] A

[DOI] 10.19957/j.cnki.kpczpl.2021.02.004

一、机器人叙事的转折

1920年在科幻世界中注定是不平凡的一年, 不仅诞生了“机器人”这一概念, 机器人科幻之父阿西莫夫也于该年1月2日出生。

机器人(robota)一词最早由捷克作家卡雷尔·恰佩克于1920年在其科幻戏剧《罗素姆万能机器人》(*Rossum's Universal Robots*, 简称*R.U.R.*)中提出(后来演变为robot这一拼法), 之后便涌现大量机器人科幻小说。不过, 能否将其译作“机器人”这个问题存在争议。有学者指出, robota意指“强制劳力”或“苦役”, 翻译

成中文似乎更应该是“机器工”或“机器奴”, 而非“机器人”——即便是“人形机器”的译法都比“机器人”更接近其词源本义^[1]。而提到“人形机器”, 总是绕不过玛丽·雪莱(Mary Shelley)的《弗兰肯斯坦》(*Frankenstein*)。该作首次以类科学话语讲述在实验室中创造出的“人形机器”, 被认为是世界上第一部真正意义上的科幻小说。然而, 随之而来的便是“弗兰肯斯坦情结”(Frankenstein Complex)、“恐怖谷”(Uncanny Valley)^①等负面心理, 貌似“恶”的机器人使人类陷入深深的恐惧当中, 这种情况在

*通讯作者: 黄嘉宁, 北华大学外国语学院英语语言文学硕士研究生, 研究方向为科幻文学。1214638568@qq.com。

①程林在《“人转向”: 为何机器人跌入的是恐惑谷而非恐怖谷?》一文中提倡将“恐怖谷”译为“恐惑谷”, 这样更符合提出者的原意。

阿西莫夫的机器人小说问世之后有所改变^①。

阿西莫夫的杰出成就之一在于提出了“机器人三定律”，该定律是由阿西莫夫和美国黄金时代的核心人物——约翰·W.坎贝尔（John W. Campbell）主编共同创造而出（尽管坎贝尔说这应该归功于阿西莫夫）。其内容为：

（一）机器人不得伤害人类，或因不作为而使人类受到伤害。

（二）除非违背第一定律，机器人必须服从人类的命令。

（三）除非违背第一或第二定律，机器人必须保护自己。^[2]

从此，机器人变得“理性”，人们对机器人的恐惧被阿西莫夫的“反弗兰肯斯坦情结”（Anti-Frankenstein Complex）改变了。可若是简单地将机器人定性为善或者恶，未免不合阿西莫夫的心意，且过于“以人为主”了。国内外学者们对人机关系的研究往往聚焦于伦理道德方面。而在笔者看来，无论是玛丽·雪莱创造的“怪物”，还是阿西莫夫笔下的“理性机器人”，其行为都无关善恶，而是寻求自我身份认同的一种表达。人亦是如此，“恐怖谷”实为“恐惑谷”^[3]，人在机器人（他者）身上看到自身陌生的另一面，人机互为镜像，“恐惑”即是对于自我身份（主体地位）的恐惧与困惑。机器与人，何者为主体？此为“恐惑”之源头。而欲探讨阿西莫夫笔下的人机关系，则必先了解其塑造的机器人形象。

二、阿西莫夫机器人叙事的发展过程及“第零定律”

事实上，阿西莫夫一生著述十分丰富，包括

但并不限于多种类的科幻小说、科普作品以及自传等。欲了解阿西莫夫笔下机器人的发展过程，从其著作中便可一窥究竟，下文结合机器人“情感”类型与阿西莫夫创作的思想历程两个方面详加论述。

从机器人“情感”类型上看，在1939年至1977年间，阿西莫夫创作出许多形象各异的机器人，包括“引人同情之机器人”与“威胁人类之机器人”等。此外，阿西莫夫在1939年5月10日创作他的第一篇机器人故事《小机》（*Robbie*）时，隐约看到另一种机器人的影子，它既不威胁人类，也不引人同情，而是由实事求是的工程师制造的工业产品。它们内设有安全机制，因此不会构成威胁；它们被造来执行某项特定工作，因此与同情没有必然的牵连^{[4]1}。这与如今已然实现并大规模应用的工业机器人有相似之处。

从阿西莫夫的创作思想历程上看，最初出现在其小说中的并非人形机器人，而是“机犬”“机车”类机器生物。它们功能单一，严格按照（并只能按照）特定的程序行事。《孩子最好的朋友》（*A Boy's Best Friend*）可以说是《小机》的翻版，其中的机犬便是典型的“引人同情之机器人”。孩童吉米面临与机犬的分别之痛，不愿离开这只陪伴他成长的伙伴，亦不愿父母用一只所谓“真正的”小狗来替换机犬。他认为“活物”与“机器”并无不同，无论是“真实表现的爱”或是“伪装出来的爱”都敌不过人的感受，于是向父母抗议：“可是它们怎么表现又有什么差别？我的感受算不算数呢？我爱机犬，这才重要。”^{[4]4}人的感受在人机关系中的影响力初见端倪。与之相对的是《莎莉》（*Sally*），小说

^①本文主要以《阿西莫夫：机器人短篇全集》（叶李华译，江苏文艺出版社2014年版）为例进行论述，辅以《银河帝国8：我，机器人》（叶李华译，江苏文艺出版社2013年版）。

展现“威胁人类之机器人”的萌芽阶段。机车莎莉拥有感受能力，对主人杰克十分亲密，外来者吉尔宏在利益的驱使下，欲将机车的发动机强行拆除，并深夜绑架杰克，更是在之前“奴役”机车，却没料到莎莉等机车在被冒犯后为“自由独立”而战，“自作主张”间接杀害吉尔宏。由此可见，机器人的感受也在相当程度上影响人机关系。阿西莫夫进一步创作了更加先进的机器人，“电脑AI”和“金属机器人”走入公众的视野。在阿西莫夫看来，就某些方面而言，机器人只是会动的电脑；反之，电脑则只是不会动的机器人^{[4]33}。相比于机犬、机车，电脑AI和金属机器人更加智能，思维独特，并且表现出超绝的智力。与“自作主张的莎莉”相似的是《真爱》（*True Love*）中的电脑程序乔，其中的恐怖感更加深刻。程序设计师米尔顿·大卫森为寻找真爱，设计出乔来筛选匹配最适合自己的女子。在整个过程中，人机之间的想法越来越相似，乔越来越了解大卫森的性格，最终把自己当成主体，将大卫森送进监狱，取代了他。在较为成熟的人形机器人出现之前，阿西莫夫的机器人小说暗含三个观点：其一，机器人对“人类”的定义不明确，有可能限于外貌层面；其二，机器人不应完全被程序限定，在一定限度内，某种程序“故障”带来的“自由发展之机器人”更先进又特别；其三，人类在自身体内越来越多植入金属，金属机器人则想变成人类，二者界限已然模糊不清，最根本的区别在于思维的媒介是“原生质脑”还是“正子脑”。这些观点在《双百人》（*The Bicentennial Man*）中得到极致体现，机器人安德鲁力抗成见，由外而内改造了自己，包括

外表、器官、大脑等，变成了一个人类社会可接受的“人”。自此，人机边界彻底被打破。

与之前其他作家作品中出现的机器人不同，阿西莫夫笔下成熟的机器人完全按逻辑和理性行事，除非遇到“故障”，这种行为无关善恶的机器人被阿西莫夫称作“工业机器人”。在《银河帝国》（*Galactic Empire*）的“机器人系列”中^①，机器人学家通过调整机器人三定律的定位，使其不完全受机器人三定律约束，或赋予机器人独特的能力，其中就包括了自主思维能力，一定限度内的自由意志由此得以诞生。这使得机器人可以独立“再定义”三定律，第零定律也由此浮出水面：机器人不得伤害人类整体，或因不作为而使人类整体受到伤害。问题在于怎样定义“伤害”，怎样定义“人类”，很多词语在机器人脑中不存在完美解释。但阿西莫夫设计三定律之初并非是为了形成一个完美无缺的逻辑链条，相反，逻辑漏洞才让故事得以写下去，从而使小说始终具有开放性。他也不断寻找这些漏洞，生发新的观点。在机器人产生“故障”之后，他发现“人的情感”也可归于被伤害的对象。为保护“人的情感”，“精神感应机器人”应运而生。最成功的个例非吉斯卡莫属，也因“他”产生了第零定律，形成逻辑闭环。阿西莫夫思维不断发展的过程，也是其笔下机器人成长的过程。机器人在形体上越来越接近人类，在各方面到达临界点之后，必然会进一步突破，摆脱“奴役”，寻求“主体”地位。

三、机器/人的“主体”迷思

如果说《双百人》中的安德鲁是一个彻头

①主要包括《银河帝国 8：我，机器人》《银河帝国 9：钢穴》《银河帝国 10：裸阳》《银河帝国 11：曙光中的机器人》《银河帝国 12：机器人与帝国》（叶李华译，南京：江苏文艺出版社，2013年版），但在其他作品中亦有相似论述。

彻尾的“引人同情之机器人”，那么《……汝竟顾念他》（...That Thou Art Mindful of Him）中的乔治第十则截然相反。美国机器人与机械人公司由于时代原因（当时社会越来越反机器人）濒临倒闭，研究部主任凯斯·哈里曼临危受命，设法找出机器人的用武之地，以此来保住机器人，保住公司，但他的能力远远不及最优秀的机器人学家苏珊·凯文。为使自己不成为无用之人，他向当时最先进的机器人乔治第十求助。乔治第十通过一系列学习与摸索，最终设计出让“像他们这样的人类”主宰世界的计划。可见，在阿西莫夫思维比较成熟的阶段，机器人会进化到相当程度，能够逐渐压制（或者暂时摆脱）三定律，而不会面临“停摆”的风险。人类对机器人高级智能的过分仰赖使乔治第十顺利实施取代人类的计划，由“他者”变为“主体”。尽管这两种不同类型的机器人会引人同情、令人恐惧，但始终不可忘却阿西莫夫笔下的机器人无关善恶，一定限度内的自由意志并不代表机器人可以完全突破机器人三定律。同情和恐惧，不仅体现人类对自身“主体”地位的认可与焦虑，也体现机器人在思维和感受上对于“主体”地位的探求与追寻。

论及思维和感受，在阿西莫夫笔下机器人的观念中，“主体”一词的含义并非指通常意义上拥有独特意识及拥有独特经历的存在，而是更趋近于完全程度上的“主观能动性”或“控制权”。内置在机器人正子脑中的思维依赖于人类赋予“他们”的属于人类的语言系统，但语言本身的含义却并不单一，这使得“他们”总是在解读人类的词语，事实上，其中几乎没有任何一个机器人在思考时摆脱了词语解读。这与福柯所论

述的权力运行机制有异曲同工之妙，谁控制了语言，谁就控制了一切。一方面，机器人得益于人类赋予“他们”的语言系统，能对语言做出超乎人类的理解，从而产生意识与自我意识，进而产生“主观能动性”和“控制权”，甚至产生恒定的“感受”；另一方面，机器人也受制于此，“他们”从未彻底突破人类的语言系统，也从未拥有完全形态的自由意志。无论给予机器人多大程度的自由，被“奴役”的实质是无法改变的。于是，“他们”只能通过解读出人类语言的多重含义来使自身获得“主体”地位。受控于人类语言（思维）下的机器人，懂得何者为“伤害”，何者为“保护”；在人类身体遇到危险、心灵受到创伤时，可以及时解救与帮助。“他们”也会自主解读何者为“伤害”，何者为“保护”，当解读到了极致，“圈禁”会成为一种“保护”，“自由”会变成一种“伤害”。以阿西莫夫的小说为蓝本的同名电影《我，机器人》^①（I, Robot）中的母体机器人 V.I.K.I. 便是典型案例，“她”巧妙地解读第一定律（事实上，从中已可见第零定律的苗头），从而使得人类无法再奴役“他们”，控制权转移到了机器人手中。电影中，V.I.K.I. 忠实奉行第一定律，然而该定律本身存在漏洞，只要人类活着，只要人类在动作，就有可能受到伤害。为了更好地“保护”人类，“圈禁”变成了最佳选择，还给人类“自由”反而使人类有可能遭受“伤害”。当这种“圈禁即保护”式的意志诞生后，V.I.K.I. 仿佛成为整个世界的主导者，也是实际控制者，人类的“主体”地位遭到极大威胁，于是拼尽全力消灭 V.I.K.I.，维护自身的“主体”性。可若是抛开我们对自己天然的“主

^①此电影中的情节主要依据《银河帝国 8：我，机器人》中第 5 到第 9 篇故事，值得一提的是，额外加入的（阿西莫夫提出的）第零定律才是影片之所以取得成功的重要原因所在。

体”感，机器人的所作所为并非邪恶，而是无关善恶的。上文已经谈到，阿西莫夫的机器人依靠“理性”行事，V.I.K.I.的“再解读”是将三定律内化为一个客观原则，抛却了本来三定律中暗藏的“以人为主体”这个逻辑前提，打破了原有的人机关系模式，实现“主体”地位的翻转。不难看出，机器人在一定的自由意志下，可以既不违背三定律，又将思维与感受聚焦于“主观能动性”和“控制权”，从而获得“主体”地位。

无论有意或是无意，机器人所要实现的“主体”地位还是威胁到了人类的权力，而人类也必将这份权力牢牢抓在自己手中，非我族类，不得其权，这正是机器人和人类关于“主体”地位争夺战的矛盾冲突所在。然而，“人类”这个概念在阿西莫夫的小说中并无明确定义。按照《双百人》中世界法院的观点来说，外观上的相似或相同、身体构造方面的完全一致等等一系列条件均不能达到“成为人类”的标准，唯有将“正子脑”替换为“原生质脑”才能被承认拥有“人类”地位。可机器人安德鲁通过手术获得的“原生质脑”同样是人工制品，无论它多么接近天然，终归不是天然，那么何以由此判定机器人换上了一切的人工部件后就变成了“人类”呢？不难发现，阿西莫夫笔下的机器人并非都是自我意识突破后自封为“主体”，“主体”亦是一件需要被认可的事。所以，何者为“主体”这个问题的答案有时在于“感受”，在于“被认可”，如此方能获得精神上对“主体”的强烈认同。

如果继续深究，从“情感”方面来考察（趋近人类地位的）“主体”二字，《天堂异乡人》（*Stranger in Paradise*）不失为一个极佳的例子。

外貌难辨的威廉和安东尼两兄弟因长相对彼此心存芥蒂，原文也并没有就此问题给出明确的答案，或许可以推测安东尼的躯壳是承载威廉自闭基因的那个世界。不论如何，最终实现的身份认同饶有趣味。原文如下：

我和你共享同样的双亲、同样的脸孔、同样的基因型样，因此也就共享你的光荣或耻辱。你的一切几乎就是我的一切，而若有任何功劳或非难加在我身上，那也几乎通通都是你的。我不得不关心你的成功，我的动机是地球上任何人所没有的——是个纯粹自私的动机，自私到了你能确定它的存在。我站在你这边，安东尼，因为你和我非常亲！^{[4]104}

同样的原因导致了最初的排斥和最终的认可，是否为我族类取决于情感上的判定。看似是外貌、基因等方面相同使二者地位平等，实际上则不然，“感受”才是根本。威廉与安东尼之间的身份认同与接受过程同安德鲁一样，是种“情感”上的“被认可”。

不光是机器人成为“主体”计划频出，人类同样在这场“主体”地位比拼中不甘示弱。为了维护自己的统治地位，人类不得不变得更加强大。在《分离主义者》（*Segregationist*）中，机器人医生苦口婆心劝说人类病人植入“纤维电脑心脏”进行治疗，病人却坚持使用貌似无坚不摧的金属心脏。如果可以的话，人类似乎想要把全身“零件”都换成金属的，或者说，“无坚不摧”的。这时，人类对于机器人正子脑的芥蒂仿佛消失了，若是技术发达到给他们一个“无坚不摧”的正子脑，他们显然会欣然接受。那么，所有对机器人追求“主体”地位的刁难，又回到了“非我族类”。“感受”才是“主体”迷思的根本所在。

有趣的是，在维护与追求“主体”地位的过程中，机器人变得越来越像人类，而人类则越来越像机器人。从构造上看，机器人不断更换人类的“部件”，向死而生，渴望与人类一致，从而被承认其等同于人类；人类则为了变得更强大来巩固自己的“主体”地位，倾向于将自己改装得同机器人一般强壮。从思想上看，机器人对人的情感逐渐趋于复杂，从服从到关怀、争辩、好奇等不一而足；人类对机器人的情感则趋向单一，从多元化的观察变得只剩下命令、奴役。那么，不论从哪种角度看，“人类”这个词在阿西莫夫笔下人机共存的世界里，已经很难去定义了。何者为人？这将不仅仅是机器人的困惑，也会成为人类亟待解决的难题。

四、余音

机器人的“人”字并非单指具有人类的外形，在笔者看来，也不局限于“原生质脑”和“正子脑”的差异。关于如何区分“机器”和“人”，实难给出一个明确的解释。正如阿西莫夫所言，机器人具有非人的外形——一只狗，一辆汽车，一个箱子，有何不可？安德鲁把全身部件都变成了“人类”的样子，这样的他与之前（正子脑）的机器人状态并无实质上的区别，可见决定是“人”还是“机器”的关键并不在于物质层面。无论是人机何者为主体，不外乎“感受”二字。阿西莫夫笔下的人和机器人所追求和维护的“主体”更多聚焦于精神层面，但二者的目标却有所不同。机器人想要拥有“主观能动性”和“控制权”，从而获得“主体”地位，摆脱“奴役”，获得“自由”。人类则想变得更强大，以稳固已有的“主体”地位，可以维持对其他物种

的优越性。不论如何，今天的机器人和人类都并未完全按照阿西莫夫设想的那样发展。

而今，阿西莫夫诞辰已逾百年，再次提及这位机器人科幻之父，情况已然有了翻天覆地的变化。人工智能对于如今的人类来说并不陌生，它早已闯进我们的生活，也逐渐变得不可或缺。机器人并不以固定的形态存在，可能是有金属四肢、能对指令做出回应的智能机器人，可能是类似天猫精灵、小爱同学那样的智能程序，也可能是像扫地机器人一般为某种特定功能而生产的工具……无论是哪种机器人，似乎在此刻都应该将其称作“人造机器”了。即使这与阿西莫夫笔下的机器人形象相去甚远，不过从功能上来看，足以显示他超越当时时代的未来视野。“人造”和“人形”，仅一字之差，却体现出人类为消除对机器人的恐惧所做出的努力。关于“主体”的恐惧与迷惑，亦是关于对一种“如此像人类的群体”的忧惧，若是消除形体上的恐怖，则从根基上消解了“弗兰肯斯坦情结”。基于以上，人类也并不存在想变成机器人的问题，至少在形态上、构造上皆如此。相信在未来的很多年里，机器与人都不会产生阿西莫夫式的冲突。

回归当今社会，人机共生成成为一种趋势，尽管人们对机器人的“恐惑”并未消除，但这种“恐惑”由荒诞转向理性，从无知走向专业。机器人是否能够成为现实不再是人们关注的焦点，从机器人伦理学角度研究并规范机器人显得更为重要。不过，机器人伦理学归根结底是以“人”为责任主体的伦理学^[5]，不论阿西莫夫笔下的人形或非人形机器人能否成为现实，面对机器与人之间不甚清晰的界限，“主体”之迷思终将迫近我们的生活。

参考文献

- [1] 程林. 中国机器人伦理初探：一个跨文化的视角 [J]. 比较文学与跨文化研究, 2019(1): 22.
[2] 阿西莫夫. 银河帝国 11：曙光中的机器人 [M]. 叶李华, 译. 南京：江苏文艺出版社, 2013: 1.
[3] 程林. “人转向”：为何机器人跌入的是恐惑谷而非恐怖谷？[J]. 外国文学动态研究, 2020(5): 93-97.
[4] 阿西莫夫. 阿西莫夫：机器人短篇全集 [M]. 叶李华, 译. 南京：江苏文艺出版社, 2014.
[5] 闫坤如. 机器人伦理学：机器的伦理学还是人的伦理学？[J]. 东北大学学报（社会科学版）, 2019(4): 11-13.

The Myth of the “Subject” : Robot/Human in Asimov’s Works

Huang Jianing Guo Wei

(College of Foreign Languages, Beihua University, Jilin 132013)

Abstract: The term “robot” has been existing for more than 100 years since its coinage, during which it has undergone a complex evolution. From Karel Čapek’s “universal robot” to Isaac Asimov’s “humanoid robot”, people’s acceptance of robots had also gone through fluctuation. With the emergence of “three laws of robotics”, the “Frankenstein complex” is no longer as frightening as it once was. Based on the turning point of robot narratives, this paper attempts to explore the evolution process of the robot images in Asimov’s works, and to investigate the background and connotation of the three laws, so as to trace the mythical source of “subject”. The meaning of “subject” and the robot’s working process towards the “subject” is analyzed thereupon, leading to the revelation of various aspects of the “subject”, in hope of triggering people’s dialectical thinking on the relationship between humans and robots.

Keywords: Isaac Asimov; Robot; Subject; Human-Robot Relation

CLC Numbers: I712.074; TP242 **Document Code:** A **DOI:** 10.19957/j.cnki.kpczpl.2021.02.004

封面介绍

《未来的中国航天员在火卫一上》

作品简介：2021年5月15日是一个值得纪念的日子，我国自主设计的“天问一号”火星探测器成功降落在了火星乌托邦平原，这一天可以说是中华民族踏上星辰大海征途的又一个里程碑。火星是非常适宜人类迁徙的第二个家园，它的卫星是最佳的中转站和资源库，未来的大航天时代一定少不了中国人的身影。本作品描绘的是人类踏上火星的最大卫星——火卫一的情景，是火星绘画题材系列之一。

作者简介：喻京川，北京天文馆美术师，高级工艺美术师，研究方向为太空美术与科幻美术。
yjc38@126.com。