

“图灵测试”技术叙事的哲学追问^{*}

吕其镁 涂良川

〔摘 要〕如何揭示当代技术规范及原则的重要价值，呈现其社会历史意义，是表征时代精神之哲学建构的重要之事。“图灵测试”作为判定“机器能否思维”的技术叙事，既是人工智能技术的“金规则”，又是对“何谓智能”的哲学追问。“图灵测试”以技术的逻辑与操作的实践肯定了机器实现功能表征思维的存在论意义；其“模仿游戏”的逼近为我们提供了通过己身对象化产物来理解自我的技术路径；其关系认同“功能实现”的原则既呈现了人工智能技术改变世界的现实效应，也以其实践影响着理解存在的理论思维。在人工智能技术高速发展的今天，直面“图灵测试”技术叙事的哲学追问，既是我们真正理解人工智能技术的哲学路径，亦有助于我们反思人工智能技术的社会历史效应及其未来可能。

〔关键词〕图灵测试 模仿游戏 思维 人工智能 关系认同

〔中图分类号〕TP18；N02

“图灵测试”通常被视作判断机器是否具有“智能”的技术标准与人工智能的起源。然而，自诞生之日起，图灵测试就不断受到各种质疑与挑战。其中，既包括阿兰·图灵自己关于计算普适性的说明以及对各种反对意见的辩护，用以维护图灵测试对“机器能否思维”之回答的确定性（参见博登，第56页），也包括诸如塞尔“中文之屋”对于“机械思维是否理解意义”的哲学诘问（参见同上，第93—99页），彭罗斯减弱图灵测试要求的普适性扩展以判定机器思维精神性品质的物理学与哲学分析（参见彭罗斯，第9页），以及珀尔（J. Pearl）通过“迷你图灵测试”简化测试问题以便考察机器思维“因果推理能力”，进而祛除机器思维对经验依赖的技术性努力（参见珀尔、麦肯齐，第15页），等等。尽管如此，人们总体上还是相信，图灵在《计算机与智能》一文中提出的“图灵测试”构成了对“什么是智能定义”的技术转换与机械表达，“几乎也是关于智能存在与否的唯一可操作的标准”（博登，第2页）。图灵测试基于思维即计算的逻辑，给出了决定机器运行“如何表达思维”“以何表达思维”的技术指标。概括地讲，图灵测试以智能机器系统的运行机制、步骤和结果为对象，在与智能生活活动结果的定量比较中给出“机器是否具有智能”的结论。其技术叙事表达为本质演绎的逻辑明晰、能行过程的操控清晰、运行结果的确定分析。这既是对思维机械过程化和逻辑抽

^{*} 本文系国家社会科学基金一般项目“马克思主义哲学视域中的人工智能奇点论研究”（编号21BZX002）的阶段性成果。

象化的技术操作，又是对思维结果的物化与量化。正因为如此，尽管图灵提出的“70%”和“5分钟”标准在技术与操作方面存在诸多争论，但图灵测试提出的技术表达智能的问题却远远超越了技术本身，蕴含着更深层次的哲学追问。从理论特质上讲，图灵测试回避了“机器能否思维”之统计学定义的含混，实现了对“机器能否思维”的确定性描述与判定；从技术操作上讲，图灵测试将机器运算与智能效应统一起来，使“模仿相似度”成为判定准则，而且“提出了让程序通过测试的策略”（珀尔、麦肯齐，第15页），为判定“机器能否思维”提供了真实可行的技术标准；从存在逻辑上讲，图灵测试重述了自亚里士多德以来坚信确定经验之客观实在性的传统，将机器计算实现的成效提升为判定其是否智能的原则，从而消解了功能对主体性的依赖，技术地肯认了“智能即计算”的哲学观念。由此看来，作为人工智能的“金规则”，图灵测试既是人工智能技术操作的核心标准，又引发了对人工智能诸多问题的思考，并追问了智能的本质、智能的表达等关乎“人类的特权”（彭罗斯，第1页）的重要问题。因此，在人工智能技术高速发展的今天，我们直面图灵测试技术叙事的哲学追问就不是简单地探讨图灵测试的技术逻辑问题，而是要真正揭示技术叙事如何完成人通过自身对象化产物理解智能之本质与本性的内在逻辑。人工智能不仅“能够彻底地改变我们的社会、经济和我们居住的星球，也因为它关乎我们自身——我们是谁？我们如何思考、沟通？是什么让我们成为人类？”（扎卡达基斯，第287页）而图灵测试作为判定人工智能的技术叙事，从本质上说就是对这些问题的技术性展开与对象化判定，是从学理性到实践性的路径澄明。因此，当图灵测试以“能行”（指技术逻辑完备）来呈现机器系统对事物判定、分析和预测的可行性、有效性时，其本质就是对智能本身的存在论追问。

一 技术测试的判定与智能本质的揭示

作为针对“机器能否思维”这一问题给出的可操作的判定程序，图灵测试本身既包含着对机器运行效果的判定，又包含着对智能表达方式的设定，还包含着对人机关系的区分。通过确定“机器”和“思维”的词义来回答“机器能否思维”，是人们习惯的理解方式。（参见博登，第56页）但这一方面容易使问题的解决陷入统计学的泥潭，在技术领域重述“多数人的暴政”；另一方面，词义能否技术化成“能行”的操作也是一个问题。图灵测试从设计思维开始就坚持以技术测试判定能否思维的路径，既避免了“以歧义的语言来表达”（同上，第56页）的问题，又形成了过程清晰、标准明确、结果确定的技术判定操作。因而其技术判定原则就成为以技术的方式对智能本质的一种回答。

直观上看，图灵测试针对的是“机器能否思维”的技术问题，同时也是以机器计算及其结果的可描述性、可判定性和可转化性的工程逻辑对智能本质的描述。其基本逻辑是“有智能就能思维，而能思维就是能计算”（高新民、罗岩超，第57页）——能计算就是能操作，能操作就是能判定。若简化掉其测试过程，我们显然可以说“判定”本身就是对智能本质的描述，因为机器制造者能通过图灵测试“宣称这种仪器实际上是在思维，他们也许还宣布它们真正是有智慧的”（彭罗斯，第4页）。虽然通过图灵测试一直是人工智能技术追求且又尚未完全达到的目标，但该测试的判定路径从根本上肯定了基于存储量大、计算能力强、逻辑单元多、数据积累广的机器系统之行为能力的拟人性。从判定逻辑来看，图灵测试重要的不是理解“何谓智能”的问题，而是应该描述智能驱动的思维“能做什么”的问题。这一思维的转换通常被称作对智能本质的计算主义、机器功能主义、行为后果主义等的描述。尽管图灵测试的技术判定逻辑是依据机器系统拟人化的程度和时效性来判定“机器能否思维”，却从根本上设定了机器在算法趋动下对存在本身进行描述与操作的可能性，以及机器逻辑与人类行为在结果上的不可区分性。也就是说，计算不过是其过程，功能不过是其表现，后

果不过是其结果。虽然这些都是判定“机器能否思维”的直接标准，但机器能够思维并非简单的结果同一性，而是以其独有的方式重述了人工智能与人类智能在对世界进行理解、操作和表达上的同构性。因此，图灵测试特别强调的是“结果不能区分哪个是人，哪个是机器”。

基于机器“行为”的拟人性，图灵测试认可以机器计算与人类思维的同构性作为判定机器能否思维的技术标准，从而给出了“智能存在可以独立于生物体”的技术结论。图灵测试延续技术的逻辑，“在人的体力能力和智力能力之间划出一条截然分明的界线”（博登，第58页）。这就一方面确认了，机器及其运行并非思维的生理结构与运作过程之再现，而是以非生物的电力和机械力驱动逻辑器件、电子元件和机械装置来“非人格”地实现思维；另一方面也预示着，图灵测试只关注在智力驱动下的行动及其现实效应，通用图灵机的一般性与能行性是这一效应的普遍性表达。因此，当图灵测试对机器在人类需要花费心力的领域中实现的任务成果进行判定时，显然并非要对机器的运算结构、运算机理和运算逻辑进行描述，而是把机器整体活动能力视作具有“心力”的系统。图灵对于作为测试对象的机器作了三个基本规定：其一，机器具有通用性，适用于每一种工程技术；其二，机器机理对其制造者和使用者而言具有系统封闭性，允许人们不对它的运算方式作出令人满意的描述；其三，机器系统具有独立于生物机体的系统性，“以正常方式出生的人不在机器之列”（同上，第59—60页）。因此，图灵测试的技术判定不是要求所有当前的数字计算机都具有通过测试的可能性，而是要从理论角度追问“是否存在着通过这一测试的数字计算机”。在图灵看来，“用数字计算机作为对机器的规定，就像我们有关‘思维’的判据一样”（同上，第60页）。因此，尽管图灵测试的技术判定直接针对的是计算机，本质上却是要对机器体系（独立于生物体生命活动逻辑）运行的内在驱动力之实质与特性作出判定。

图灵测试不仅肯认了机器系统在不同状态之间映射出的与人类思维的同构性，而且认可了机器系统运行结果的智能意义。这一方面强化了我们前述的“智能独立于生物机体之可能性”观念，另一方面则表明，智能本质可以用系统某一形式特征进行刻画。图灵测试和图灵机虽然以技术叙事的方式强调了“计算的形式的普适性”（派利夏恩，第59页）和计算与思维的等价性，但从图灵测试对测试对象的规定以及对测试结果的判定上看，这种逻辑并非内在主义的推理，而是行为主义的判定。正因为图灵测试兼具内在主义对思维本质的探讨与行为主义对智能的后果主义的判断，图灵测试才形成了回答智能本质问题的同构性逻辑。我们知道，在《计算机与智能》一文中，图灵不仅给出了判定的具体准则，而且想象出了能够通过测试的图灵机。但能够通过图灵测试的机器显然“不是模仿运动（像早期的钟表机械那样），而是模仿（在这个词适用的范围内）某些不可观察的内部过程”（同上，第53页），且必须在运行效果的评价中才能确认其思维与否。因此，图灵测试本质上并不以结构的“同一性”与成效的绝对相同来认定“机器能否思维”，它看重的是实现不可观察内部过程的机器系统整体。或者说，图灵测试的技术判定既不执着于智能的物质基础，也不执迷于智能的运行机理和控制内容，而是看重以机器编辑来实现智能读取事实、调用控制的能力，用等价“机械计算”的图灵机来具体化智能实现的对象映射及其具体物象化的可能性。一旦机器运行的效果体现了数据结构与算法结合，则既表征了智能表达主体，又尊重了客体的实际效用。（参见吕陈君、邹晓辉，第123页）

因此，图灵测试的技术判定，既从技术的逻辑与操作的程序上给出了“机器能否思维”的确定结论，又“表达了一种全新的、具有革命意义的智能观”（高新民、罗岩超，第58页）。智能既依赖一定的物质结构，又具有超越物质结构的一般形式，还具有表达其特性的客观效应。图灵测试的技术叙事表明，虽然在技术的逻辑上我们可以从具身性来研究与实现智能（比如神经网络、深度学习），也可以从非具身性来设计智能（比如符合表征、结构功能），但是理解智能本质的第一前提却是区分

智能与非智能的界限，然后才是猜测智能实现的基本要素，最后则是以智能效应为目标的逻辑分析或定义。因为，对于能够思维的机器而言，一方面，其所用的推理过程“没有必要得到最严格的逻辑学家的首肯”，只需要实现确定的归类（参见博登，第88页）；另一方面，以“占重要地位的那些规定有关逻辑系统规则的实施顺序的命令”（同上，第88页）实现不同对象之间的映射。此即是说，通过图灵测试的机器思维是智能的技术实现而非智能的本质定义和整体还原。也正是在此意义上珀尔明确指出，图灵测试其实可以在判定机器系统是否具有因果推理能力的基础上，判定机器是否具有智能。（参见珀尔、麦肯齐，第16页）这也完全符合当代哲学的一般看法，即人工智能的发展使我们“摆脱了任何关于实际的有机体如何做它们想做之事的问题，而代之以这样的追问：我们依靠哪一种能力簇，可在原则上使得人工物有智能”（布兰顿，第4—5页）。若是依据图灵测试的等价性，则我们显然可以说明哪一种“能力簇”是智能的基础。因为从人类智能本身的形成过程与表达形式来看，递归地调用某些基础性的过程，恰恰是获得智能和表达智能的过程。因此，图灵测试就不仅仅是“机器能否思维”的判定标准，更在其关系性评价中呈现了机器思维与人类思维形式同构性和效应性的同源性，即虽然在人工智能时代，厘清智能的本质定义的确有助于人们把握智能，但这种方式是一种固定化的知识论逻辑，而要真正把握智能本质就必须全面呈现智能活动的逻辑。

二 本质主义的定义与“模仿游戏”的逼近

在图灵测试中，“模仿游戏”对思维活动的逼近给出了“机器能否思维”的技术判定规则和操作程序，而非基于本质主义，以结构、功能的完全一致来确定机器能否像人一样思考。假如有机器能够通过图灵测试，那就意味着思维不为碳基生命体所独有，其他自治系统同样可以具有。（参见涂良川，第79页）尽管人造自治系统特别是人工智能系统首先是通过植入定义好的逻辑运算、神经网络、深度学习等机制来构建的，但这并不意味着本质主义的定义就从根本上决定了“机器能否思维”。因为本质主义的定义既面临着我们是否理解思维机制的物理难题，又遭受着思维运行环境之社会难题的诘问，还必须回答思维与意识的经验“附随性”问题。（参见查默斯，第8页）因此，“这个问题的新形式可通过一个游戏来描述”的判定就不失为一种可以逼近思维活动过程、呈现思维活动结果进而判定“机器能否思维”的新方式。

作为一种思想建构的技术判定原则，图灵测试既使“模仿游戏”具有技术操作的明晰性、程序性和可判定性，又赋予了“模仿游戏”能够逼近智能的本体论意义。其一，“模仿游戏”以技术的一致性消弥了机器与人的区别（参见博登，第59页）；其二，“模仿游戏”本身的重点不在于模仿人的行为逻辑和行为过程，而在于“设法提供那些人类会自然而然得出的答案”（同上，第59页）；其三，“只允许数字计算机参与我们的游戏”（同上，第60页），这就回避了思维过程的具体性，确保得出的答案能够真实地表征思维本身；其四，不指定“模仿游戏”的具体主体，而是强调存在一种可以想象出来的计算机。因此，当具体的操作实现图灵测试的理论想象时，需在判定“机器能否思维”的实践中确立以下观念：第一，思维本质是一种表现为具体实行的抽象规定性，既不能简单地从机器的逻辑单元构成、算法运行逻辑、机械驱动机制等的分析中进行推演分析判断，也不能从本质规定性的呈现程度来判定，只有不可区分的行为才是根本标准；第二，思维不依赖系统结构，虽然系统结构看起来是人或机器“智能”与否的关键性因素，但是“模仿游戏”之模仿者具有开放性与可修正性，可以超越既成系统的限定性；第三，“模仿游戏”之逼近所呈现的不是机器的功能问题，而是特质问题，即通过图灵测试的机器具有思维的独立性与自治性，虽然其存储力和算力是其“思维”的物理性限制，但是测试的时间长短和辨别智能的概率既保证了测试的可行性，又确证了机器“模

仿游戏”表征思维能力的可能性。

对于“模仿游戏”的设想，图灵表明自己并非打算研究游戏理论，而是要确认机器所完成的“模仿游戏”自然而然就是思维本身的表达方式。（参见同上，第59页）“模仿游戏”之逼近原则一经提出，就把思维的本质从近代哲学限定的理性推理能力中解放了出来。作为判定能否思维的中介，“模仿游戏”并不依赖计算机计算结果的准确性、理性决策的经济性、知识贮量的丰富性、逻辑进程的清晰性 etc. 来实现模仿的逼真或效应的一致——其本身就是思维存在。因为，就技术自身而言，“如果‘图灵测试’特指逻辑运算和推理能力，那么人们已经无法区分智能机器和人类”（王阳，第19页）。无论是“深蓝”计算机战胜卡斯帕罗夫还是 AlphaGo 战胜李世石，都从技术上说明逻辑推理、数量统计与实例分析等都容易被机械分解，而且算力和存储力就足以解决问题，根本用不着“模仿游戏”。但图灵要求机器“非常接近地模仿人类计算机的行动”（博登，第63页），且要完成这样的“模仿游戏”只需要机器实现能使游戏得以运行的规则，比如递归地调用指令以保证游戏运行的经济性，随机地产生场景以保证游戏运行中可以遭遇偶然情形，无限扩充容量以保证游戏运行过程可以回溯并在时间上无限制，等等。由此看来，图灵之所以要把“机器能否思维”这一问题转换成“是否存在可以想象得到的能够在模仿游戏中干得出色的数字计算机”（同上，第67页），是因为“模仿游戏”并不追求计算机与人行为模式的相似，而致力于在需要思维能力的游戏中有效地应对时间、策略和陌生等依赖思维与智能的情形。

由此看来，图灵测试中“模仿游戏”对思维活动的逼近是一个极具深意的技术性原则。图灵测试检验的是“机器能否思维”，而“模仿游戏”的实现过程使其具有更广阔的技术测试场景，通过隐喻表达的对智能的探问也就更为深入。“模仿游戏”所蕴含的技术思维使其既能够有效处理和准确判断“机器能否与人在所有纯智力领域竞争”，又能够应对“狡猾策略”对计算机的伪装问题，还能够解决“计算机沃森”的智力问答问题。“模仿游戏”是开放性的测试方式，它作为游戏具有最抽象的基本规定性，而作为“模仿游戏”则具有开放的适用性。无论是20世纪50年代图灵给出的原初规则，还是20世纪60年代排除“狡猾策略”后的修正规则，抑或21世纪“机器能否思维”就是“能否像人一样思考”的强化规则，“模仿游戏”的逼近都具有技术上的可操作性。也正因为如此，通过图灵测试一直是人工智能技术专家与工程技术人员的梦想。因此，“模仿游戏”对思维活动的逼近既迫使我们反思自笛卡尔以来的本质主义定义智能的理论思维，又敦促我们必须重视呈现思维能力活动的过程及其结果。从根本上看，这说明了以下问题：其一，思维有其基本的规定性，这恰恰就是“模仿游戏”强调“游戏”的意义之所在。虽然图灵说这里不需要讨论游戏，但是他以游戏模仿的拟人性作为重要的原则就足以说明，无论是机器思维还是人类思维都不能在本质主义的定义中得到规定。思维在活动中体现出的一般特性作为游戏的基本规定，仅适用于保障游戏顺利而开放地运行。其二，“模仿游戏”是图灵测试的核心规则，这就意味着图灵并没有对智能作本质主义的规定，而是在智能展开的现象学中对“机器能否思维”作出判断。因此，丹尼特（D. Dennett）认为，在人工智能追求通过图灵测试的过程中，重要的也许并非“能否通过”，而是有一天“我们或许同时能够学到一些关于社会心理学的无法忽视的知识”——这应该就是“模仿游戏”试图隐喻性地传达的内容。（参见王阳，第19页）其三，“模仿游戏”是从人的活动中抽象出来的规则，而非从智能的本质定义中推导出来的程序，其作为“机器能否思维”的测试原则说明，智能只有在驱动对象行动的场景中才能显现出来。在此意义上说，图灵测试“模仿游戏”的技术叙事一方面宣告了对智能进行本质主义定义的破产（因为智能本身是在多场景活动中呈现其存在与力量的），另一方面则说明，意义和理解与智能所实现的对象性驱动未必就是本质上截然分裂的，因为“模仿游戏”从未强调必须要把体现

“机器能否思维”的效应等同于无意义的对象本身。

因此,“模仿游戏”对思维活动的逼近作为图灵测试的核心原则,既说明了思维呈现过程的重要性,又强调了认可思维成效的判定主体理解思维观念的重要性。这就是说,“模仿游戏”既不以机器思维的效果来规定思维本质,又不以思维的本质来规定“机器能否思维”,而是在两者的关系判定中呈现思维的特性,这恰恰符合人类理解自我的一般方式。

三 功能主义的实现与关系主义的认同

图灵测试以行为效果的不可区分性肯定了不同自治系统行为成果的等效性。从直观上讲,图灵测试意味着对功能相同对象之间等价性的认同;从技术上讲,图灵测试扩展了通用图灵机的普适性;从理论上讲,图灵测试以“数学上的功能相似性”(博登,第64页)验证了机器思维的可能性;从存在论上讲,图灵测试创造了理解存在的新方式,“图灵所设想的‘图灵测试’采用的也是关系定义,即如果一台机器在与人对话中能够被识别为人,那么它就是一个智能存在”(赵汀阳,2020年,第4页)。因此,图灵测试对机器实现功能和能力的多重肯定,既体现了人工智能时代人们对客观实存效应的形而上学认可,又体现了人工智能时代人们理解存在方式的全新变化。

图灵测试对机器功能表征思维能力的确认并非着眼于理性解析机器基本原理、机器构成和机器逻辑上,而是着眼于机器整体效应。因此,图灵测试在判定“机器能否思维”的技术逻辑中自然地把笛卡尔“我思故我在”所表达的现代存在论从“我思”延展到“我行”再到“我成”。或者说,图灵测试以机器实现功能的被认可程度来确认机器能否思维,实际上强调了“思维”两个方面的特质。其一,机器应该能够以行之有效的理性逻辑产生客观可判定的对象化成效;其二,机器运行成效在主体缺失的前提下可以通过认可度的关系加以确认。因此,图灵测试对“实现与人思维相似功能的机器即可被视为具有思维能力”的肯定,显然不再把机器视作延长人自然肢体的简单工具,而把它看作具有整体行为能力的自治系统。这样的系统所产生的功能能够获得关系性的认可。也就是说,虽然能够通过图灵测试的机器(本质上就是人工智能)是由人定义并创制出来的,但是其高阶自动化运行的成效并非完全由定义本身所规定,而且其成效与人类行为的不可区分性更使得自治的机器系统表现出了似人的思维能力——一旦成效被关系认可,机器所实现的功能就是思维能力的表达。

图灵测试的关系主义是具有认可实现功能的技术叙事,虽然对于“机器能否思维”的判断目前尚存诸多争议,但关系主义认可不是无“功能性错误”和无“结论性错误”意义的确定性认可,而是通过人认可机器成效(包括对与错、好与坏等可能的成效)的意义得出的判断。或者说,图灵测试的技术叙事本身既包括数字计算机应该预设的基本范式,以避免人们基于归纳形成的对机器特质的一般性判断,又意味着机器与人作为关系对等的关系者项是可以交换的,以此解决了判断中机器前置还是主体前置的问题。按照图灵的观点,机器之所以能够超越“功能性错误”和“结论性错误”对机器思维的否定,并不是因为机器每次运行都必须产生被判断者设定的“正确目标”,而是因为从某种意义上说“机器毫无疑问地可以是自己的题材”(博登,第77页)。换言之,尽管形成反身性的自我意识仍然是人工智能的一个难题,但如今的计算机已经可以用能够预见的方式,“通过观察自己行为的结果……修改自己的程序,以便更有效地实现某种目标”(同上,第77页)。

图灵测试的关系主义认同强调的是实现目标的机器所呈现的作为关系者项的独立性。尽管图灵的这一论证与设想是技术性的,但一旦将其在关系中进行判断,以技术驱动的机器整体就不再仅仅是机械逻辑,而是自治系统。因此,当图灵测试把实现了功能的机器作为关系者项时,并非把机器作为“机械逻辑的集成与人能力的再实现”来展开功能与效果的比对,而是通过语言对话的方式来实现对

机器作为实体的认可。与此同时，它亦回避了通过建立规则来描述作为主体之人其行为结果的现实难题，以实现功能的具体性澄清了“指导规则”与“行为规律”之间可能存在的某种混淆，进而解决了行为非形式特性的问题；机器因为实现功能成为得到了关系认同的关系者项，事实上是对“科学观察”这一发现规则以及“关系认同是主体社会承认的重要方式”这一科学和哲学原则的充分运用。（参见同上，第81页）换言之，关系认同在操作上肯定了机器能够思维，在理论观念上肯定了被认同的主体与人具有相同的主体性。因此，图灵测试对功能实现的关系认同在技术判定上具有规律描述的周延性与判定结果的确定性，因为“如果我们用‘制约他的生活的行为规律’来替换上述论点中的‘制约他的生活的指导规律’，那么不周延中项就不再是不可逾越的”（同上，第81页）。在思维逻辑上，图灵测试与中国传统儒家“把人定义为人际关系中的人”（赵汀阳，2020年，第3页）的哲学思维具有同构性。机器与人能够实现同样的功能而无法相互区分的条件是区分者对被区分者的同构性认可，更是关系性定义。也正因如此图灵才认为，讨论“机器能否思维”之间本质上就等价于讨论如下情形是否可能：一台经过改进的特定的数字计算机，已具备合乎要求的存储、运行速度以及恰当程序等条件，在模仿游戏中，由人担任角色B而该计算机能够令人满意地扮演角色A。（参见博登，第67页）在此，该特定计算机因为实现了扮演角色A之“功能”，其实是可以与其角色直接互换的。这种互换性使得计算机与角色A能够“同一”而非相同，两者自然也就构成了相互独立且相互承认的关系者。进而按照图灵测试的逻辑，对实现功能的关系认同不仅意味着不同且独立的两个对象，更意味着两个均具有思维能力的独立对象。

更为重要的是，在图灵测试中得到关系认同的实现功能是“语言”，这一技术操作对象显然大有深意但又最难处理。本文主要关注语言的关系认同问题，即对于实现语言功能之对象的不可区分而言，如何规避意义理解的问题。这一问题是塞尔“中文之屋”诘问的核心。该诘问本身是在本质主义的定义中来呈现思维的逻辑、过程和要素的，并要求实现功能的对象严格符合之。然而，在“模仿游戏”中，机器是因为功能实现而被识别为人或能够像人一样思维的，所实现的功能能否被机器理解本身并不重要，因为关系识别逻辑本身就如同前述论证的那样承认了对象的独立性。卡普兰以“机械随机性”反驳塞尔，强调“能模拟大脑的机器必须有这样的行为表现，即好像有自由意志”（Turing, p. 464）。这就从根本上强调了关系认同的行为表现，至于行为表现的生成逻辑与生成过程就只能作为“黑箱”悬置。

由此看来，图灵测试中所蕴含的关系主义认同既是回答“何为智能”这一问题的本质主义的技术路线，又是对人造智能体与人类智慧本质关系的一种存在追问。或者说，图灵测试的技术逻辑远远超越了技术本身，是一个本质的存在论追问。比如，赵汀阳在尝试构造“升级版的图灵测试”时就提出这样的原则：“在涉及自身利益的博弈中，如果电脑能够在博弈中与人类对手成达均衡解……那么，这个电脑可以被认为是具有与人类共通的理性。”（赵汀阳，2022年，第22页）此即强调人类智能与人工智能之间的共通性（而非智能本质的共同性）才是关键。无论是“模拟游戏”的图灵测试还是“升级版的图灵测试”，抑或逼问存在的“终极版超图灵测试”（同上，第22页），其实都是从对思维和意识的追问中自然地延伸到了追问存在本身。

因此，图灵测试以技术叙事的明晰性表达了一种追问思维和意识的路径。其一，从图灵测试以关系认同实现功能机器的思维能力看，智能行为是依据，关系认可是关键。这一方面符合人类思维和智能发展的历史事实，即思维和智能对人而言本身就是一个社会历史性的事实，体现社会历史本身的认可才是真正的判断标准。也正是在这个意义上可以说，图灵测试给出的只是一个开放的原则，这也是至今我们尚未制造出能够通过图灵测试的机器的重要原因之一。其二，图灵测试虽然是在人工智能技

术发展初期提出来的技术性原则，但其关系认同的技术操作提出了一种在技术时代理解存在的思维逻辑，一方面真实体现了本质定义思维范式的式微，另一方面则强调了关系主义追问人类特质的意义。或者说，图灵测试以其技术叙事的方式追问我们研究人类智能的理论思维之合理性和有效性问题。对于人而言，“人应该在实践中证明自己思维的真理”（《马克思恩格斯选集》第1卷，第134页）。同理，“能思维”的机器也必须在实现功能中证明自身。显然，关系认同实现功能的图灵测试就意在表达这一原则。

结 语

图灵测试技术叙事的哲学追问表明，当人工智能技术以对象化的方式实现了人在社会历史层面的人性时，重新追问思维的本质和智能的本性就成为一个极具时代紧迫性的问题，需要从理论层面展开全新思考。一方面，人工智能在技术逻辑上是以物理逻辑的方式实现对人类思维智能之操作的，这既证明了思想的真理性，又体现了智能的巨大力量；另一方面，人工智能对人类思维能力的实现，既说明硅基人造物的运行与碳基生命主体的感性实践在结果上具有相似性甚至同一性，又需要我们以诸如图灵测试这样的原则对两者“是否同一”“因何同一”“如何同一”等问题作出判定。因此，图灵测试表达的具体技术逻辑，既是我们对物体系（硅基人造物）的人工智能的描述性认知，也是我们对人类思维能力的技术性想象。而正是这一想象的技术原则使得我们今天既需要重新审视我们理解人的思维逻辑与理论方式，亦需要正视人的对象化产物，并深入理解对象化在诸如硅基物体系中的社会历史性在何种意义上改变着我们的现实行为，进而深层地影响我们生物体的内在结构。当然，依循图灵测试的哲学追问，我们还需要进一步探析图灵测试的技术逻辑，这亦更加要求我们正视这种追问，直面人本身。

参考文献

- 博登，2001年：《人工智能哲学》，刘西瑞、王汉琦译，上海译文出版社。
- 布兰顿，2019年：《在理由空间之内：推论主义、规范实用主义和元语言表达主义》，孙宁、周靖、黄远帆、文杰译，上海人民出版社。
- 查默斯，2013年：《有意识的心灵：一种基础理论研究》，朱建平译，中国人民大学出版社。
- 高新民、罗岩超，2021年：《“图灵测试”与人工智能元问题探微》，载《江汉论坛》第1期。
- 吕陈君、邹晓辉，2020年：《通用人工智能的数学基础初探》，载《自然辩证法研究》第3期。
- 《马克思恩格斯选集》，2012年，人民出版社。
- 派利夏恩，2007年：《计算与认识：认知科学的基础》，任晓明、王左立译，中国人民大学出版社。
- 彭罗斯，2018年：《皇帝新脑》，许明贤、吴忠超译，湖南科技出版社。
- 珀尔、麦肯齐，2019年：《为什么：关于因果关系的新科学》，江生、于华译，中信出版社。
- 涂良川，2020年：《马克思历史唯物主义视阈中的人工智能奇点论》，载《东北师大学报（哲学社会科学版）》第1期。
- 王阳，2016年：《图灵测试六十五年——一种批判性的哲学概念分析》，载《科学技术哲学研究》第2期。
- 扎卡达基斯，2017年：《人类的终极命运——从旧石器时代到人工智能的未来》，陈朝译，中信出版社。
- 赵汀阳，2020年：《技术的无限进步也许是一场不可信任的赌博》，载宋冰编著《智能与智慧：人工智能遇见中国哲学家》，中信出版社。
- 2022年：《人工智能的神话或悲歌》，商务印书馆。
- Turing, A. M., 1951, *Programmer's Handbook for Manchester Electronic Computer*, Manchester: University of Manchester Computing Laboratory.

（作者单位：太原理工大学马克思主义学院、山西大学乡村振兴研究院；华南师范大学马克思主义学院）

责任编辑 孙婧一