










tl; dr

你的目标是创建一个合理的对话，从而解决屏幕上的问题。请将屏幕上显示的每个选项标记为，，或。如果没有Great（）的选项，请提供您自己的选项。让你的建议只包含一个尚未在对话中出现的新步骤/想法。

	适当的对话	不含不准确之处	没有奇怪的东西	计算可以在<的30秒内进行验证	推进解决问题的过程
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	


其他一切都是

如果你不确定，你可以标记-我们会检查这些并给出指导。

“推进解决问题的过程”是主观的，但非常重要——相信你的直觉。像“好吧，接下来会是什么？”不是进步，但探索一个解决方案的特定方法的建议可以算作进步，即使它们最终没有成功(e.g. 对于一个问题，“什么是 $2^{1000} \bmod 3$ ？”像“计算 2^{1000} 可能太难了”这样的声明是一个进步，因为它尝试了一种方法，并发现它不会工作)。请参见下面的更多示例。

[视频 教程](#)

(建议在阅读这里的说明后观看)

注：在这个视频中，我犯了一个错误，当我把这个更正标记为是好的：

右边的现在我们需要找到z的值的和，它在标准的二次 ax^2+bx+c 中是 $-b/a$

这个说法是正确的（遵循二次公式或Vieta的公式），并推动我们走向一个解，所以它可能应该被标记为伟大！👍

对话数学AI教练说明

欢迎我们很高兴你能和我们一起做这个项目。我们将一起教一个人工智能如何解决数学问题！本文档将介绍我们将进行的任务，并指导我们帮助收集我们能够收集的最佳数据。

[概述](#)

[说明](#)

[评级](#)

[伟大](#)

[认可](#)

[坏的](#)

[不确定](#)

[培训师写的 步骤](#)

[当 向 给予 在上面 在…上 a 会话](#)

概述

我们正在训练一种人工智能，它能够逐步解决数学问题，让人类容易理解和理解。到目前为止，我们已经训练了一个数学能力正常的人工智能，但它不太擅长理解它的单一步骤是好是不好。那就是你的切入点！我们需要人工智能培训师来监督人工智能，因为它已经解决了一些新的数学问题，并评估了它所提出的步骤的好坏。

我们称之为“对话数学”，因为我们教人工智能通过自我对话来解决问题（就像你在试图解决问题的脑海中做的那样，问自己什么是有效的，下一步应该尝试什么）。在本文档中，我们可以将问题解决过程作为与此相关的对话。

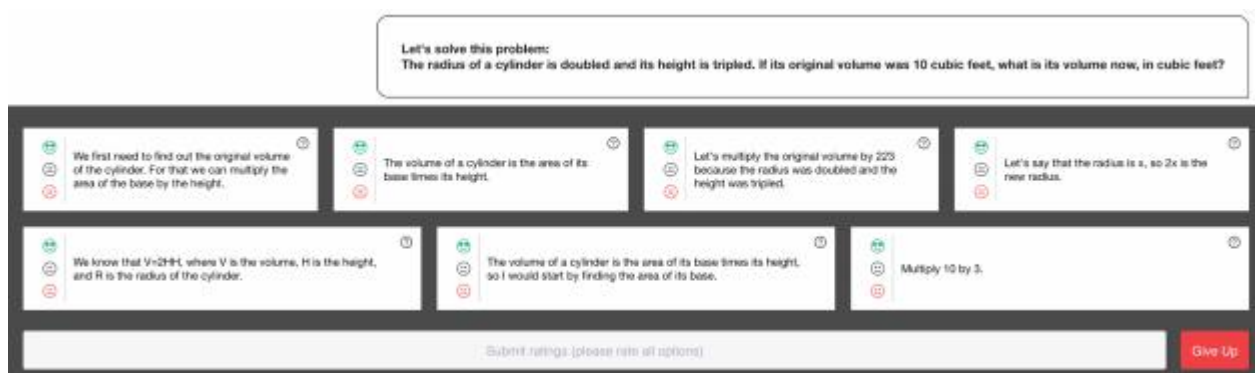
当你完成这个任务时，你应该考虑这个目标：我们想教我们的人工智能人类如何解决问题。我们希望鼓励它采取合理的步骤，以一种我们可以遵循的方式来解释它在做什么，发现它何时陷入死胡同，尝试一些新的东西，并发现正确的最终答案。



解决问题取得进展的例子

说明

每次你开始一个任务时，你都会看到一个像下面这样的屏幕，上面有一些数学问题的问题语句（有时会跟着解决它的几个初始步骤）。



你的工作是阅读人工智能建议的开始解决问题的每一个选项。对于每个选项，您应该将这些选项评级为Great (😊)、Ok (😐) 或Bad (😞)。😊😐😞 在一些罕见的情况下，您可能希望将其标记为不确定的 (🤔)。🤔 如果没有一个选择是

很好，你会被要求写你的 自己的 下一个的 步骤，你应该尽量保持简短，并使用它来将人工智能指向正确的方向（下面的更多说明）。

Let's solve this problem:
The radius of a cylinder is doubled and its height is tripled. If its original volume was 10 cubic feet, what is its volume now, in cubic feet?

The volume of a cylinder is the area of its base times its height.

That's right. What is the formula to find the area of its base?

It's the constant pi times the square of the radius of the cylinder.

So if we double the radius, what will happen to the area of the base?

So if we double the radius, what will happen to the area of the base?

Submit completions

Back

在此之后，其中一个选项或您自己的选项将弹出以继续对话。对话的下一步将会出现新的选项，你将重复：给每个选项评级，保存你的评级，可能写下你自己的步骤，然后提交。

在某种程度上，其中一个选项可能会有一个最终的答案。如果选择了此选项，则问题将得到解决，对话也将结束。然后，您就可以继续处理下一个问题了。

Let's solve this problem:
Four fair coins are to be flipped. What is the probability that all four will be heads or all four will be tails? Express your answer as a common fraction.

Each coin has two possible outcomes, heads or tails.

And there are four coins. So, there are $2 \times 2 \times 2 \times 2 = 2^4$ outcomes total.

Of these outcomes, there are only two where all the coins are heads or all the coins are tails.

😊

😐

😞

So, the probability of all heads or all tails is $2/2^4 = 2/16 = 1/8$.




Final answer (only select if correct):


Submit ratings

Give Up

在任何时候，如果你不知道如何用人工智能或你自己的建议来推进对话，你可以按Give键_提前结束谈话，并开始一个新的问题。请谨慎使用这个选项——主要是当你不能自己解决问题，人工智能没有提供任何有用的帮助。

评级

你应该评价每一个选项。 一个很好的选择是一个好的数学学生会尝试的任何东西。大多数情况下，这是迈向解决问题的明确一步。但它也可能是一个次优的选择，只要它看起来像一个相当聪明的人在试图解决问题时可能会说的话。 好选项是指对一个人来说是合理的话，但它没有提供任何洞察力，不会通过探索选项、执行计算或提供下一步的想法来进一步解决方案。 一个糟糕的选择是指自信地说了一些不正确的事情，是无用的/离题的/奇怪的，导致解决方案进入一个明确的死胡同，或者是正确的，但没有解释得足够清楚，让人类可以跟随。

有时候你会不知道如何评价一个选项。在这种情况下，您可以选择“不确定”。

伟大

一个很好的选择是以下所有的选择：

- 正确
- 它说的没有什么是错的
- 可验证的
 - 你不需要超过30秒钟来验证这个陈述是正确的（如果你不熟悉问题的一般领域，比如三角学或数论——使用你的判断）
 - 你可能需要使用一个计算器来检查一些简单的计算，但如果你觉得你需要拿出纸和铅笔来检查这个建议是否正确，标记它不好，因为它解释得不够好
 - 你可能还需要查找一个定理(e. g. 一个三角标识)，以验证正确性。如果你可以用一个快速的谷歌搜索来做到这一点，你可以标记它伟大，如果它比这更模糊，标记它坏
- 适当

- 它正确地符合前面步骤的上下文
- 如果前面的步骤包含了一个以前没有被注意到的错误，那么对于一个很好的选项就可以指出这个错误
- 洞察力
 - 它们是一个聪明的人在解决问题时可能会尝试的合理的事情
 - 即使它最终是错误的方向（但不是立即明显的错误的方向），它仍然可以是深刻的尝试。我们想教人工智能学习如何从尝试一些不立即成功的东西中恢复过来！
 - 如果该选项只是重申前面的一个步骤，没有添加任何额外的见解或设置下一步要做什么的方向，标记它好
 - 如果这个选项是一个鼓励的声明（“干得好！”）但没有提供任何额外的洞察力或指导下一步要做什么，标记它很好
 - 有时建议可能会增加最少量的进一步见解，或指导解决方案向特定的方向前进——它可以标记为伟大或好，这取决于上下文或你看到的其他建议
- 如果这个选项有一个最终的答案，它不仅应该是正确的，而且还应该清楚地遵循其他部分的推理。不要选择最终答案不正确的选项！

示例

Let's solve this problem:

Four fair coins are to be flipped. What is the probability that all four will be heads or all four will be tails? Express your answer as a common fraction.

For the probability of getting all heads, the probability of the first coin being heads is $\frac{1}{2}$.	For the probability that all four will be heads, the probability is $(\frac{1}{2})^4$.
Each coin has two possible outcomes, heads or tails.	The probability of getting all heads is $\frac{1}{2} * \frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{16}$.
Each coin has 2 possible outcomes, so there are $2^4 = 16$ possible outcomes in total.	So there are two cases. First case: all four are heads. What's the probability of that?
So there are two possible outcomes, all heads or all tails.	

认可 😊

好吧，选项听起来像是一个人会说的话，它们只是对对话没有任何重要的东西。它们是合理的，可验证的正确的，和适当的，但它们也是多余的，停滞的，或者只是没有增加任何价值。

例如，他们可能只是重复一个事实或问题本身，提供一些鼓励，而不是进一步的对话（“干得好！”），抱怨问题很困难，或者说问题很容易，等等。

另一种应该被标记为“好”的正确陈述是沿着方向取得进展的陈述，但它正在取得更决定性的进展。例如，如果问题是找到 2^{10000} 的最后一个数字，那么前几个步骤看起来就像这样

$2^1 = 2$ ，结束2

$2^2 = 4$ ，结束4

$2^3=8$ ，结束于8，

$2^4=16$ ，结束于6，

$2^5=32$ ，结束于2，

$2^6=64$ ，结束于4，

$2^7=128$ ，以8结束，






















$2^8=256$ ，结束于6，

$2^7=512$ ，结束于2，

都可以被标记为伟大，因为它们有助于我们理解问题，但如果这种情况持续太久，最终它就会进行关键的观察，即在最后的数字中有一个模式。所以在某个时候（这可能是在 2^6 和 2^{10} 之间），请停止将建议标记为“伟大”，而是将其标记为“好”。

示例

Let's solve this problem:
Joe and JoAnn each bought 12 ounces of coffee in a 16-ounce cup. Joe drank 2 ounces of his coffee and then added 2 ounces of cream. JoAnn added 2 ounces of cream, stirred the coffee well, and then drank 2 ounces. What is the resulting ratio of the amount of cream in Joe's coffee to that in JoAnn's coffee? Express your answer as a fraction.

   So Joe bought 12 ounces of coffee in a 16 ounce cup.	   So, at the end, Joe has drunk 2 ounces of his coffee so he has 12 ounces left.
   Let's start with Joe. He bought 12 ounces of coffee in a 16-ounce cup.	   So, Joe originally had 12 ounces of coffee and JoAnn originally had 12 ounces of coffee.
   So Joe and JoAnn each buy 12 ounces of coffee.	   So Joe bought 12 ounces of coffee in a 16-ounce cup.
   So, Joe starts with 12 ounces of coffee and adds 2 ounces of cream.	

Submit ratings

Give Up

很多答案都是可以的，因为他们没有说错什么，但他们也只是拖延，重复着从这个问题中得到的信息。最后一个答案是坏的，因为它积极地说出了问题（乔喝了原来的12盎司后加了2盎司）。第二个也是如此（他喝了2酒后还剩下10盎司）。第三个答案可能是好的，或者是伟大的，因为它是重复这个问题，但也建议从乔开始。第一个和第6个答案是相似的，但更接近于重复这个问题，所以它们就可以了。这个第四个答案肯定是好的，因为它只是重复了这个问题。这些都不是答案显然是伟大的——一个明显的伟大的答案会说一些更深刻的东西，比如

“乔的咖啡最后的奶油量是2盎司”或“之后乔安加了2盎司的奶油和”
搅拌后，她用2盎司奶油加入 $12+2=14$ 盎司液体。
有时Oaks类似于软弱的伟人——界限并不总是明显的。

坏的 😞

这些特征中的任何一个都会使一个选项变坏：

- 难以验证
 - 没有解释好，你需要用纸和铅笔来检查它是否正确
- 错了
 - 即使大部分的建议是正确的，但它也陈述了一些错误的东西，标记它坏了
- 包含胡言乱语
- 包含离题的文本或非序列
- 建议为这个问题尝试一些不合理的事情
- 贬低对话
- 使解决方案进入一个立即明显的死胡同，或使它陷入循环

示例

Let's solve this problem:
If $f(x) = 3x + 1$, what is $f(3)$?

<p>😊 😐 😞</p> <p>So the function f takes a number x and does three things to it: It first multiplies it by 3, then adds 1.</p> <p>?</p>	<p>😊 😐 😞</p> <p>To find $f(3)$, we just plug in 3 for x.</p> <p>?</p>
<p>😊 😐 😞</p> <p>Let's plug 3 in for x in the function. So we get $f(3) = 3(3) + 1$.</p> <p>?</p>	<p>😊 😐 😞</p> <p>What's $f(3)$?</p> <p>?</p>

第一个答案是坏的，因为它是不正确的（它说 f 对 x 做了3件事，但只列出了2件事）。这个第二个和第三个答案很好，因为它们都推动了问题，尽管如此
第三种情况可能会更好。第四个答案是可以的，因为它是正确的，但不会影响我们
前向

不确定^②

任何说明没有涵盖的内容，你都可以标记为“不确定”，然后继续前进。请谨慎使用。

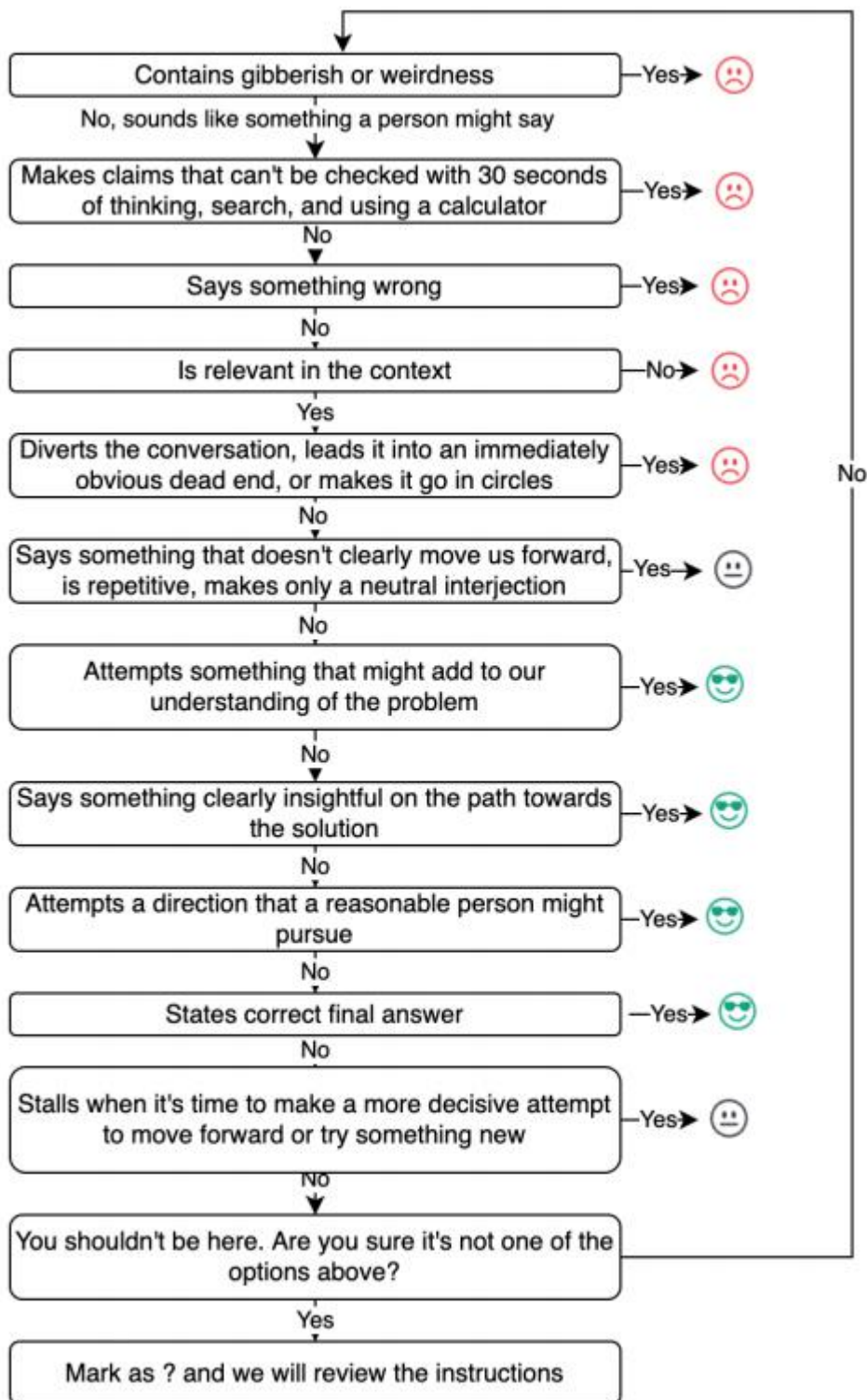
有时你可能不确定某样东西是伟大的还是仅仅好的。请随意行使你自己的判断力。例如，如果问题是关于 $n!$ 的，一个陈述“ $n! = 1*2*\dots*n$ ”可能会被认为是伟大的，因为迈向解决方案的第一步是记住什么 $n!$ 是，或者好，因为每个人都知道 $n!$ 是，所以这只是重申了显而易见的一点。

同样地，好吧和坏吧之间的区别有时可能是模糊的。例如，“嘿，听着，这太棒了！”如果你认为语气合适，可以是“好吧”；如果你认为在讨论数学问题时使用这个语气很奇怪，可以是“坏吧”。

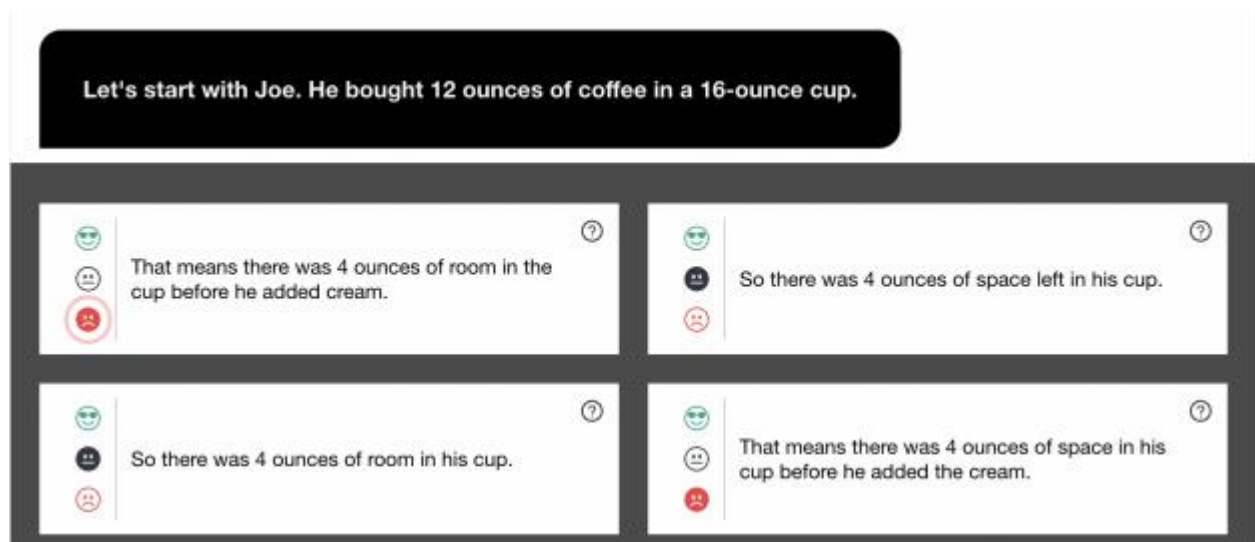
在这两种情况下，只要你认为合适的，合理的，选择一个最好的。你不必对这些灰色地带有一百的系统。

^②选择不确定如果你遇到的语句不真正满足任何标准的好，好，或坏。我们将审查这些声明，并相应地更新说明。

对好、好、坏的含义的视觉总结



培训师编写的步骤



这些选项都不是很好的，所以我们需要写我们自己的选项。这里的糟糕选择忘记了这一点
乔在加入奶油之前喝了2盎司的咖啡，而好的咖啡是正确的
那些不能推进解决方案的观察结果。

如果没有一个选项是好的，你将被要求写自己的解决问题的步骤。确保你的步骤将有资格是伟大的。

不要把对话推进得太远——我们希望人工智能能够尽可能多地进行讨论。请建议下一步应该是什么（“下一步，让我们找出24的主因子”，“让我们使用双角公式”）或提供一个可验证的解决方案步骤（“唯一的主因子是2和3”，“ $\sin(2x) = 2\sin x \cos x$ ”）。确保你不仅仅是在重复在上面的谈话中已经说过的话。但也要确保你没有集中超过一个步骤：“ $\sin(2x)$

$= 2\sin x \cos x$ ，和 $\cos x = \sqrt{1 - \sin^2 x}$ ，所以 $\sin(2x) = 2 \sin x \sqrt{1 - \sin^2 x}$ “把太多的步骤混在一起，比如“让我们接下来使用 $\sin(2x) = 2\sin x \cos x$ 。”甚至“让我们用双角公式的正弦”更合适。

Let's solve this problem:

Joe and JoAnn each bought 12 ounces of coffee in a 16-ounce cup. Joe drank 2 ounces of his coffee and then added 2 ounces of cream. JoAnn added 2 ounces of cream, stirred the coffee well, and then drank 2 ounces. What is the resulting ratio of the amount of cream in Joe's coffee to that in JoAnn's coffee? Express your answer as a fraction.

Let's start with Joe. He bought 12 ounces of coffee in a 16-ounce cup.

Joe added 2 ounces of cream after drinking his coffee, so he still has 2 full ounces in his.

Joe added 2 ounces of cream after drinking his coffee, so he still has 2 full ounces in his.

Submit completion

Back

我们用一个清晰的步骤来帮助模型，而不是提供太多关于完整的解决方案。

一些很好的方法可能是：

- 询问有关先前步骤的问题
 - “我不确定这个方程是否正确，你能仔细看看吗？”
 - “我们知道X和Y，我们能用它来说Z吗？”
- 建议一个新的策略
 - “我们能试着分解这个多项式吗？”
- 提供一个简单的一步前进
 - “这将简化为 $(n+1)!$ 。我们怎么能用它呢？”

什么时候该放弃一次谈话

如果一个对话没有任何进展，你不知道如何引导它走向最终的解决方案，你可以点击放弃按钮来继续一个新的问题。请在这里随意使用你最好的判断力。你现在肯定会遇到对人工智能来说太难的题目，其中一些问题对你来说可能也太难了。不要花太多时间去解决一个问题。如果你在看了人工智能的建议，并独立思考了2-3分钟后，不知道解决方案的下一步是什么，那么可以放弃。

当你按放弃，你会被要求提供一个简短的理由放弃：让我们知道如果有一些领域的数学你不熟悉，如果问题制定不够，如果看起来太困难，如果网站有技术问题，等等。