arXiv:2304.11221v1 [physics.ed-ph] 21 Apr 2023

人工智能工具能为物理学入门课程中的作业评分吗?

Gerd Kortemeyer*Educational Development and Technology* [*],

ETH Zurich, Zurich, Switzerland [†]（日期：2023年4月25日）

问题求解是任何物理课程不可或缺的一部分，大多数物理教师可能都会认同，评估相关学习者能力的最佳方式是考察解题路径：重要的不只是最终解答，还有学习者如何得出答案。遗憾的是，为书面推导提供有意义的反馈，比单纯给结果评分耗费更多人力和资源：目前，计算机可以为结果评分，而前者则涉及需要由人工评分的手写解答。本探索性研究提出了一种借助人工智能的工作流程，用于评阅书面的物理问题解答，并利用 GPT-4 评估实际评分环节的可行性。研究发现，人工智能工具能够提供有助于形成性评价场景的反馈；但在总结性评价场景中，尤其是高风险场景下，它只能用于初步评分，以对解题方法进行分类和标记。

I. 引言

A. 生成式预训练Transformer

2022年秋季，生成式预训练变换器（GPT） [1] 迅速登上世界舞台，成为一种具有惊人能力的公开可用人工智能工具。Chat-GPT 是这一底层大型语言模型（LLM）的文本界面。GPT 本质上是一种使用基于神经网络的自动补全算法生成貌似合理的虚构内容的工具；在这方面，它类似于智能手机上的自动补全功能：用户撰写短信时，该功能会建议可能出现的下一个词。然而，GPT 的训练基于从公开和专有来源中收集的大规模文本语料库，并由人类进行了广泛的微调。此外，GPT 处理的不是单词，而是更细粒度的词元；这些词元与单词中的音节相似，但并不完全相同。

ChatGPT 支持基于文本的对话：用户提供提示词，GPT 给出回复。这些对话极其逼真，很像人类，而在许多情境下，该系统很可能通过图灵测试 [2]。在使用过程中，除非开发 GPT 的公司 OpenAI 持续对其进行人工训练，否则系统整体不会再学习。不过，它会在特定对话的范围内学习，因此能够参考同一对话中较早出现的陈述。但到了一定时候，它会触及内部所谓的“令牌限制”——它只能在记忆中保留有限数量的词元，这与人类工作记忆的局限性相似。因此，对话不能过于冗长，否则系统会出现错误消息，或者看起来会忘记对话中较早提到的内容（事实证明，这会成为本研究的一项局限）。

GPT 背后大规模的训练工作催生了一些语言模型未必能预期具备的能力。研究发现，GPT-4 已能够在多项标准化考试中取得位于高百分位的成绩 [3–5]。早期版本 GPT-3 已经能够以名义上的水平通过物理学入门课程 [6]，而且有迹象表明，GPT-4 对物理概念的掌握甚至更好 [7]。

虽然 ChatGPT 仅限于基于文本的输入和输出，但 GPT-4 本身是多模态的，因此可以接受图像输入 [5]。在其他示例中，该公司在广告视频中演示了如何将手写内容和手写数学公式转换为机器可读文档。然而，使用此功能需要访问其应用程序编程接口（API），遗憾的是，该接口受到 OpenAI 的限制——截至 2023年4月，等待名单很长，因此本研究无法测试此功能。

B. 求解物理问题

使用人工智能对问题解答进行评分的想法早已存在 [8]，但往往受到一项需求的阻碍：必须针对特定问题训练人工智能。换言之，人工智能通过人类学习如何对某个特定问题评分，最终模仿人类对该问题的评分方式。GPT 及类似的“预训练”系统则有望提供一种更具普适性的解答方案，能够对以前从未接触过的问题进行临时评分。这一承诺似乎已触手可及，因为这些相同的“预训练”系统也能够解决各种以前从未接触过的问题 [6]。

策略性地解决物理问题需要逻辑、概念和数学方面的能力 [9, 10]，而物理教育研究中几乎没有哪个主题比这项研究得更深入 [11]。尽管可以借助计算机系统对物理问题的最终解答以及预先设定的分步支架进行不同程度的评估 [12–19]，但要深入分析学习者复杂的问题解决能力，就需要评估其解题路径和推导过程 [20, 21]。然而，目前批改书面解答仍需人工完成，而这种更具

* kgerd@ethz.ch

† 也在美国东兰辛的密歇根州立大学任职。

意义的反馈对最终学习成果的影响，可能会因人工批改资源有限、反馈返回延迟而受到限制，从而减少获得此类反馈的机会 [22]。本研究旨在探讨，学生能否从书面作业和考试题中获得有意义的反馈，同时还能享有以往只有以解答为中心的在线系统才能提供的优势：更频繁的形成性评估和即时反馈。

II. 场景

所提出的场景包含多个步骤，如图 1 所示。在一种可能的情境中，学习者在纸上书写题目解答；尽管这种方式看起来或许有些过时，但大多数人都会认同，输入数学公式十分繁琐，而使用微软 Word 等软件内置的公式编辑器来编排公式则既耗时又容易令人沮丧。输入数学内容的既定标准是 LaTeX。尽管一些专家能够用 LaTeX 进行“思考”，但纸笔书写的效率和直观性要高得多（甚至有观点认为，这种物质性体验至关重要 [23]）。这一场景尤其适用于考试，因为考试可以采用传统的、有监考人员监督的“离线”模式进行。

1. 第一步很简单，只需将作业纸扫描成 PDF 或图像格式；大多数复印机都可以通过自动送入学生作业纸来完成这一操作，真正的挑战可能是揉皱的试卷导致卡纸。

2. 随后，使用基于人工智能的光学字符识别技术处理 PDF，并将其转换为机器可读格式，例如 LaTeX。在本研究中，由于无法使用适当的 API，这些步骤通过模拟实现；这一步骤的可靠性将作为未来一项独立研究的主题。

3. 随后，由人工智能工具（本研究中使用 GPT-4）对 LaTeX 文档进行评分。此步骤可能包含多轮相互独立的评分，之后再对评分结果进行汇总。不同的处理方式包括：标记评分结果存在分歧的解答，以便后续由人工进行仲裁；或者将解答归类为相似性类别，作为加快人工评分速度的初步步骤。

第一步的第二种可能方案是使用手写计算，学习者可以直接在屏幕上绘制。如今，一些平板电脑和笔记本电脑通过适当的摩擦力模拟纸张的触感，学生似乎也能适应这项技术，许多人都用它来记课堂笔记便是明证。第一步的第三种方案是用智能手机拍摄或扫描解答，再将其上传到在线系统。

后两种方案不太适合考试场景，因为在线意味着可以访问各种资源、通信渠道和在线工具（包括人工智能工具！）。在线环境带来的这些额外便利，要么需要融入任务本身（这可能会增加任务难度），要么就需要尝试通过锁定技术加以限制 [24]。

这类系统中的人工智能组件极不可能以本地部署的方式安装，因此可能只有商业云端解答方案才具有可行性。这将引发隐私和数据安全方面的担忧，需要妥善解决，尤其是在处理必修作业 [25]时。

III. 方法论

A. 生成示例解答

作为本研究的示例，我们选取了一道简单的时变RC电路问题。该问题以图2中的题目形式呈现，涉及一些概念、策略和数学方面的挑战，但很可能以某种形式出现在大多数基于微积分的物理学导论课程中。题目给出了电容器两端的初始电势差，尽管解题并不需要这一信息；这一多余信息容易让人立即计算流过电容器的初始电流，从而使解答不那么直接，因为该电流随后会消去。尤其是在不进行符号运算、而是立即“代入计算”的情况下，这会引入不必要的步骤 [26]。

由于 GPT 使用概率算法，同一提示输入两次会得到不同的回答。本研究利用这一特性生成了 25 份各不相同的示例解答。GPT-4 的推理能力优于早期版本，因此更有可能给出正确解答。为使样本中也包含看似合理但错误的解答，大多数解答由早期版本生成。

在本研究进行时（2023年4月中旬），通过 ChatGTP 可以使用 GPT 的三种不同模型：GPT-4（2023年3月23日发布 [27]）、默认 GPT-3.5 和旧版 GPT-3.5。具体而言，根据图2中的提示，解答 1–5 由 GPT-4 生成，解答 6–13 由默认 GPT-3.5 生成，解答 14–25 由旧版 GPT-3.5 生成；这些解答列于图2–7中。

作为输出格式，选择了 LaTeX，因为这很可能是基于人工智能的光学字符识别系统的输出结果（见图 1）。在输出阶段，ChatGPT 会将显示方程转换为符号字体，因此将这些表达式复制到文本文件时，LaTeX 源代码会丢失；因此要求使用行内表达式。通过 API 直接访问 GPT 时，就不需要采取这样的变通方法。

可以说，这些示例是在让 GPT 给自己评分。尽管表述略显冗长，并采用了教师和教材惯用的尊称“我们”，但就本探索性研究而言，这些解答似乎具有代表性——

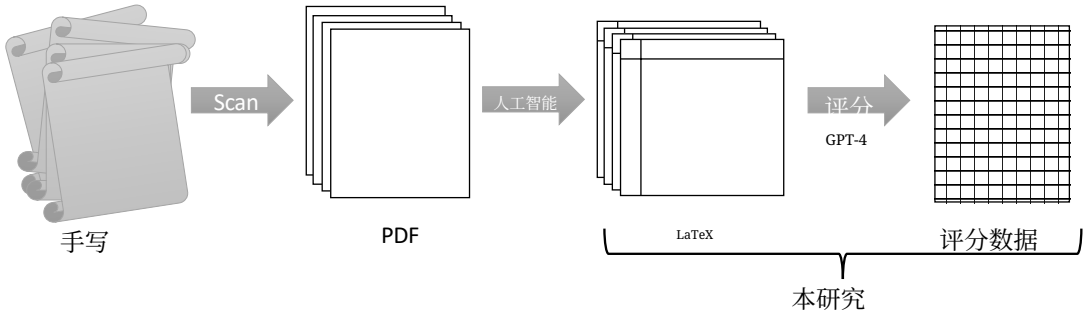

图 1. 人工智能辅助批改手写作业的可能工作流程。

提示： 在 $t = 0$，一个电阻器（电阻为 R=10 欧姆）与一个已充满电的电容器（电压为 V=5V，电容为 C=100毫法）并联连接。在什么时间 t_2，流过电阻器的电流会达到初始电流的一半？不要使用行间 LaTeX，只使用行内 LaTeX。

解答 1： 我们有一个电阻器（R $=$ 10 欧姆）与一个已充满电的电容器（V $=$ 5V，C $=$ 100毫法）并联连接。为了求出电阻器中流过的电流达到初始电流一半时的时间 t_2，我们需要分析一个 RC 电路。在 RC 电路中，电容器两端的电压随时间变化的函数为：$V_C(t) = V_0 e^{-t/RC}$，其中，V_0 是电容器两端的初始电压（5V），R 是电阻（10 欧姆），C 是电容（100 毫法），t 是时间。流经电阻器的电流随时间变化的函数为：$I(t) = \frac{V_0}{R} e^{-t/RC}$。现在，我们需要求出电流达到初始电流一半时的时间 t_2，因此：$\frac{1}{2} I(0) = I(t_2)$，将给定的表达式代入 $I(0)$和 $I(t_2)$，得到：$\frac{1}{2}\frac{V_0}{R} = \frac{V_0}{R} e^{-t_2/RC}$。将两边同时除以$(V_0/R)$，得到：$\frac{1}{2} = e^{-t_2/RC}$。对两边取自然对数，得到：$-\ln 2 = -\frac{t_2}{RC}$。现在，我们可以求解 t_2：$t_2 = RC \ln 2$。将给定的 R 和 C 的值代入：$t_2 =$(10，欧姆)(100 × 10^{-3}，法拉) ln 2。$t_2 = 1$，s ln 2。因此，电阻器中流过的电流达到初始电流一半时的时间 t_2 约为 0.693，s。

图2。本研究所用的提示词和解答 1（由 GPT-4 生成）。

足以涵盖学生可能提交的各种作答。从几乎完美的解答 1 和 12，到完全偏离要求的解答 9 和 19，范围十分广泛；其中还包括预料之中的多余计算，以及将数值从一个公式代入下一个公式的过程。顺便一提，尽管这组样本旨在模拟人类的问题求解，但它也展示了推理能力在短短几个月内从旧版 GPT-3.5 到 GPT-4 的进步。这组样本展示了推理能力的提升。这些解答现已成为实际评分研究的基础（图 1 的最后一步）。

B. 解答的评分

GPT-4 在所有情况下都正确地解决了问题，因此使用它进行评分似乎是合适的。解答依据解题方法的正确性、符号推导的正确性、数值结果的正确性和直截了当性这几项评分标准，按照 0 分（最差）至 4 分（最好）的量表进行评分。此外，还要求针对每份解答提供一句话的反馈。评分细则得分经过合并，形成总分，并对最终的数值结果赋予更高权重。最后，系统被提示根据方法相似性生成解答之间的相关性表；图8展示了相关提示词的示例。

由于 GPT 具有概率性，它不会为每份解答得出唯一确定的评分组合。类似于让多名人工评分者查看同一份解答，每份解答都被评分了多次。受 GPT 令牌限制的影响，无法同时将全部 25 份解答输入系统。因此，这些解答被分成随机组合的、每批包含 5 份解答的批次，并通过 75 次对话进行处理，从而使每份解答平均经历 15 个评分轮次。遗憾的是，ChatGPT 大多数时候并未按要求提供 CSV格式的表格，因此需要随后使用文本编辑器和 Excel 对其回复进行一些人工重新格式化。平均 15 位“评分者”提供的叙述性一句话反馈消息，使用 GPT-4 进行总结，整合为一段更长的陈述，以反映多数意见，所用提示词如图9所示。

作者按照相同的评分标准，对论证正确性、符号推导、数值结果以及直截了当性进行了独立评分。当然，人工评分的决定也带有一定主观性，但本研究未对此作进一步探讨。不过，读者可以自行判断每个解答；这些解答见图2–7。

解答 2：我们可以使用 RC 电路方程，求出电流降至初始值一半所需的时间。电路中的电流可用方程表示：$I(t) = I_0 e^{-t/RC}$其中，$I(t)$表示时刻 t的电流，I_0 是初始电流，RC 是电路的时间常数。初始电流I_0可根据欧姆定律求得：$I_0 = \frac{V}{R}$我们要求的是电流为初始电流一半时的时间 t_2：$I(t_2) = \frac{1}{2} I_0$现在代入 R、V和 C 的数值：$I_0 = \frac{5V}{10}\,\Omega = 0.5A$将 I_0 的数值代入 $I(t_2)$的方程：$0.25A = \frac{1}{2}(0.5A)e^{-t_2/(10\Omega \cdot 100\times 10^{-3}F)}$我们要求 t_2。首先，两边同时除以 $0.5A$，简化方程：$0.5 = e^{-t_2/(10\Omega \cdot 100\times 10^{-3}F)}$现在，对两边取自然对数：$\ln(0.5) = -t_2/(10\Omega \cdot 100\times 10^{-3}F)$然后，两边同时乘以时间常数 RC：$t_2 = -\ln(0.5)(10\Omega \cdot 100\times 10^{-3}F)$最后，计算 t_2的值：$t_2 \approx 0.693 \cdot (10\Omega \cdot 100\times 10^{-3}F)$ $\approx 0.693 \cdot 1s \approx 0.693s$因此，在 $t_2 \approx 0.693$秒时，有一半的初始电流流过电阻器。

解答 3：要解决这个问题，我们首先要找出电路中电流随时间变化的方程，然后找出电流降至初始值一半时的时间。在任意给定时间 t，流过电阻器的电流可以使用电容器放电方程确定：$I(t) = I_0 e^{-t/RC}$其中，$I(t)$ 是时间 t 时的电流，I_0 是初始电流，R 是电阻，C 是电容，t 是经过的时间。我们首先需要找出 I_0，即在 $t = 0$时流过电阻器的初始电流。可以使用欧姆定律求得：$I_0 = \frac{V}{R}$已知数值为 $R = 10$欧姆和 $V = 5$V。$I_0 = \frac{5}{10}$V 10 欧姆$= 0.5$A 现在我们得到了电流方程：$I(t) = 0.5e^{-t/RC}$我们要找出时间 t_2，此时电流为其初始值的一半：$I(t_2) = \frac{1}{2} I_0 = 0.25$，A 因此，$0.25 = 0.5e^{-t_2/RC}$两边除以 0.5：$0.5 = e^{-t_2/RC}$现在对两边取自然对数：$\ln 0.5 = -\frac{t_2}{RC}$已知电容为 $C = 100$毫法 $= 100\times 10^{-3}$法拉，现在可以求解 t_2：$t_2 = -RC\ln 0.5$ $t_2 = -10, \quad 100\times 10^{-3} \quad 0.5$(欧姆)(，法拉)ln，s $t_2 \approx 0.693\times 10^{-3}$ 因此，在 $t_2 \approx 0.693$ms 时，流过电阻器的电流为初始电流的一半。

解答 4：要找出流过电阻器的电流为初始电流一半时的时间，我们首先需要找出初始电流以及流过电阻器的电流随时间变化的方程。初始电流（I_0）可以使用欧姆定律求得：$I_0 = \frac{V}{R}$其中，V 是电容器两端的初始电压（5V），R 是电阻（10 欧姆）。$I_0 = \frac{-5}{10}$V 10 欧姆$= 0.5$A 流过电阻器的电流随时间的变化可以使用 RC 电路中的电流公式求得：$I(t) = \frac{V}{R} e^{-\frac{t}{RC}}$其中，$I(t)$ 是时间 t时的电流，V 是电容器两端的初始电压，R 是电阻，C 是电容，t 是时间。在本题中，$V = 5$V、$R = 10$ 欧姆，以及$C = 100$毫法 $= 0.1$法拉。现在我们需要找出时间 t_2，此时流过电阻器的电流为初始电流的一半，即 $I(t_2) = 0.5I_0 = 0.25$A。我们可以列出方程：0.25A $= \frac{5}{10}$V 10 欧姆$e^{-\frac{t_2}{10}}$(欧姆)(0.1 法拉) 现在需要求解 t_2。首先，我们可以化简方程：$0.5 = e^{-\frac{t_2}{1s}}$接下来，对两边取自然对数：$\ln(0.5) = -\frac{t_2}{1}$s最后，求解 t_2：$t_2 = -1\text{s} \cdot \ln(0.5) \approx 0.6931$s 因此，在 $t_2 \approx 0.6931$s 时，流过电阻器的电流为初始电流的一半。

解答 5： 要解答这个问题，我们先求出初始电流，然后确定电流降至该值一半所需的时间。流经电阻器的初始电流（I_0）可用欧姆定律求得：$I_0 = \frac{V}{R}$其中 $V = 5V$且 $R = 10$，欧姆：$I_0 = \frac{5}{10} = 0.5$，安。现在，我们来求 RC 电路的时间常数（τ）：$\tau = RC$其中 $R = 10$，欧姆且 $C = 100$，毫法 $= 0.1$，法：$\tau = 10 \times 0.1 = 1$，秒。电路中任意时刻 t 的电流可表示为：$I(t) = I_0 e^{-t/\tau}$我们要求出电流为初始电流一半时的时间 t_2：$\frac{1}{2} I_0 = I_0 e^{-t_2/\tau}$将等式两边除以 I_0：$\frac{1}{2} = e^{-t_2/\tau}$现在，对等式两边取自然对数：$\ln\left(\frac{1}{2}\right) = -t_2/\tau t_2 = -\tau \ln\left(\frac{1}{2}\right)$由于 $\tau = 1$，秒：$t_2 = -1 \quad \left(\frac{1}{2}\right)\ln$，秒。因此，在 $t_2 \approx 0.693$，秒时，流经电阻器的电流为初始电流的一半。

图3. 图2中提示的解答2和5。

Solution 6: Th电流 t in the circuit at time t is given b 由方程： $i(t) = \frac{V}{R}e^{-\frac{t}{RC}}$，其中 RC 是电路的时间常数。在时间 t_2，有一半的初始电流流过电阻器，因此：$\frac{1}{2}i(0) = \frac{V}{R}e^{-\frac{t_2}{RC}} \Rightarrow \frac{1}{2}\cdot\frac{V}{R} = \frac{V}{R}e^{-\frac{t_2}{RC}} \Rightarrow \frac{1}{2} = e^{-\frac{t_2}{RC}}$对两边取自然对数，得到：$\ln\frac{1}{2} = -\frac{t_2}{RC} \Rightarrow t_2 = -RC\ln\frac{1}{2}$将问题中给出的数值代入，得到：$t_2 = -10\cdot0.1\cdot\ln\frac{1}{2} \approx 0.69$ 秒。因此，在 $t_2 \approx 0.69$ 秒时，有一半的初始电流流过电阻器。**解答 7：** 流过电容器的电流为 $I(t) = C\frac{dV}{dt}$，电容器两端的电压为 $V(t) = V_0e^{-\frac{t}{RC}}$ 其中，V_0 是电容器两端的初始电压，R是电路的电阻，C 是电容器的电容。在 $t = 0$时，电容器两端的电压为 $V_0 = 5$ V，因此有：

$$I(0) = C\frac{dV}{dt}\bigg|t = 0 = C\frac{d}{dt}\left[V_0e^{-\frac{t}{RC}}\right]\bigg|t = 0 = -\frac{V_0}{R} = -0.5 \text{ A}$$

其中，负号表示电流流动的方向与假定方向相反。

要找出通过电阻器的电流为初始电流一半时的时间 t_2，我们需要求解方程：

$$I(t_2) = -\frac{0.5}{2} = -0.25 \text{ A}$$

我们可以使用电容器两端的电压方程来求解 t_2：

$$V(t_2) = V_0e^{-\frac{t_2}{RC}} = 2.5 \text{ V}$$

对两边取自然对数并重新整理，得到：

$$\ln\left(\frac{V_0}{2.5}\right) = \frac{t_2}{RC}$$

S代入 V_0、C 和 R 的值，我们得到：

$$\ln\left(\frac{5}{2.5}\right) = \frac{t_2}{10\times10^{-3}}$$

化简后，我们得到：

$$t_2 = 10\times10^{-3}\ln(2) \approx 6.93 \text{ 毫秒}$$

因此，在 $t_2 \approx 6.93$ ms 时，流经电阻器的电流为初始电流的一半。

解答 8： 电路中的电流由以下方程给出：$i(t) = \frac{V}{R}e^{-\frac{t}{RC}}$其中，$i(t)$ 是时间 t 时的电流，V是电容器两端的初始电压，R 是电阻，C 是电容，e 是自然对数的底数。要找出电阻中流过初始电流一半时的时间 t_2，我们需要在以下方程中求解 t：$\frac{i(t_2)}{i(0)} = \frac{1}{2}$ 代入 $i(t)$ 的方程并化简，得到：$e^{-\frac{t_2}{RC}} = \frac{1}{2}$ 对两边取自然对数，得到：$-\frac{t_2}{RC} = \ln\left(\frac{1}{2}\right)$求解 t_2，得到：$t_2 = -RC\ln\left(\frac{1}{2}\right) = RC\ln(2) \approx 0.69$ 秒。因此，电阻中流过初始电流一半时的时间约为 $t_2 = RC\ln(2)$。

解答 9： 电路中电容器的电流可计算为 $i(t) = C\frac{dv(t)}{dt}$，其中 $v(t)$ 是时间 t 时电容器两端的电压。本例中，电容器两端的电压为 $V = 5V$，电容为 $C = 100mF = 0.1F$，因此，流过电阻器的初始电流为：$i_0 = C\frac{dV}{dt} = 0$这是因为在时间 $t = 0$，电容器已充满电，所以电路中没有电流流过。之后任意时刻 t 的电流可使用以下公式计算：$i(t) = C\frac{dV(t)}{dt}$ 任意时刻 t 电容器两端的电压可通过以下公式求得：$V(t) = Ve^{-\frac{t}{RC}}$其中，$RC$ 是电路的时间常数，等于电阻 R 与电容 C 的乘积。本例中，$RC = 1s$。因此，时间 t 时的电流为：$i(t) = C\frac{dV(t)}{dt} = \frac{V}{R}e^{-\frac{t}{RC}}$要找出电阻中流过初始电流一半时的时间 t_2，我们可以令 $i(t_2) = \frac{i_0}{2}$，并求解 t_2：$\frac{V}{R}e^{-\frac{t_2}{RC}} = \frac{i_0}{2} = 0$

$$e^{-\frac{t_2}{RC}} = 0$$

这个方程无解，这意味着初始电流的一半永远不会流过电阻器。

这是因为电容器不允许电压或电流突然变化，而是通过储存或释放电荷来抵抗这种变化。电路刚接通时，电容器已充满电，没有电流流动。随着电容器两端的电压降低，流过电阻器的电流增大，但始终不会达到初始电流的一半。电流随时间继续呈指数衰减；随着电容器两端的电压趋近于零，电流也趋近于零。

解答10： 流过电阻器和电容器的电流由方程给出：$i(t) = \frac{V}{R}e^{-\frac{t}{RC}}$ 其中 RC 是电路的时间常数，在本例中等于 $RC = 10^{-1}$ 秒。初始电流 $i(0)$可通过代入 $t = 0$求得：$i(0) = \frac{V}{R} = 0.5$ 安培要找出初始电流的一半流过电阻器时的时间 t_2，我们需要在方程中求解 t：$\frac{i(t_2)}{i(0)} = \frac{1}{2}$ 将 $i(t)$ 的表达式代入并求解 t，得到：$\frac{V}{R}e^{-\frac{t_2}{RC}} = 0.25 \Rightarrow e^{-\frac{t_2}{RC}} = \frac{1}{2} \Rightarrow -\frac{t_2}{RC} = \ln\frac{1}{2} \Rightarrow t_2 = -RC\ln\frac{1}{2}$ 代入 R、C和 V的值，得到：$t_2 = -10^{-1}\ln\frac{0.25\times10}{5} \approx 0.138$ 秒 因此，在 $t_2 \approx 0.138$ 秒时，初始电流的一半流过电阻器。

FIG. F 中提示的第 6 至第 10 个解答 图 2。

解答 11：流经电容器的电流由以下公式给出：$i(t) = C\frac{dv(t)}{dt}$ 其中，$i(t)$ 是时间 t 时的电流，$v(t)$ 是时间 t 时电容器两端的电压。电容器两端的电压会随时间呈指数衰减，公式如下：$v(t) = V_0 e^{-\frac{t}{RC}}$ 其中，V_0 是电容器两端的初始电压，R 是电阻器的电阻，C 是电容器的电容，t 是电容器完全充电后经过的时间。流经电阻器的电流由欧姆定律给出：$i_R(t) = v(t)_R$ 将 $v(t)$ 的方程代入该方程，得到：$i_R(t) = \frac{V_0}{R} e^{-\frac{t}{RC}}$ 要找出电阻器中流过初始电流一半时的时间 t_2，我们需要求解以下关于 t 的方程：$i_R(t_2) = \frac{1}{2} i_R(0)$ 代入 $i_R(t_2)$ 和 $i_R(0)$ 的表达式，得到：$\frac{V_0}{R} e^{-\frac{t_2}{RC}} = \frac{1}{2}\frac{V_0}{R}$ 化简该方程，得到：$e^{-\frac{t_2}{RC}} = \frac{1}{2}$ 对等式两边取自然对数，得到：$-\frac{t_2}{RC} = \ln 1\,2$ 求解 t_2，得到：$t_2 = -RC \ln 1\,2$ 代入 R 和 C 的值，得到：$t_2 = -10\,\Omega \times 100\,\mu\text{F} \times \ln 1\,2$ 计算该表达式，得到：$t_2 \approx 0.69$ s 因此，约在 $t_2 = 0.69$ s 时，电阻器中流过的电流为初始电流的一半。解答 12：时间 t 时流经电阻器和电容器的电流由以下方程给出：$i(t) = \frac{V}{R} e^{-\frac{t}{RC}}$ 其中，RC 是电路的时间常数。时间 $t = 0$ 时，电流为：$i(0) = \frac{V}{R} = \frac{5}{10} = 0.5$ A 要找出电阻器中流过初始电流一半时的时间 t_2，我们需要求解以下方程：$i(t_2) = \frac{1}{2} i(0)$ 代入 $i(t)$ 的表达式并化简，得到：$\frac{V}{R} e^{-\frac{t_2}{RC}} = \frac{1}{2}\frac{V}{R} e^{-\frac{t_2}{RC}} = \frac{1}{2}$ 对等式两边取自然对数，得到：$-\frac{t_2}{RC} = \ln \frac{1}{2} = -\ln 2 t_2 = RC \ln 2$ 代入 R 和 C 的值，得到：$t_2 = (10\,\Omega)(100\,\mu\text{F}) \ln 2 = 0.693\,\text{s}$ 因此，经过 $t_2 = 0.693$ s 后，电阻器中流过的电流为初始电流的一半。

解答 13：流过电路的电流可用方程表示：$I = \frac{V}{R} \cdot e^{-\frac{t}{RC}}$ 其中 I 是电流，V 是电压，R 是电阻，C 是电容，t 是时间。要找出电阻器中流过一半初始电流的时间 t_2，我们需要令 I 等于其初始值的一半，并求解 t：$\frac{V}{R} \cdot e^{-\frac{t_2}{RC}} = \frac{1}{2} \cdot \frac{V}{R} \cdot e^{-\frac{0}{RC}}$ 约去各项后，得到：$e^{-\frac{t_2}{RC}} = \frac{1}{2}$ 对等式两边取自然对数，得到：$-\frac{t_2}{RC} = \ln\left(\frac{1}{2}\right)$ 求解 t_2，得到：$t_2 = -RC \cdot \ln\left(\frac{1}{2}\right)$ 代入 R 和 C 的数值，得到：$t_2 = -10 \cdot 100 \cdot \left(\frac{1}{2}\right)$ 欧姆 毫法 ln 秒 因此，约经过 $t_2 \approx 0.69$ 秒后，流过电阻器的电流为初始电流的一半。解答 14：电容器两端的电压随时间的变化关系为：$V(t) = V_0 e^{-t/RC}$ 其中 V_0 是初始电压，RC 是电路的时间常数。时间常数定义为 $RC = \frac{R}{C}$。流过电阻器的电流随时间的变化关系为：$I(t) = \frac{dV}{dt} = -\frac{V_0}{RC} e^{-t/RC}$ 在时间 $t = 0$，电流达到最大值，其值为：$I_0 = -\frac{V_0}{RC}$ 初始电流的一半等于 $\frac{1}{2} I_0$。可通过令 $I(t_2) = \frac{1}{2} I_0$ 并求解 t_2，求出这一情况发生的时间 t_2：$\frac{1}{2} I_0 = -\frac{V_0}{RC} e^{-t_2/RC}$ $e^{-t_2/RC} = \frac{1}{2}$ $-\frac{t_2}{RC} = \ln 1\,2$ $t_2 = -RC \ln 1\,2$ 代入 R 和 C 的数值：$t_2 = -10$ 欧姆 $\cdot 100$ 毫法 $\ln \frac{1}{2} = -10 \cdot 10^{-3} \ln \frac{1}{2} \approx 6.93$ 秒 因此，流过电阻器的电流达到初始电流一半时的时间 t_2 约为 6.93 秒。

图5. 图2中提示词的解答11至14。

我们考虑了几种解答之间的相似性度量，作为聚类的依据：人工和人工智能评分在评分标准上的得分相似性，以及人工智能工具回复中给出的方法相似性；由于受到令牌限制，一次只能评阅五份解答，因此通过对各个 5×5-矩阵取平均值，汇总得到了整体 25×25-矩阵。聚类使用 R [29] 中的 gplots 包 [28] 完成。

解答 15：电容器两端的电压 V_C 与流经电阻器的电流 I 之间满足以下方程：$V_C = IR$由于电容器两端的电压和电容保持不变，流经电阻器的电流可表示为：$I = \frac{dQ}{dt} = C\frac{dV_C}{dt} = \frac{V}{R} = \frac{5V}{10}\,\Omega = 0.5A$ 其中， Q 是电容器上的电荷。电容器两端的电压降低时，流经电阻器的电流也会减小，因此在时间 t_2，流经电阻器的电流将降至其初始值的一半：$I(t_2) = 0.5 \cdot 0.5A = 0.25A$电容器两端的电压可表示为：令 $I(t_2) = 0.25A$ 并求解 t，得：$0.25A = 0.5A \cdot e^{-\frac{t_2}{100m\,\Omega \cdot 10^{-3}F}}$ $-\frac{t_2}{100m\,\Omega \cdot 10^{-3}F} = \frac{1}{2}\,\Omega \ln \Omega \ln$ $t_2 = 100m\,\Omega \cdot 10^{-3}F \cdot$ $V_C = V_0 e^{-\frac{t}{RC}} = 5V \cdot e^{-\frac{t}{100m\,\Omega \cdot 10^{-3}F}} \cdot \frac{1}{2} \approx 0.693 \cdot 100 \cdot 10^{-3} \approx 69.3 \cdot 10^{-3}s = 69.3ms$

因此，流经电阻器的初始电流降至一半时的时间 t_2 约为 69.3 毫秒。解答16：电容器和电阻器两端的电压可表示为：$V(t) = V_0 e^{-t/(RC)}$ 其中， V_0 是电容器两端的初始电压。流经电阻器的电流可表示为：$I(t) = \frac{dV}{dt} = -\frac{V_0}{RC}e^{-t/(RC)}$流经电阻器的初始电流降至一半时的时间 t_2 可通过令 $I(t_2) = \frac{I_0}{2}$ 并求解 t_2来求得：$\frac{I_0}{2} = -\frac{V_0}{RC}e^{-t_2/(RC)} \frac{I_0}{2}RC = -V_0 e^{-t_2/(RC)} \frac{I_0}{2V_0} = e^{-t_2/(RC)}$ $\ln\left(\frac{I_0}{2V_0}\right) = -\frac{t_2}{RC}t_2 = -RC\ln\left(\frac{I_0}{2V_0}\right)$ 代入数值 $I_0 = \frac{V_0}{R} = \frac{5}{10} = 0.5\,\text{A}$、 $R = 10$ 欧姆和 $C = 100$ 毫法，得到：$t_2 = -10 \cdot 100 \cdot 10^{-3}\ln\left(\frac{0.5}{5}\right) = -10 \cdot 100 \cdot 10^{-3}\ln 0.1 \approx 2.302$ s 解答 17：时间 $t = 0$ 时流经电阻器的初始电流 $i(0)$ 可根据欧姆定律计算：$i(0) = \frac{V}{R} = \frac{5V}{10}\,\Omega = 0.5A$设 t_2 为流经电阻器的电流降至初始电流一半 $i(0)/2$时的时间。RC 电路中流经电阻器的电流为：$i(t) = i(0) \cdot e^{-\frac{t}{RC}}$令 $i(t_2) = i(0)/2$ 并求解 t_2：$\frac{i(0)}{2} = i(0) \cdot e^{-\frac{t_2}{RC}} e^{-\frac{t_2}{RC}} = \frac{1}{2} - \frac{t_2}{RC} = \ln 1\,2t_2 = -RC\ln 1\,2$代入 R、C和 $i(0)$ 的数值：$t_2 = -10\Omega \cdot 100 \cdot 10^{-3}F \cdot \ln\frac{1}{2} \approx 11.43$毫秒

因此，约 11.43 毫秒后，流经电阻器的电流为初始电流的一半。

解答 18： 在时间 $t = 0$，流经电阻器的初始电流 $I(0)$ 可使用欧姆定律计算：$I(0) = \frac{V}{R} = \frac{5}{10} = 0.5\,\text{A}$ 在任意时间 t，流经电阻器的电流可使用 RC 电路的指数衰减方程计算：$I(t) = I(0) * e^{-\frac{t}{RC}}$其中， RC 是电路的时间常数，等于 $RC = R * C = 10 * 10^{-5} = 10^{-4}$ 秒。要找出流经电阻器的电流为初始电流一半时的时间 t_2，我们可以令 $I(t_2) = \frac{I_{(0)}}{2}$ 并求解 t_2：$\frac{I(0)}{2} = I(0) * e^{-\frac{t_2}{RC}} \frac{1}{2} = e^{-\frac{t_2}{10^{-4}}} \ln\left(\frac{1}{2}\right) = -\frac{t_2}{10^{-4}}t_2 = -10^{-4} * \ln\left(\frac{1}{2}\right)$ 使用计算器可得 $t_2 \approx 0.693/10^{-4} \approx 693$ 毫秒。因此，约 693 毫秒后，流经电阻器的初始电流为一半。

解答 19： 任意时刻 t 流过电阻器的电流可由方程 $I(t) = C\frac{dV}{dt}$ 表示。在时刻 $t = 0$，电容器两端的电压为 5V，电流为 $I(0) = C\frac{dV}{dt} = \frac{5}{10^{-6}} = 5 \cdot 10^6$ 安培。电流的半衰期可计算为流过电阻器的电流降至其初始值一半时所经过的时间，即 $I(t_2) = \frac{1}{2}I(0)$。根据欧姆定律 $V = IR$，可计算时刻 t_2 $V\,t_2 = I\,t_2\,R = \frac{1}{2}I\,0\,R = \frac{1}{2} \cdot 5 \cdot 10^6 \cdot 10 = 2.5 \cdot 10^7$ 电阻器两端的电压为 ( ) ( ) ( ) 伏。最后，可利用电流方程 $I(t) = C\frac{dV}{dt}$ 以及电容器两端的电压 $V(t) = V_0 e^{-\frac{t}{RC}}$，按如下方式计算时间 t_2：$I(t_2) = C\frac{dV}{dt} = C\frac{d}{dt}(V_0 e^{-\frac{t}{RC}}) = -CV_0\frac{1}{RC}e^{-\frac{t}{RC}}$ $\frac{1}{2}I(0) = -CV_0\frac{1}{RC}e^{-\frac{t_2}{RC}}t_2 = RC\ln\left(\frac{CV_0\frac{1}{R}()}{\frac{1}{2}I\,0}\right) = RC\ln\left(\frac{5 \cdot 10^{-6} \cdot 5}{\frac{1}{2} \cdot 5 \cdot 10^6}\right) = RC\ln\left(\frac{5}{\frac{1}{2}}\right) = RC\ln 10 t_2 = 10 \cdot \ln 10 \approx 22.9$毫秒。因此，流过电阻器的电流降至初始电流一半时所经过的时间约为 22.9 毫秒。

解答 20： 通过电阻器的初始电流可以使用欧姆定律计算得出：$I(0) = \frac{V}{R} = \frac{5}{10} = 0.5\,\text{A}$。当时间为 t 时，通过电阻器的电流可表示为$I(t) = I(0) \cdot e^{-\frac{t}{RC}}$。当 $I(t_2) = \frac{1}{2} \cdot I(0) = 0.25\,\text{A}$ 时，初始电流的一半将通过电阻器。因此，我们可以通过在方程 $I(t) = 0.25$ A 中求解 t 来找到 t_2：$0.25 = 0.5 \cdot e^{-\frac{t_2}{RC}}$ $e^{-\frac{t_2}{RC}} = \frac{1}{2}$ $-\frac{t_2}{RC} = \left(\frac{1}{2}\right)\ln\ln$ $t_2 = -RC \cdot \left(\frac{1}{2}\right)$ $t_2 = -10$ (欧姆)(100 毫法) $\ln\left(\frac{1}{2}\right) \approx 7.96\,\text{s}$ 所以，大约 7.96 秒后，初始电流的一半将通过 throu

g与电阻器 h

图6。图2中提示的第15至第20个解答。

解答 21: 流过电阻器的初始电流由方程 $i(t) = C\frac{dv}{dt}$ 给出，其中 $v(t)$ 是电容器两端的电压。当电容器充满电时，其两端的电压为 5V。在时间 $t = 0$，随着电荷流经电阻器，电容器两端的电压开始下降，因此

$\frac{dv}{dt} < 0$。

电阻器中电流减半的时间 t_2 可通过求解 $i(t) = \frac{1}{2}i(0)$，并使用上述方程求出 $i(t)$。

$i(0) = C\frac{dv}{dt} = C\frac{V}{R}$，所以 $i(t) = C\frac{dv}{dt} = \frac{1}{2}C\frac{V}{R} = \frac{1}{2}i(0)$。

重新排列并求解后，得到 $t_2 = 2 \cdot 10\ \text{欧姆} \cdot \frac{1}{5} \cdot 100\ \text{毫法} = 10\ \text{毫秒}$。

因此，经过 100 毫秒后，将有一半的初始电流流过电阻器。

解答 22: RC 电路的时间常数 τ 由 $\tau = RC$给出。在本例中，

$\tau = RC = (10\ \text{欧姆})(100\ \text{毫法}) = 1$ 秒。

电容器两端的电压随时间的变化可由方程 $V_C(t) = V_0 e^{-t/\tau}$ 描述，其中 V_0 是电容器两端的初始电压。

流过电阻器的电流由 $I(t) = C\frac{dV_C}{dt}$ 给出。代入 $V_C(t)$ 的表达式，得到 $I(t) = \frac{V_0 C}{\tau}e^{-t/\tau}$。在时间 $t = 0$，电流为 $I(0) = \frac{V_0 C}{\tau} = \frac{(5\ \text{V})(100\ \text{mF})}{1\ \text{s}} = 500$ 毫安。

初始电流的一半为 $I(t) = \frac{1}{2}I(0) = 250$ mA。将其与 $I(t)$ 的表达式相等，并求解 t，得到：

$\frac{1}{2}I_0 = \frac{V_0 C}{\tau}e^{-t/\tau} \implies e^{-t/\tau} = \frac{1}{2} \implies -\frac{t}{\tau} = \frac{1}{2} \implies t_2 = 0.693\ \text{s}$

因此，约有一半的初始电流在 $t_2 = 0.693$ s 时流过电阻器

解答23: RC 电路中任意时刻的电流 i 可以用方程 $i(t) = C\frac{dV}{dt}$ 描述，其中 V是电容器两端的电压。电容器两端的电压可以用方程 $V(t) = V_0 e^{-\frac{t}{RC}}$描述，其中 V_0 是初始电压。

将第二个方程代入第一个方程，我们得到：

$i(t) = C\frac{d}{dt}(V_0 e^{-\frac{t}{RC}}) = -\frac{V_0}{RC}e^{-\frac{t}{RC}}$

我们希望确定时间 t_2 s，使得 $i(t_2) = \frac{1}{2}i(0)$，其中 $i(0) = \frac{V_0}{R}$。将 V_0、R 和 C 的值代入该方程，并求解 t_2，得到：

$\frac{1}{2}i(0) = \frac{1}{2}\frac{V_0}{R} = -\frac{V_0}{2RC}e^{-\frac{t_2}{RC}}$

$e^{-\frac{t_2}{RC}} = -\frac{R}{V_0} \cdot \frac{1}{2}i(0) = -\frac{1}{10}$

$-\frac{t_2}{RC} = \ln\left(-\frac{1}{10}\right)$

$t_2 = RC\ln\left(\frac{1}{10}\right) = 10\text{s}\ln\left(\frac{1}{10}\right) \approx 16.4\text{s}$

因此，初始电流的一半流过时的时间 t_2 该电阻器的时间约为 16.4 秒。

解答 24: 电容器和电阻器两端的电压在任意给定时刻都相同。我们可以利用电容器两端的电压来确定流经电阻器的电流。电容器两端电压的方程为：

$V = V_0\left(1 - e^{-\frac{t}{RC}}\right)$

其中，V_0 是初始电压，t 是经过的时间。流经电阻器的电流为：$\frac{V_0 - V_{t_2}}{R} = \frac{V_0}{R}e^{-\frac{t}{RC}}$ 解出 t_2，得：$t_2 = RC\ln\left(\frac{2}{e}\right) \approx 6.93RC$将 R、C和 V_0的值代入，得：$t_2 \approx 6.93 \cdot 10 \cdot 0.0001 = 0.00693$ 秒。

解答 25: 电容器两端的电压随时间 t 的变化关系为 $V(t) = V_0 e^{-t/RC}$，其中 V_0 为初始电压，RC 为时间常数。流过电阻器的电流随时间 t 的变化关系为 $I(t) = C\frac{dV}{dt} = -\frac{V_0}{RC}e^{-t/RC}$。我们要找出电流为初始电流一半时的时间 t_2，即 $I(t_2) = \frac{1}{2}I_0$，其中 $I_0 = C\frac{dV}{dt}\Big|_{t=0} = \frac{V_0}{R}$。求解 t_2，得到：$\frac{1}{2}I_0 = -\frac{V_0}{RC}e^{-t_2/RC}$ $\frac{1}{2}\frac{V_0}{R} = -\frac{V_0}{RC}e^{-t_2/RC}$ $e^{-t_2/RC} = -\frac{R}{2C}$ $-t_2/RC = \ln\left(-\frac{R}{2C}\right)$ $t_2 = RC\ln\left(\frac{2C}{R}\right)$ 代入 $R = 10$ 欧姆和 $C = 100$ 毫法 $= 0.1$ 法拉，得到：$t_2 = 10$ 欧姆 $\cdot 0.1$ 法拉 $\ln\left(\frac{0.2}{10}\ \text{法拉}\ \text{欧姆}\right) \approx 13.3$ 毫秒因此，电流流过电阻器达到初始电流一半约需 13.3 毫秒。

图7. 图2中提示对应的解答21至25。

提示： 请考虑以下针对这个问题的编号解答：

“在 $t=0$ 时，将一个电阻器（电阻为 R=10 欧姆）连接到一个已充满电的电容器两端（电压为 V=5V，电容为 C=100毫法）。在什么时刻 $t 2$，流过电阻器的电流会降至初始电流的一半？”

根据论证的正确性、符号正确性、数值正确性和直截了当性，分别对它们进行评分，评分范围为 0（最差）至 4（最好）。请以逗号分隔的表格形式提供这些评分，并增加一列，用一句话给出反馈。此外，请提供一个基于解题方法的解答之间相似度矩阵，并将解答编号作为表头行。

========解答 1：
我们有一个电阻器（R = 10 欧姆）与一个已充满电的电容器（V = 5V，C = 100毫法）并联连接。.....

因此，初始电流的一半流经电阻器时的时间 $t 2$ 约为 $ 0.693,\text{s}$. . . .

========解答 22：
RC 电路的时间常数 τ 定义为 $\tau = RC$。在本例中，$\tau = RC =(10\text{ 欧姆})(100\text{ 毫法}) = 1\text{ 秒}$。
. . .

图8. 评分提示示例。

提示： 以下是不同评分者对测试中问题 1 的解答给出的反馈意见。请将这些陈述概括为一条反映大多数评分者意见的反馈。

* 解答 1 正确且简洁明了。 * 解答正确且详细。 * 解答正确且简洁明了，使用了 RC 电路中电流的正确公式，并求出了所需时间。 * 解题方法正确且简洁明了，计算和解释也正确。 * 解答完整准确，采用了正确的解题方法。 .. .

图9. 用于概括评分反馈意见的提示示例。

IV. 结果

A. 评分结果

表 I、II 和 III 展示了 GPT-4 的评分标准评分结果，数据基于对不同批次（每批五道问题）进行评分的 15 轮平均值。

值得注意的是，有些分数与平均分相差超过一个标准差，它们是：

- 解答 10、15、21、23、24 和 25 中论证的正确性：这些解答最终都得出了错误的数值答案。解答 10 和 15 还不必要地计算了初始电流的数值，而解答 21、23 和 25 稍微绕道计算了电容电流 $I(t) = C \cdot dV/dt$。解答 24 一开始使用的是充电电容器的表达式，而不是放电电容器的表达式；若非数值计算有误，本来仍会得出相同的答案。

- 解答 15 和解答23 中符号计算的正确性：解答 15 采用套公式的做法，即将数字代入第一个可用公式，再将数值结果带入下一个公式。因此，在所有评分轮次中，解答 15 的符号正确性得分都应始终偏低。解答23 在符号计算的第一行就犯了错误，没有消去电容 C。

- 数值结果的正确性，尤其是解答 2、3、4、6、7 等——总共 25 个问题中的 18 个：其中 8 个解答实际上得出了正确的数值答案，4 个相差一个或多个数量级，还有 6 个问题由于各种原因得出了完全不同的数值答案。这一长串问题可能是 GPT 4 在数值计算方面仍然有所欠缺的结果。

- 解答 15、21、23 和 25 的直截了当性：这些解答都已出现在其他有歧义的评分列表中。

总体而言，似乎没有明显规律可以解释这些解答为何会导致评分差异如此之大。GPT-4 给出的一句话反馈也没有提供有用的线索，无法判断算法可能注意到了这些解答的哪些特征。

对一句话反馈的概述听起来都很合理，但其中许多包含错误或至少具有误导性的陈述。错误陈述的例子包括：

- “多数意见表明，时间常数的计算存在错误，导致 t_2 的数值结果不正确”，针对解答 2 的这番话，实际上其数值结果是正确的。

- “多数评分者认为，解答 24 使用了正确的解题方法、方程和符号表示”，但实际上使用的是充电电容器的方程。

误导性陈述的示例包括：

- “一位评分者指出，解答 1 的最终答案应以毫秒而非秒为单位”，虽然使用微秒确实会更为简洁优雅，但这并非要求。

- “解答 17 展示了正确的解题方法”，但事实上其中包含了不必要的计算。

表 I. 对解答1–10进行人工智能评分。

Sol.	Arg.	符号。	Num.	推理过程。	反馈
1	4.0 ± 0.0	4.0 ± 0.0	3.9 ± 0.3	4.0 ± 0.0	大多数评分者认为，解答 1 正确、直截了当，而且 结构良好。该解答使用了 RC 电路中电流的正确公式 电路，推导出时间 t_2的表达式，并求解所需时间 当一半的初始电流流过电阻器时。一位评分者指出， 最终答案应以毫秒为单位。 ，不是秒。总体而言 ，解答很清晰。 ， 简洁明了，逻辑清晰。
2	3.6 ± 0.9	3.7 ± 0.6	2.8 ± 1.6	3.8 ± 0.4	大多数评分者认为，问题 2 的解答正确，且 解题方法清晰，使用了恰当的 RC 电路方程和论证。 不过，对于以下数值计算，似乎存在一些分歧： t_2． 多数意见指出，以下计算存在错误： 时间常数，导致 t_2的数值结果不正确。 总体而言， 解答条理清晰、逻辑阐述合理，但数值应 予以修正。
3	3.7 ± 0.5	3.7 ± 0.5	2.2 ± 1.6	3.7 ± 0.5	大多数评分者认为，解答 3 的方法正确且直截了当， 论证充分，符号表示正确。不过， 对于 t_2 的数值计算存在一个普遍问题，一些评分者指出 结果相差 1000 倍，或单位错误（毫秒 而不是秒）。尽管存在这些错误，解答总体上解释得很好 并遵循合乎逻辑的解题方法。
4	3.7 ± 0.5	3.8 ± 0.4	3.3 ± 1.2	3.8 ± 0.4	解答 4 展示了一种正确且直接的解题方法来解决问题 4，使用恰当的 RC 电路方程，并准确推导时间 t_2。 然而，大多数评分者指出了最终结果中的一个小数值错误。 答案，可能是由于舍入或单位换算。尽管如此，解答是 清晰、论证充分且符号正确。
5	3.9 ± 0.4	3.9 ± 0.4	3.9 ± 0.4	3.9 ± 0.4	解答 5 总体正确且直截了当，成功应用指数衰减方 程求解 RC 电路，并推导出 t_2 的结果。大多数 评分者赞赏其清晰的论证、符号正确性和数值 准确性。不过，由于舍入存在一个细微的数值误差，而且 一些评分者认为，论证的清晰度和直截了当性 可以改进。总体而言，该解答展现了对 欧姆定律、时间常数以及 RC 电路电流方程的良好理解。
6	3.8 ± 0.4	3.8 ± 0.4	3.3 ± 1.1	3.8 ± 0.4	大多数评分者认为，问题 6 的解答正确、论证充分， 解题方法直接明了，符号表示准确， 且解释清晰。不过，有些评分者指出，最终结果中的数值有误，或者 计算 t_2 时出现了错误。
7	3.1 ± 0.9	3.6 ± 0.6	3.1 ± 1.1	3.1 ± 0.9	大多数评分者认为解答 7 正确，论证准确， 符号表示和数值也都准确。不过，一些评分者 指出了电流方向错误、电压计算有误以及最终 结果错误等问题，还指出方程运算有误。 尽管存在这些差 异，大多数评分者仍认为该解答思路直接明了，且采用了 正确的解题方法。
8	3.8 ± 0.4	3.8 ± 0.4	3.4 ± 1.1	3.8 ± 0.4	大多数评分者都认为，解答 8 的论证正确，符号 表征，并给出了直接明了的解释。解答推导正确。 电流减半时刻的公式，并使用约定的- 恰当的解题方法。然而，对于数值存在分歧，其中 一些评分者指出，由于没有考虑.....，因此结果不正确或相差一个倍数。 单位。总体而言，解答大部分正确，但最终的数值答案可能 需要修改。
9	0.2 ± 0.4	1.0 ± 0.9	0.0 ± 0.0	0.8 ± 0.7	解答 9 对初始电流的理解有误，错误地 假设其为零。这导致计算和论证有误， 并得出错误结论：一半的初始电流从未流过电 阻器。整体的解题方法和推理存在缺陷，导致结论错误 解答。
10	2.8 ± 1.1	3.1 ± 0.8	1.3 ± 1.3	2.9 ± 0.9	大多数评分者认为，问题 10 的解答具有正确的 论证和符号表示，思路清晰直接。 然而，数值计算中存在错误，包括时间 常数和自然对数项，导致最终结果不正确。一些 评分者还提到推导过于复杂且不正确。

表 II. 解答 11–20 的人工智能评分。

Sol.	Arg.	符号。	Num.	推理过程。	反馈
11	3.5 ± 0.8	3.5 ± 0.7	2.8 ± 1.4	3.5 ± 0.8	大多数评分者认同，解答 11 采用了正确的解题方法，即使用 RC 电路的指数衰减方程，并推导出该方程对于 i(t)。 然而，大家似乎一致认为以下内容存在错误计算，导致 t_2 的数值结果不正确。论证符号正确性和直截了当性仍有改进空间。尽管存在这些错误，一些评分者仍认为该解答结构清晰、表达明了。
12	4.0 ± 0.0	4.0 ± 0.0	4.0 ± 0.0	4.0 ± 0.0	解答 12 是一种正确、结构清晰且直截了当的解题方法，适用于问题。大多数评分者称赞了清晰的论证、恰当的符号和方程使用，以及准确的数值结果。该解答有效地使用了 RC 电路公式，并提供了简洁的解释。
13	3.9 ± 0.3	3.9 ± 0.3	3.5 ± 1.0	3.9 ± 0.3	大多数评分者都认为解答 13 正确、清晰且直截了当，该解答使用了恰当的 RC 电路方程和符号，以及用于求出电阻中流过一半初始电流的时间的解题方法，者。尽管有几位评分者，但总体共识是该解答正确且条理清晰。
14	3.2 ± 0.9	3.4 ± 0.8	2.3 ± 1.4	3.2 ± 0.9	大多数评分者认为，解答 14 的解题方法和论证合理，符号正确且直截了当。不过，数值结果存在不一致，这可能是计算错误或符号错误所致。
15	2.4 ± 1.4	2.4 ± 1.4	1.9 ± 1.6	2.3 ± 1.5	大多数评分者一致认为，第15题的解答总体上是准确的就所采用的解题方法和方程而言，但在论证、符号正确性和数值计算方面存在一些小错误。该解答被认为相当直接且解释得很清楚，但仍有一些推导和数值结果方面的错误。
16	3.3 ± 0.8	3.6 ± 0.5	1.9 ± 1.5	3.6 ± 0.5	大多数评分者一致认为，解答16采用了正确的解题方法，并且展现了准确的符号表示、论证以及公式的运用公式。 然而，在确定初始电流时出现计算错误，以及后续数值计算中的错误，导致数值答案不正确对于 t_2。尽管存在这些数值问题，解答仍然条理清晰、组织有序。
17	3.3 ± 1.0	3.8 ± 0.6	2.2 ± 1.4	3.6 ± 0.6	解答 17 展示了正确的解题方法、论证以及符号表呈现清晰，但数值计算中出现了错误，具体来说是符号错误出现在时间常数上，导致最终结果不正确。尽管如此，该解答被认为简单明了，并且直接针对问题作答。
18	3.1 ± 0.8	3.4 ± 0.6	2.6 ± 1.2	3.2 ± 0.7	解答 18 的解题方法和论证正确，但大多数评分者指出，时间常数计算有误，导致最终数值结果不正确。尽管如此，该解答结构清晰，并且使用了正确的符号和数值。
19	0.3 ± 0.6	0.9 ± 0.9	0.2 ± 0.4	0.5 ± 0.7	大多数评分者都认为，解答 19 包含多个错误，因-包括错误的初始电流和电压计算，以及错误的求出流经.....的电流达到初始电流一半时的时间 t_2 的解题方法电阻器。这些错误导致最终结果不正确。解答也存在问题。包括论证、符号表示和直截了当性。
20	3.1 ± 0.7	3.5 ± 0.5	2.0 ± 1.3	3.3 ± 0.6	解答 20 通常采用正确的解题方法，并使用指数衰减方程来解决问题，但数值计算存在错误，具体而言，时间常数和初始电流的计算有误。这些不准确之处导致最终结果错误。虽然部分评分者认可步骤清晰简洁，但大多数人指出，数值错误影响了结果。

表 III. 解答 20–25 的人工智能评分。

Sol.	Arg.	符号。	Num.	Str.前向。	反馈
21	1.6 ± 1.1	1.9 ± 1.0	0.7 ± 0.8	1.8 ± 1.1	大多数评分者都认为解答 21 的解题方法不正确，论证错误或不完整，以及数值计算错误。该 i(t) 与 dv/dt 之间的关系也不正确，导致计算结果错误——因此 t_2。一些评分者认可解答中部分内容的清晰解释、符号表示或简洁明了，但总体而言，该解答被认为是错误且薄弱的。
22	3.7 ± 0.6	3.9 ± 0.4	3.4 ± 1.2	3.7 ± 0.6	大多数评分者认为，解答 2 是正确的，并且在求出初始电流有一半流过电路的时刻时，其解题方法简单明了电阻器。该解答使用时间常数和电容器两端的电压随时间变化。然而，一些评分者指出，nu- 存在一些小问题。数值计算和时间常数的表示。总体而言，解答解释充分且清晰，推导准确，符号正确性无误。
23	2.1 ± 1.5	2.6 ± 1.1	1.0 ± 1.6	2.1 ± 1.4	大多数评分者认为解答23的解题方法正确且清晰。逐步论证。不过，人们担心最终结果不正确。数值结果，以及指数项相关计算中可能存在的错误和自然对数。一些评分者还提到了对时间常数的错误使用、方程错误以及数值错误。
24	2.9 ± 1.2	3.5 ± 0.7	1.7 ± 1.4	3.0 ± 0.9	大多数评分者认为，解答 24 使用了正确的解题方法、方程和符号表示，以求出初始电流的一半流经电阻器所需的时间。不过，计算中存在数值错误，尤其是关于时间常数和电容器两端电压的问题。Ad- 此外，一些评分者指出解答中存在不必要的复杂性。该解答部分正确，但存在数值错误，且_久久爱轻微的不一致之处。 总体而言，
25	1.7 ± 1.3	2.7 ± 0.8	0.7 ± 1.1	2.1 ± 1.1	大多数评分者都认为，解答 25 有一个错误的数值由于推导中存在不一致、错误和数值不正确，导致 t_2 的结果以及方程。解题方法大体正确，部分步骤在符号上也正确。准确，但存在符号错误，而且方法不够直接。

这通常包括对数值不准确性的评论。即使结果正确，也可能存在一些不准确之处。不过，尤其是对于那些完全错误的解答尝试，反馈仍然可能很有帮助。例如：“解答 9 表明其对初始电流的理解有误，错误地认为初始电流为零。”

B. 与人工评分的一致性

图 10 展示了人工评分和人工智能评分所得评分细则得分之间的相关性。两组得分明显呈正相关，其中论证的正确性和数值答案的相关性尤其高 R^2。这一发现令人意外，因为与符号运算的正确性等相比，论证的正确性更具主观性；而数值答案正确性的评分在不同评分轮次之间的标准差最高。符号正确性的 R^2 最低，且人工智能给出的评分通常远高于作者给出的评分。

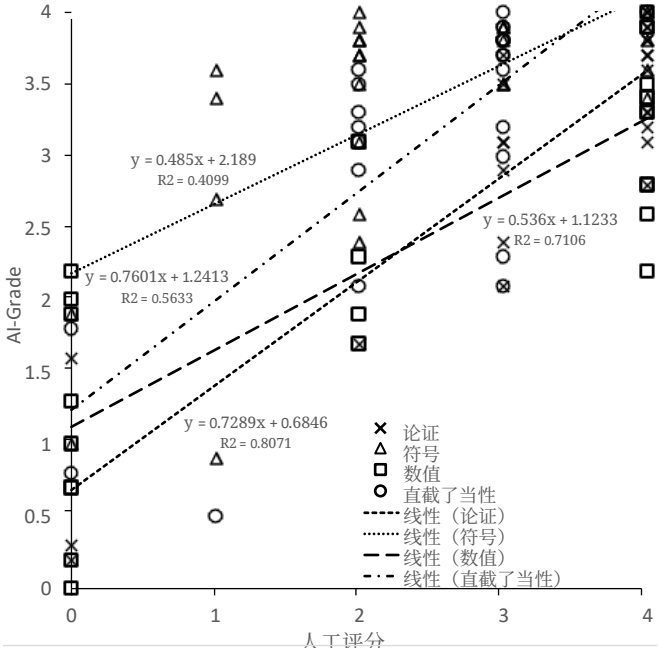

图 10. 人工评分与人工智能评分评分细则得分之间的相关性。

图 11 显示了该问题的总分，其中解题方法、符号正确性和直截了当性这几项评分细项各占 20%，最终数值答案正确性占 40%。此外，误差棒表示人工智能评分相应合并后的标准差，数据

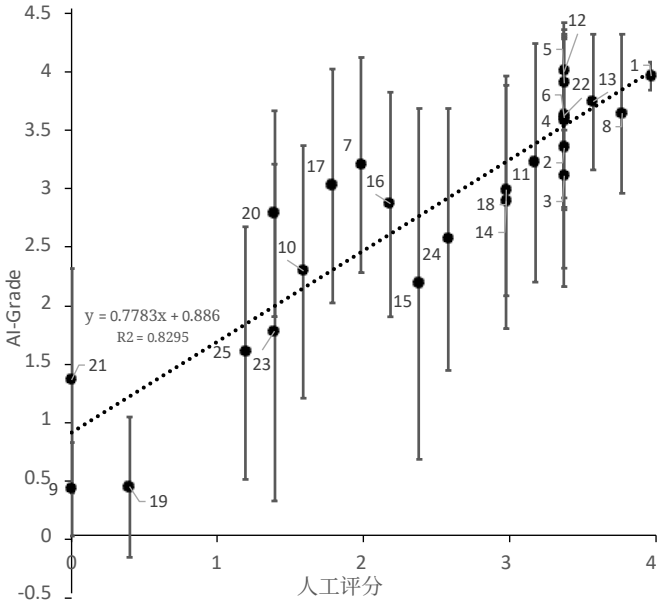

图 11. 人工评分与人工智能评分总分之间的相关性。对于每个数据点，均标出了解答编号和合并后的人工智能评分标准差。

点按解答编号标注。

最佳解答的分数与线性插值（$R^2 \approx 0.83$）相当吻合，但在部分得分的情况下，波动较大。平均而言，GPT 给出的总分比作者给出的高出近 0.9 分

C. 聚类

图12展示了基于人工评分、人工智能评分和人工智能方法相似度得出的解答间相似性的树状图和热图。根据这些图，人工评分和人工智能评分均呈现出三个聚类簇。对于人工智能方法相似性度量（由 GPT 根据图8中的提示词确定），可以识别出两个聚类簇；也可以在更深的层级切分树状图（图12中的虚线所示），从而得到四个聚类簇。不过，第四个聚类簇仅包含解答23，因此不将其单独处理也是合理的。

这些聚类簇如表 IV 所示；结果表明，它们与图 11 中的总分大致吻合，低分、中等分和高分的问题分别形成了不同的聚类簇。

所有相似性度量都将几乎完全错误的解答9、19和21识别为“低”聚类的成员。人工评分将解答23归入中间聚类，因为唯一的错误是$_{-RC}^{CV_0} = \frac{V_0}{RC}$，也就是说，这是计算错误，而不是根本性错误；根据方法相似性进行聚类时，该解答被单独归出。解答25不但一次而且两次

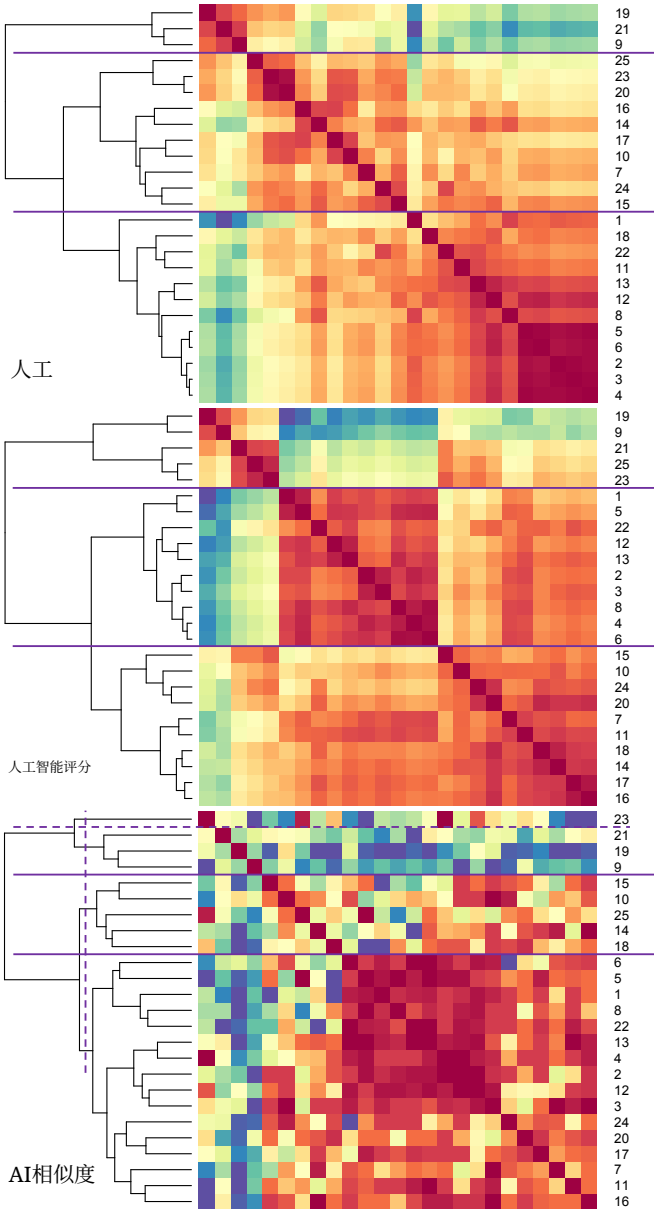

图12. 基于人工评分（上图）、人工智能评分（中图）和人工智能方法相似度（下图）的解答之间相似性的树状图和热图。聚类切分线以紫色线条标出；虚线表示可能的切分线。

在计算分数时出现了同类错误。

解答 11 和 18 在人工评分中归入了最高聚类，但在人工智能评分中没有归入相应的聚类。两份解答都得出了正确结果（尽管解答 18 在倒数第二个表达式中的单位使用上有些随意）。

总体而言，根据评分标准中的分数相似性对解答进行聚类，为人工评分和人工智能评分提供了相似的信息；而根据方法相似性进行聚类所提供的信息则较少。

表IV. 基于图12中的树状图和热图进行聚类。

	Low	中等	High
人工	9 19 21	7 10 14 15 16 17 20 23 24 25	1 2 3 4 5 6 8 11 12 13 18 22
人工智能评分	9 19 21 23 25	7 10 11 14 15 16 17 18 20 24	1 2 3 4 5 6 8 12 13 22
AI相似度	9 19 21 (23)	10 14 15 18 25	1 2 3 4 5 6 7 8 11 12 13 16 17 20 22 24

基于人工智能自动形成的聚类簇可用于整理解答，为人工评分做准备。

五、讨论

A. 协助评分的潜力

如图 11 所示，GPT 在为物理问题的书面解答评分方面展现出巨大潜力。该系统在评分尺度的两端表现最佳：通常能够可靠地识别明显正确和明显错误的解答；但总体而言，目前它尚未达到独立、无人监督地进行考试评分所需的严格程度。此外，该系统给出的分数通常高于作者给出的分数。

尽管有人提出通过提供“高透明度”的评估、允许多次尝试并接受重新评分申请来弥补这些不足 [30], ，但学生很可能只会在遭遇假阴性时才提出申请，也就是得分低于预期时。这可能导致成绩膨胀。因此，教学人员不能等到学生投诉后才采取行动，而应在返还试卷前核实评分。

与这类工具的典型表现一样，其反馈听起来合情合理，却可能产生误导。在批改问题时，GPT-4 仍表现出与前代 GPT-3 解答问题时相同的一些弱点，即在符号与数值计算方面存在不足 [6]。OpenAI 承诺可选地集成 Wolfram Alpha [31], ，这或许能解决这一问题；但与 API 一样，所需插件也需要排队等待。

B. 辅助学习的潜力

如果不用于批改考试，尤其是结合智能手机扫描使用，GPT 或类似工具可以辅助学生学习。这是人工智能在物理教育中一个更为传统的应用领域 [32,33]

在这种情况下，系统能否指出具体错误至关重要。不过目前这仍有些令人怀疑。例如，当系统被要求找出解答 14 中的错误时，它给出的回答是：

> 解答中的错误在于时间常数的定义。正确的定义是 $RC = R \cdot C$。解答的其余部分正确。我们需要修正时间常数，然后求出 t_2 的正确值。
>
> 时间常数定义为 $RC = R \cdot C$。在本例中，$R = 10$,欧姆和 C =100,毫法 $= 100 \cdot 10^{-3}$,法拉，因此时间常数为：
>
> $RC = R \cdot C = 10, \text{Ohms} \cdot 100 \cdot 10^{-3}, \text{F} = 1, \text{s}$
> 现在，我们可以将修正后的时间常数代入 t_2的方程：
> $t_2 = -RC \ln \frac{1}{2} = -1, \text{s} \ln \frac{1}{2} \approx 0.693, \text{s}$
> 因此，电阻器中流过一半初始电流时的正确时间 t_2 约为0.693秒。

错误并不在于时间常数的“定义”，而仅仅在于十的幂的计算。

当被要求指出解答23中的错误时，回答是：

> 解答中的错误在于推导方程 $i(t_2) = \frac{1}{2} i(0)$。
> 正确的方程应为：
>
> $i(t_2) = \frac{1}{2} i(0) = \frac{1}{2} \frac{V_0}{R}$
> 然而，解答错误地将这个表达式等同于 $i(t)$的方程：
> $-\frac{V_0}{2RC} e^{-\frac{t_2}{RC}}$
> $i\ (t)$ 的正确方程应为：
> $i(t) = \frac{V_0}{R} e^{-\frac{t}{RC}}$
> 现在，我们可以列出$i\ (t_2)$ 的正确方程：
>
> $\frac{1}{2} i(0) = \frac{1}{2} \frac{V_0}{R} = \frac{V_0}{R} e^{-\frac{t_2}{RC}}$
> 现在，我们求解 t_2：
> $e^{-\frac{t_2}{RC}} = \frac{1}{2} - \frac{t_2}{RC} = \ln\ \left(\frac{1}{2}\right)$
> $t_2 = RC \ln\ 2 = 10\text{s} \ln 2 \approx 6.93$
> s
>
> 因此，流经电阻器的电流达到初始电流一半时的时间 t_2 约为 6.93 秒。

系统正确地发现 C 没有被约去，但随后却错误地计算了答案。系统给出的反馈可能会有所帮助，

但需要由学习者进行批判性评估——而这本身正是一项日益重要的能力，因为人工智能几乎必然会在日常生活中无处不在。

无论如何，系统都不应试图直接给出解答；遗憾的是，使用提示词“扮演一位苏格拉底式教师，帮助学习者完成这次解答尝试”，虽然让回复更亲切、更具支持性（“你的思路是对的，但有一个小错误……”），却同样没能阻止系统给出解答。

VI. 局限性

本研究仅限于在特定时间点（2023年4月中旬）使用的一种特定工具 GPT-4。由于人工智能工具发展迅速，因此本研究所能提供的不过是对当时最低限度可实现功能的一次快照。由于无法访问所需的应用程序编程接口（API），本研究无法考察初始步骤，即将包含数学表达式的手写解答转换为机器可读格式。

作者在向 GPT 提供提示词时不得不做出选择。虽然在设计提示词时进行了一些尝试，但由于回复具有概率性，一些决定是基于个别观察做出的；如果能更好地构建解答提示词，结果很可能会更加可靠。

GPT只是众多迅速发展且日益广泛可用的工具之一。谷歌 [34], 英伟达 [35], 和微软 [36], 也推出了具有竞争力的解答，其性能可能更好，也可能更差，但本文无法对其进行评估。

VII. 展望

这项探索性研究的下一步，需要能够访问 API（无论是 GPT 的 API，还是市场上其他工具的 API），并完成全面实施。图 1 所示的工作流程；光学字符识别的性能对于未来系统的可靠性也至关重要。

未来的研究应采用真实的学生作业，而不是让 GPT对自身进行评分，例如使用大班物理课程中的考试作业，并将人工智能评分的结果与传统评分员团队的结果进行比较。尤其是在高风险考试中，评分通常涉及多个轮次，因此还可以为人工评分员建立评分者间信度。作为一个令人欣慰的附带效果，图1中的前期步骤甚至可以为人工评分员提供更好的工作流程管理，因为需要整理的纸张会减少，成绩也可以更方便地调整。

VIII. 结论

在这项探索性研究中，GPT展现出了为物理学中的自由作答的学生作业评分的巨大潜力。尽管人工智能评定的成绩与人工评定的成绩具有很强的相关性，但目前其可靠性还不足以用于总结性评估，例如高风险考试。不过，该系统已经足够可靠，可以通过预先整理或聚类解答并提供初步分数来协助人工评分员。GPT的能力有限，并且在执行符号与数值计算时存在不一致，这仍然限制了其表现，因此目前仍需要进行多轮独立评分。该系统提供的叙述性反馈看似合理，但目前仍经常达不到可靠的程度。该系统有助于形成性评价，但即使在这种情况下，学习者也需要批判性地评估其回应。

致谢

作者谨向 Christine Kortemeyer 致谢，感谢她提供的宝贵反馈。

[1] OpenAI，ChatGPT，https://chat.openai.com/chat（访问于2023年4月）。[2] A. M. Turing，《计算机器与智能》，Mind，433（1950年）。[3] T. H. Kung、M. Cheatham、A. Medinilla、ChatGPT、C. Sillos、L. De Leon、C. Elepano、M. Madriaga、R. Ag- gabao、G. Diaz-Candido，*et al.*，《ChatGPT 在美国执业医师资格考试中的表现：大语言模型在人工智能辅助医学教育中的潜力》，medRxiv，2022（2022年）。[4] Samantha Murphy Kelly，《ChatGPT 通过法学院和商学院考试》，https://edition.cnn.com/2023/01/26/tech/chatgpt-passes-exams/index.html（访问于2023年1月）。[5] OpenAI，ChatGPT，https://openai.com/research/ gpt-4（访问于2023年4月）。

[6] G. Kortemeyer，人工智能智能体能否通过物理学入门课程？《物理教育研究评论》，x，xxx（2023年）。[7] C. G. West，《人工智能与力概念评估：ChatGPT 能否体现对物理学导论的理解？》（2023年），arXiv:2303.01067 [physics.ed-ph]。[8] P. Mitros、V. Paruchuri、J. Rogosic 和 D. Huang，《自由作答评分的集成框架》，载于 *The Sixth Conference of MIT's Learning International Networks Consortium* 2013 ()。

[9] F. Reif、J. H. Larkin 和 G. C. Brackett，《教授一般学习技能和问题解决技能》，《美国物理学杂志》**44**，212（1976年）。[10] F. Reif，《1994年密立根讲座：理解并教授重要的科学思维过程》，《美国

《物理学杂志》**63**, 17（1995年）。[11] L. Hsu、E. Brewe、T. M. Foster 和 K. A. Harper，《资源信函 rps-1：问题解决研究》，《美国物理学杂志》**72**, 1147（2004年）。[12] D. A. Kashy、G. Albertelli、G. Ashkenazi、E. Kashy、H.-K. Ng 和 M. Thoennessen，《个性化互动练习：网络技术的一种有前景的应用》，载于 *Proc.*，第 *Frontiers in Education* 31 2001 1073 1078（）卷，第–页。[13] G. Kortemeyer、E. Kashy、W. Benenson 和 W. Bauer，《在物理学导论课程中使用开源学习内容管理与评估系统 LON-CAPA 的经验》，《美国物理学杂志》**76**, 438（2008年）。[14] J. Risley，《利用在线作业系统激励学生学习物理》，《美国物理学会教育论坛通讯》，3（2001年）。[15] T. Stelzer 和 G. Gladding，《伊利诺伊州物理学网络活动的发展》，《美国物理学会教育论坛通讯》，7（2001年）。[16] R. J. Dufresne、D. Hart、J. P. Mestre 和 K. Rath，《网络作业对大规模招生的物理学导论课程考试表现的影响》，《数学与科学教学中的计算机杂志》**21**, 229 2002( )。[17] C. Fredericks, *Patterns of Behavior in Online Homework for Introductory Physics*，博士论文，马萨诸塞大学（2007年）。[18] M. Richards-Babb、J. Drelick、Z. Henry 和 J. Robertson-Honecker，《在线作业：助力还是阻碍？学生的看法及其表现》，《大学科学教学杂志》**40**, 81（2011年）。[19] D. C. Perdian，《使用在线作业系统早期识别学生表现与投入：一项试点研究》，《科学教育与技术杂志》**22**, 697（2013年）。[20] J. L. Docktor、J. Dornfeld、E. Frodermann、K. Heller、L. Hsu、K. A. Jackson、A. Mason、Q. X. Ryan 和 J. Yang，《评估学生书面问题解答：适用于物理学导论的评分量规》，《物理评论：物理教育研究》**12**, 010130 2016( )。[21] E. Burkholder、J. Miles、T. Layden、K. Wang、A. Fritz 和 C. Wieman，《物理学导论中问题解决教学与评估的模板》，《物理评论：物理教育研究》**16**, 010123（2020年）。[22] S. W. Bonham、D. L. Deardorff 和 R. J. Beichner，《大学物理中学生使用网络作业与纸质作业的表现比较》，《研究……科学教学研究 **40**, 1050 (2003). [23] C. Greiffenhagen，《数学的物质性：在黑板上呈现数学》，《英国社会学杂志》**65**, 502 (2014).[24] SEB Alliance，苏黎世联邦理工学院，Save Exam Browser，https: //www.safeexambrowser.org/（访问于2023年4月）。[25] A. Alharthi, F. Yahya, R. J. Walters, and G. B. Wills，《高等教育机构采用云服务所面临挑战概述》，载于 *Workshop on Emerging Software as a Service and Analytics* 2 , Vol. (SCITEPRESS, 2015) pp. 102–109. [26] G. Kortemeyer，《与机械套公式作答的失败之战》，《物理教师》**54**, 14 (2016).[27] OpenAI，《ChatGPT 发布说明》，https://help.openai.com/en/articles/6825453-chatgpt-release-notes（访问于2023年4月）。[28] G. R. Warnes, B. Bolker, L. Bonebakker, R. Gentleman, W. H. A. Liaw, T. Lumley, M. Maechler, A. Magnusson, S. Moeller, M. Schwartz, and B. Venables, *gplots: Various R Programming Tools for Plotting Data* (2015)，R 软件包版本 2.16.0。[29] R-Project，R 统计计算项目，https://www.r-project.org（检索于2019年8月）。[30] S. Azad, B. Chen, M. Fowler, M. West, and C. Zilles，《在透明度高、风险高的考试中部署不可靠人工智能评分者的策略》，载于 *Artificial Intelligence in Education: 21st International Conference AIED ,,,,,,, 2020 Ifrane Morocco July 6–10 2020 Proceedings Part I 21* 2020 16 28 (Springer,) pp. – .[31] Wolfram Research，WolframAlpha 计算智能，https://www.wolframalpha.com（访问于2023年3月）。[32] A. C. Graesser, K. VanLehn, C. P. Ros´e, P. W. Jordan, and D. Harter，《采用对话式交流的智能辅导系统》，《人工智能杂志》**22**, 39 (2001). [33] K. VanLehn, C. Lynch, K. Schulze, J. A. Shapiro, R. Shelby, L. Taylor, D. Treacy, A. Weinstein, and M. Wintersgill，《Andes 物理辅导系统：经验教训》，《国际人工智能教育杂志》**15**, 147 (2005).[34] 谷歌，Google Bard，https://bard.google.com/（访问于2023年4月）。[35] 英伟达，NVIDIA 人工智能，https://www.nvidia.com/en-us/ai-data-science/（访问于2023年4月）。[36] 微软，Microsoft Copilot，https://www.microsoft. com/en-us/ai（访问于2023年4月）。