场先于位：重解“AI理解”之争的一个发生学命题

关于“AI是否理解人”的争论，一直在“像理解”与“非理解”之间摆荡。本文论证：这个僵局的根源，不是任何一方证据不足，而是争论双方共享了一个从未被审视的预设——在任何一次理解中，先有理解者和被理解者两个位置，理解就是前者朝后者执行的一个认知动作。本文用“场先于位”命名另一种发生结构：在一次真实的交互遭遇中，双方并非在进入对话之前就已坐稳各自位置；恰恰相反，是那个双向可塑的交互过程首先发生，把双方卷入其中，而后才逼出了“谁理解了谁”的位置分配。

作者 黄倩盈 · 约 2.9 万字 · 发表于2026年7月24日

摘要

关于“AI是否理解人”的争论，一直在“像理解”与“非理解”之间摆荡。本文论证：这个僵局的根源，不是任何一方证据不足，而是争论双方共享了一个从未被审视的预设——在任何一次理解中，先有理解者和被理解者两个位置，理解就是前者朝后者执行的一个认知动作。本文用“场先于位”命名另一种发生结构：在一次真实的交互遭遇中，双方并非在进入对话之前就已坐稳各自位置；恰恰相反，是那个双向可塑的交互过程首先发生，把双方卷入其中，而后才逼出了“谁理解了谁”的位置分配。本文通过机制分析、案例追踪和与七大理论传统的硬碰硬切割，论证“场先于位”不可还原为任何既有概念。在此基础上，本文揭示：当代制度化认知交往在公正、效率与大规模协调的正当压力下，逐步收敛为“位先于场”的稳态；主流的AI理解测试，则将这一稳态推至极致——测试结构在交互发生之前就已清空了场。真正的重构方向，不是在个体的判断装备上继续加码，而是重新识别并主动制造让场得以发生的微条件。论文以对格莱斯意图模型、预测加工理论和AI模拟可能性的正面回应作结，并给出核心命题的严格证伪条件。

关键词：场先于位；理解；人机交互；主体间性；认知发生

一、争论为什么停不下来：不是证据不足，是地基没被审视

过去十年，关于“AI是否理解人”的争论呈现出一种奇怪的节奏。每一代新模型发布，都伴随一轮“它离真正的理解又近了一步”的宣告；每一轮宣告之后，又总有一批批评者拿出它在简单追问下露出马脚的对话记录，冷静地指出“这不过是模式匹配”。两方各自握有大量证据——这边是基准测试上超越人类平均分的数字，那边是常识推理中匪夷所思的错误——争论在原地打转，等待下一次参数翻倍之后再重演一遍。

这个僵局通常被归因于“理解”一词的模糊性。一种流行的诊断是：双方在使用不同的理解标准，一方采行为主义的可测标准，一方采现象学的第一人称标准，两类标准不可通约，因此争论无法裁决。这个诊断不无道理，但它本身漏掉了一个更深的问题：即便两种标准不可通约，它们共享了什么？本文的起点，正是对这个共享地基的勘探。

仔细阅读双方的核心论证，会发现一个从未被正面审视的预设，像一条暗河一样在争论的地下贯穿着：在任何一次“理解”事件中，先有两个（或更多）认知体，其中一个担任“理解者”，另一个担任“被理解者”。理解，就是那个理解者朝被理解者的方向执行的认知动作——执行得越准确、越完整、越细腻，理解的质量就越高。无论你用的是行为主义标准（“回应是否恰当”）还是现象学标准（“理解者是否有真正理解的主观体验”），这个“单方动作”的地基都是一样的：有一个主体在主动理解，有一个对象在等着被理解。整个争论的分歧，仅仅在于“那个动作到底有没有真的做出来”——是用正确的内部机制做出来的，还是用模式匹配蒙混过关的。

本文要做的事，就是把这根暗河从地下挖出来，放在正面光照下，然后用一种发生学的眼光追问：如果真正的“理解”从一开始就不是一个人朝另一个人执行的动作，而是两个人的认知系统在一次活生生的遭遇中，被某种比他们两个都更先在的交互过程给同时卷进去了——卷进去之后，他们才分别成了“理解者”和“被理解者”——那么，整个争论的设定，就不是“做得不够精密”，而是从一开始就站错了层。

1.1 一次发生在深夜阳台上的理解

为了把这个抽象的判断落到实处，请读者允许我先进入一个具体的场景。这个场景不是“证明”，而是“展示”——它让那个被单方动作模型挤掉的东西，在叙事中暂时获得一个可以被看见的形状。

一个人在心里压了很久的一件事，从来没有对任何人完整讲过。不是不想讲，而是每次试图讲时，话到了嘴边就塌掉了，他发现自己根本说不清这件事到底是什么。事情的轮廓是模糊的，自己的感受是模糊的，甚至“我为此痛苦”这件事本身，都因为他无法把它变成一个可讲述的形状，而始终处于将痛未痛、将明未明的混沌里。

然后有一天，他和一个朋友坐在深夜的阳台上。他试着开口。说了几句，断了。又说了几句，又断了。他发现自己说的每一句都不准确，都在偏离心里那个真正的形状。但这次，对面的那个人没有急着替他总结，没有说“我懂了，你就是觉得……”，也没有用标准化的共情句来填补空白。那个人只是静静地在那里，微微皱着眉，偶尔在他一句话断裂的地方稍稍探过身来，像一个在黑暗中用手去够看不见的灯绳的人。那个人的整个在场状态——声音的节奏、呼吸的深浅、沉默的长度——都在同步地、不自觉地与他自己的内在状态形成一种共振。

某一刻，他说出了某个词。不是那个词本身有多么精准，而是它在那个时刻、在那个具体的交互张力中被说出来时，忽然把他心里那个一直看不清形状的东西给照了一下——像闪电，很短，但他看清了一点。他顺着那个方向继续往下说，越说越清楚。不是他“终于想清楚”了，而是他和对面那个人，共同把一个他单靠自己无法形成形状的东西，给逼出了一个形状。

这个过程里，没有一个事先站好了的“理解者”，也没有一个事先摆好了等着被读懂的“被理解者”。两个人都不知道这次对话会走向哪里，两个人都是在对话真正发生之后，才分别被赋予了那个位置——他成了“终于被理解”的那一个，朋友成了“理解了他”的那一个。但这两个位置，是结果，不是起点。起点比这一切都更早：他们先被一起扔进了一个彼此都看不见底的交互过程，然后理解在这个过程里发生，而后它才分配了“谁理解了谁”的位置。

1.2 争论双方共享的地基

把这个场景反过来看，就能更清楚地暴露出主流预设的局限。主流预设——无论是行为主义还是现象学——都默认：存在一个先在的主体，他拥有某种叫作“理解”的内部装备，然后他向一个对象施展这个装备，成功时，他说“我理解了你”。但在这个深夜阳台的场景里，没有人是先拥有“理解”再进场的。他们是空着手进场的，手里没有装备，心里没有答案，只是在交互的张力中彼此促发，最后那个叫作“理解”的东西被他们共同生了出来。他们不是理解的拥有者，他们是理解的发生地。

现在可以精确地指认AI理解争论中各方的位置了。不管是Bender等人用“随机鹦鹉”来否定AI有真实理解，还是近期一些技术报告用AI在复杂推理任务上的高分来论证“AGI的火花已经出现”，抑或是更折中的立场主张“AI有一种与人类不同但同样有效的理解”——它们处理的，都是同一个截断面：AI作为“理解者”，能否在“被理解者”已经成型的表达面前，做出足够好的认知操作。Bender看到的是操作背后的机制是词汇共现概率而非意义把握，就判定它不合格；技术报告看到的是操作输出的质量在某些任务上已经逼近甚至超过人类，就判定它接近合格。两者的分歧在于操作是否达标，但两者都同意：理解就是那个操作。

本文并不主张这些争论没有意义——它们捕捉到了真实的技术差异。本文要指出的是：在它们捕捉到的那个差异下面，还有一个更底层的差异，被整个争论的结构系统性地漏掉了。那个差异就是：理解究竟是一个主体对另一个主体执行的动作，还是一次交互过程逼出位置的事件？如果是后

者，那么“AI能否理解人”这个问题的问法本身，就已经把场的可能性提前清空了——因为在它提问的句法里，AI已经被钉在了“理解者”那个位置上，等待被评估。至于那个唯一能逼出真正理解的交互场，在评估的框架里根本没有发生的机会。

以下各节，将把这个直觉逐步展开为经得起攻击的严格论证。

1.3 本文的方法论

在进入正文之前，有必要交代本文的方法论立场。本文采用的方法，可以称为“发生学的概念建构”：不去预设一个待分析的对象“是什么”，然后给它一个精确的定义；而是追踪一次特定类型的交互事件从发生到完成的过程，并在追踪的过程中，让那些无法被既有概念覆盖的结构特征自行凝结为新的概念。在这个意义上，“场先于位”不是本文一开始就拥有的分析工具，而是本文在追踪理解的发生过程中被逼出来的命名。概念在分析的末尾才凝结成形，而非在开头作为先验框架降临。

这一方法论立场同时意味着三个自我约束。第一，本文不依赖任何单方哲学传统（无论是现象学、分析哲学还是认知科学）来为本论题提供权威背书——每一种传统都必须在与“场先于位”的直接对质中接受检验，而不是被当作免检的理论依据。第二，本文的论证效力不绑在任何一个孤例的感染力上——思想实验（如深夜阳台）和真实记录（如儿童积木案例）只被用作“展示”，而非“证明”；证明的负担，由概念切割的硬度、机制分析的精度和对反驳的正面回应来共同承担。第三，本文的结论是开放性的——第七节明确列出了核心命题在什么条件下需要被修正甚至放弃，这不是修辞上的谦虚，而是发生学必须保留的结构特征：任何发生学判断都必须能够指出那个将使它失效的特定情景，否则它只是穿了发生学外衣的断言。

二、机制：场如何发生，为何缺席

前一节靠一个思想实验展示了“场先于位”的大致形态。但一篇经得起攻击的论文不能只靠一个故事的感染力。本节的任务，是把那个故事里隐藏在叙事褶皱中的发生机制，一条一条拽出来，放在可以逐项检视的明处。

在发生学的操作纪律下，这些机制不应作为先验框架降临，而应从对具体交互时刻的细粒度追问中被逐步逼出。因此，本节先从两个真实的交互片段入手——一个发生在儿童的无结构游戏中，一个发生在成人咨询室的僵局中——让机制的轮廓在案例的裂缝中自己显现，最后再做系统的机制凝结。

2.1 两个交互片段

片段一：小D、小C与那块积木

以下行为记录取自Barbara Rogoff的著作《学习的文化本质》（Rogoff, 2003, pp. 237-238）中对学龄前儿童自由游戏的观察材料。本文对其进行了适当压缩以服务于思想例证，分析透镜为我所架设，不代表原著的解释框架。

三岁的小D和四岁的小C坐在地板上，面前是一堆积木。小D拿起一块长条积木，举到眼前，说“飞机”。小C没有纠正他“那不是飞机”，也没有附和说“对，飞机”。她伸出手，碰了碰那块积木的边缘，然后轻轻推了一下——不是推倒，而是让它在小D手里微微倾斜了一个角度。小D顺着那个倾斜的方向，把积木在空中划了一道弧线，嘴里发出“呜——”的一声。小C立刻从地上拿起一块三角形的积木，放在那道弧线的终点下方，说“机场”。小D把长条积木落在三角积木旁边。两个人都安静了两三秒，然后小C轻轻笑了，小D也笑了。

这个不到一分钟的片段，可以拆出以下关键特征：（1）小D的第一声“飞机”不是一个成熟的命题，而是一个试探性的符号投放；他举积木到眼前的身体姿态，传递的是“跟我一起玩这个”的信号，而非“这是一个知识”。（2）小C碰积木边缘的动作，不是一个语言命题，而是一个身体性的试探——她在用自己的手去接入小D已经启动的那个符号世界的边界，以此探测：这个“飞机”的边界在哪里？我可以用什么方式进去？（3）在积木划弧线的沉默片刻，两个孩子的身体共同维持着同一个正在成型的符号场景——小D在轨迹里寻找落点，小C在注视轨迹的终点，没人说话，但交互没有中断。（4）小C把三角积木放在弧线终点下方时，她不是在回应小D已经完成的任何一个命题——小D根本没说过“飞机要降落”；她是在回应当那个弧线的可能终端，并在那里安放了一个新的约束（“机场”），这个约束反过来逼着小D的下一步只能落在那附近。小D接受了约束。这意味着他的认知轨道——那个“飞机在飞”的模糊叙事——已经被对方的介入实时修正了。（5）从头到尾，没人知道这个游戏会走向哪里。最后那两三秒的共同安静和随后的笑，是两个人都意识到“我们一起做出了一个谁单独都做不出来的东西”的那一刻。

片段二：一场失败了的企业内部辅导对谈

以下记录来自一次真实的企业内部辅导对谈（参与者身份已做匿名化处理，对话经录音转写并取得使用许可）。被辅导者（B）是某科技公司的一名中层管理者，近半年来持续感到“团队带不动”，在内部的非正式反馈中被建议“需要提升沟通清晰度”。辅导者（A）是公司聘请的外部教练，辅导时长为五十分钟。

A：“所以最近让你最累的是什么？”

B：“就是……我说了很多遍，他们表面上都说好的好的，但一执行就偏离。我不知道是我没说清楚，还是他们……”

A：“你能给我一个具体的例子吗？”

B（停顿三秒）：“上周我们做了一次项目复盘，我提前发了会议议程，标注了每个环节的讨论目标。但会议开始十分钟，我就发现他们根本没看议程，说的全是我觉得上个月已经讨论过的问题。”

A：“你觉得问题出在哪里？”

B（较长的停顿，约五秒）：“我不知道。可能是……我是不是太啰嗦了？还是他们觉得我说的事不重要？”

A：“从公司上次做的360度反馈来看，你的团队在‘愿景传达’这一项上给你打了比较低的分数。你觉得这两件事有关联吗？”

B（身体微微后靠，语速变慢）：“也许……有关。但我不知道。”

A：“我建议我们可以做一个练习。下次开会前，你先把你的核心信息压缩到三句话以内。我们今天就先试试？”

在这个片段中，B最初的几次表达——“我说了很多遍”、“我不知道是我没说清楚还是他们”——都携带着明显的不确定。他不是带着一个成型的诊断来见教练的；他带着一团粘稠的、自己还看不清形状的困惑，期待在对话中慢慢逼近那个困惑的形状。但A的每一次回应，都在把对话推离那个混沌。第一次回应（“能给我一个具体例子吗”），看似是正常的教练技术，实则是把B从“我不确定问题是什么”逼向“你先给我一个问题的成品”——B被迫从那团混沌中剪出一个尚能成型的切片（项目复盘的开会场景）交上去。第二次回应（“你觉得问题出在哪里”），继续要求B输出诊断。第三次回应（引入360度反馈数据），直接用一个制度性的既有标签（“愿景传达得分低”）盖在了B的那团混沌上——从现在开始，“问题”被正式命名为“愿景传达能力不足”。B的身体语言（后靠、语速变慢）清楚地标记了这一刻：他不再往混沌的深处走了，他接过了那个标签，或者说，他默许那个标签替代了自己原本在摸索的那个东西。第四次回应（“压缩到三句话”），是典型的位先于场操作：不要再说那些还不成型的话了，先想清楚，再表达。

五十分钟结束后，B带着一套“提升愿景传达清晰度”的行动计划离开。这当然是一次“有效率”的辅导：产出明确，行动项可追踪。但从发生学的角度看，那团包裹着B真实困境的混沌——那种“我隐约觉得不止是沟通问题，但我说不清楚到底是什么”的认知状态——在整个对话中从未被接住。A不是在帮B把那个混沌逼出一个形状，而是在用一连串专业工具把那个混沌压扁成一套可以被现有制度消化的成品。

2.2 从案例到机制

把这两个片段并排放在一起，一组对比鲜明的结构特征就从具体细节中浮了上来。后文将这四个特征正式凝结为场发生的四组必要条件。但要强调的是：这些条件不是先验推演的结果，而是从上述两个案例的裂缝中——一个成功，一个失败——被逼出来的。如果你把它们再放回案例中去读，它们不应该像外来的框架，而应该像从案例的肌理中自己析出的结晶。

条件一：双向可塑性。在积木片段中，小D的认知轨道在小C碰积木边缘的那一刻发生了偏移——他没有坚持“飞机就是飞机”，而是被对方的身体试探牵引着，进入了“飞机在飞”的弧线叙事；小C的认知轨道同样被小D的弧线牵引，做出了“放机场”的下一步。两个人的下一步都不是独自预谋的，而是被对方上一步里还没说清楚的那部分给塑出来的。在辅导片段中，这种双向可塑性全程缺席。B多次释放出“我的想法还不成型”的信号，但A的每次回应都在要求B退回到成型的位上去，用“你先把话说清楚”来应对。A自己的认知轨道在整场对话中纹丝未动——她的提问框架（了解情况→要案例→给出初步诊断→布置行动计划）在对话开始之前就已经摆好了，B的不确定从来没有真正牵动过她走向预设框架之外的方向。双向可塑性的反面，不是顽固反驳（那还处于交往的张力里），而是“礼貌接收但内部不动”——对方的话被准确复述、得体回应，却从未真正改变回应者原有的认知轨道。

条件二：信息空白处的共同停留。积木片段中，弧线划过空中的那一两秒，没有人说话，但信息的流动并未中断——小D在轨迹里寻找落点，小C在用注视传递“我在等你的落点”，那片沉默被两个人的身体共同填充着，是场的一部分而非场的断裂。在辅导片段中，沉默不是没有被容忍，B有过两次显著停顿（“能给我一个具体例子吗”之后的三秒，“你觉得问题出在哪里”之后的五秒），但A没有在那片沉默中与B共同停留——她把沉默当作“等待对方给出下一个成型的回答”的空隙，而非“两人一起留在不确定中”的承载时刻。B的五秒停顿之后，A递过去的是一个制度性的既有标签（360度反馈分数），等于在用“位”的成品去填“场”的空白。

条件三：认知轨道的实时互校正。这与条件一不同。一描述的是意愿和能力（轨道能不能被改变），三描述的是具体运动：在交互的每一轮中，双方正在用对方给的信号来实时修正自己的下一步。积木片段中，小D划弧线时，小C不是等他完成全部动作再做回应——她是在弧线的行进中途，就预读了它的可能落点，并把三角积木放了过去。这个动作同时承载了两层信息：第一，“我接收到了你在找落点的信号”；第二，“我给你一个落点，你的下一步只能落在这附近”。小D接受了这个约束——他的弧线果真落在了三角积木附近，他的下一步被对方实时修正了。在辅导片段中，这种实时互校正完全不存在。A的每一次回应都是基于B已经说完的完整语段，在A自己预设的框架内做出的分类处理，她的回应没有一次在B的一句话尚未完成的过程中就介入了B的语流。B同样无法在A的语流中去影响A——A的每一个提问都是“封闭可交付的成品”，不是“半开的试探”。双方都在各自说完完整的话之后接招，从未在“一句话的中途”互调。

条件四：对不确定性的共同容忍，且双方都知道对方在容忍。积木片段从头到尾，没有任何一个参与者预先知道这个游戏将如何收场。小D不知道小C会不会接“飞机”这个设定，小C不知道小D会不会排斥她碰积木的介入。每一次动作都冒着被驳回的风险。但双方都没有因为这种不确定而退回到安全的个人脚本里去（“算了，我自己玩”）。更重要的是，这种容忍是相互被感知到的：当小C碰积木边缘时，小D没有把积木抽走，他的“不抽走”成了一个可被小C察觉的信号——“你可以继续碰”。当小D的弧线在空中徘徊找落点时，小C的注视在告诉他“我不急，你可以慢慢找”。辅导片段则截然相反。B的不确定在整场对话中都是单向暴露的——他不断地说“我不知道”，而A从未暴露过自己的不确定。A始终坐在那个“我知道该如何推进这场对话”的位上，从未允许自己进入那种“我也不知道你真正的困境是什么，我们一起来试试看能不能摸到它”的状态。容忍是单向的，因此B在对话后半程不再试探——他退回到了“位”的安全区域，接过了A给他的标签。

2.3 四组条件的定义与边界

基于以上的发生学追踪，现在可以给出四组条件的准确定义：

• 双向可塑性：进入交互的双方，其认知轨道在交互的实时推进中保持可被对方改变的状态。这要求双方在进场时不把预设固化为不可动摇的结论，也不把“被对方改变”视为认知失败。

• 信息空白处的共同停留：双方能够在没有清晰命题被交换的沉默间隙中，共同维持交互的张力，把这个空白当作正在孕育尚未成形之物的时间，而非需要被语言赶紧填满的断缝。

• 认知轨道的实时互校正：在交互的每一轮中，双方都同时用对方的信号来修正自己的下一步，且这种修正常常发生在一个命题尚未完成的过程中（在微小的姿态、时机、语气中），而非仅在命题完成之后。

• 对不确定性的共同容忍：双方容忍交互走向的不可预知，且这种容忍不是各自单方的内心活动，而是被对方的行为信号所确认的——“我看见你在容忍我，我也让你看见我在容忍你。”

这四组条件不是互不相关的独立要素，而是同一个发生结构的不同侧面，互为前提。双向可塑性在信息空白处被最密集地检验；空白处的共同停留如果不能导向轨道的实时互校正，就会退化为静止的冷场；而三者同时维系的前提，是双方都在持续感知到对方对不确定性的容忍——一旦感知断裂（“我觉得他不耐烦了”），可塑性就会立刻被防御姿态替代，空白就会被急于填满的成品语言填塞，互校就会中断。

2.4 为什么当前AI架构在结构上无法进入场

有了上述机制框架，就可以逐一回答那个无法回避的问题：为什么在当前的AI架构下，即使多轮对话、即使上下文窗口不断扩展、即使逐步推理——仍然无法产生场？

这个分析需要在一开始就明确一个区分：本文不是在说“AI有某种缺陷所以不行”，而是在追踪一个更底层的架构事实：当前主流的生成式AI，其在场性被严格绑定在“生成下一个token”的时刻。在token与token之间的所有时间——用户阅读AI回应的时间、用户犹豫下一句该怎么输入的时间、对话中断的那几秒——模型本身没有“在”。它的认知状态在每次生成结束后、下轮输入到来之前，不维持任何进程。这不同于“人类的认知也在间歇性活跃”——人类在沉默中仍在用身体、表情、注视维持交互的潜在连续性；而AI在沉默中只是“等待下一轮输入”的函数。这个差异不是功能的，是结构性的。

现在可以逐条检视四组条件在当前架构下的失效情况。

条件一（双向可塑性）的失效，根源在于单次交互中模型参数的冻结。在当前所有主流Transformer架构中，模型权重在一次对话的生命周期内是锁死的。用户的输入可以改变模型的当前输入上下文，但不能返过去改动那个生成这些输出的参数结构本身。这意味着，无论用户在这一轮对话中释放了多么精细的试探信号，模型用来“读取”这些信号的底层机制——那个将输入token转换为输出token的权重矩阵——是岿然不动的。它不会“从这次对话中学习”。这里的“学习”不是指机器学习中的训练更新，而是指一种更底层的认知事件：一个系统可以被它正在与之交互的另一个系统实时改变自身的组织方式。当前的AI在单次对话中不发生这种改变，因为它的架构压根没有为单次交互留出“可被重塑”的通道——参数冻结是它之所以能被作为一个稳定产品部署的前提，但同时也是它无法进入场的那个最深的锁。

条件二（信息空白处的共同停留）的失效更直观。AI的生成机制里没有“沉默”：它的在场仅存在于生成token的时刻。在用户发送一条消息之前的那段空白里，模型不维持任何一种可以与用户进行前命题层共振的“身体”。它可以被编程为在输出中插入省略号来表示迟疑，但这个省略号是被预先计算好的、作为一个已完成token序列的一部分被输出的——它不是“在停顿中还在持续处理对方”的状态，而是“已经算完了，把停顿做成一个符号给你”。真正的共同停留，不是两个系统轮流输出“停顿符”，而是两个系统都在那段空白中维持着彼此的在场，并在那个在场中持续进行着尚未形成命题的互调。这个“在”的要求，当前架构不满足。

条件三（实时互校正）在技术上的失效最为隐蔽。表面上，多轮对话中的AI确实会根据用户的新输入“修正”自己的下一轮输出。但仔细看这个“修正”的动作，它永远发生在用户的完整输入结束之后：用户把一整段话打完、发送，AI基于这段完成的输入生成下一个回应。场要求的实时互校正，则发生在一句话尚未完成的过程中——一个词刚出口、一个停顿刚发生、一个音调的微妙上挑，就足以让对方的下一步发生偏转。这种粒度的互校正，需要对话双方在话语的微观时序中维持连续的耦合状态，而不是以一个个完整的“话轮”为单位做轮替回应。当前AI的交互协议（输入—处理—输出）天然把话轮边界变成了互校的硬断点——因为在用户“正在说话”的那几秒里，AI是哑的、是静止的，它完全错过了那个唯一能发生实时互校的窗口。

条件四（对不确定性的共同容忍）的失效，涉及一个更深的架构特征：AI的不确定性，在当前架构中，始终是概率分布上的不确定性——给定同一输入，模型会计算出一个可能的输出token的概率空间，从中采样。这种不确定性的“形态”与人在场中经历的那种不确定，有本质区别。人的不确定是：我不知道我面前这个他者会用他的下一步把我变成什么样子。AI的概率不确定是：我不知道我这次会从预设的概率空间里抽出哪个token，但无论抽出哪个，我都没有被改变。概率分布在采样前是不确定的，但采样完成后，整个系统回到那个没有被改变的状态，等下一轮输入。真正构成场的那种不确定性，恰恰是“我会在这次交互中失去我现在的样子”——这个可能性，当前AI架构把它作为噪声去处理了，而场把它作为唯一的燃料。

总结来说，这不是一个“算法不够好”的问题——它是架构层对“在场”的定义问题。当前主流的交互架构在设计之初，就把在场定义为“当且仅当生成输出时”，而在那之外的整个时间通量里，系统被设计为“不占用任何认知资源”以提高吞吐效率。这是一个在效率约束下被合理做出的工程选择。但也正是这个选择，在发生学上等价于：场的全部微条件，已被工程规范提前清空。

2.5 “位先于场”如何成为当代认知生态的默认稳态

如果“场先于位”在发生学上是理解的更基本形态，那就必须追问一个历史问题：为什么在今日的认知生态中，“位先于场”反而成了默认值？本文的回答是：这不是一个“技术入侵”或“人性堕落”的简单故事，而是一系列彼此独立的改进运动——它们在各自的领域内都正当且有效——在认知交往的层面上共同收敛出的一个稳态。

第一步：教育的可测量化转型（约1960s–1990s）。战后大众教育体系的扩张带来一个制度难题：如何公平而高效地评估大量学生的认知能力？答案是大规模标准化测试和评分量规（rubric）的普及。这一转型的正当性毋庸置疑——它在制度层面压制了出身与主观偏袒对教育机会的干预，用透明、可申诉的标准体系替代了教师个人的隐秘裁量权。但它的副产物是一种认知主体模型的悄然重塑：在评分量规的长期反馈下，学生被反复训练为一个先在的意义拥有者——先想清楚、再表达，表达接受量规逐项评分。那些需要在与他人的实时对话中慢慢逼出形状的认知活动（课堂上的试探性发言、论文初稿中连作者自己都不知道结论在哪里的摸索、与教师的非结构化长谈），因为无法被量规有效捕捉，从正式评价中逐渐隐退，进而从学生的学习策略中被边缘化。这不是某一堂课教坏了学生，而是学生在整个教育生涯中，被反复置于一个只对“位先于场”行为给予正式肯定的环境里。其后果不是学生“丧失了进入场的能力”，而是他们学会了把场的表现形态储存在私人领域里，而在公共的制度空间中默认启动“位”的认知模式。

第二步：职场协作的效率化转型（约1990s–2010s）。信息技术在办公领域的全面推开，使得书面化、结构化、可追踪的协作方式逐步取代了面对面的、电话里的或走廊里的即兴交谈。项目管理软件把协作拆解为可分配、可量化、可控进度的“任务项”；OKR与KPI体系要求每个成员在协作开始之前就明确自己将要产出的成果及其测量方式；会议越来越强调“会前阅读材料、会中高效决策、会

后明确行动项”。这些制度的综合效果是：每一次职业交往都必须在开始时就坐稳各自的位——你带着你负责的那块已经成型的方案进会场，他带着他的，大家交换成品并做出裁决。那种“我们都还不确定这个方案应该长什么样子，让我们坐在一起聊聊看能聊出什么”的工作方式，虽然会在某些创意团队或非正式场合存活下来，但它已经不再是制度认可的默认协作动作。它变得“奢侈”——而奢侈，第一个被日程表删除。这不是谁的过错，而是一个在效率正当性驱动下的集体选择。

第三步：社交媒体与AI辅助界面的平滑化（约2010s至今）。上述两个转型已经把“位先于场”铸成了制度框架；而近十五年数字界面的演化，则将它从制度框架进一步下沉为日常认知操作的肌肉记忆。社交媒体设计的深层语法是“发表—回应”：你不是在一条帖子里和另一个人共同逼出一个新想法，你是发表一个已经成型的观点（即使它是情绪性的、尚未深思的，它仍然在形式上是成品），等待回应。评论区里当然会有冗长的争论，但争论中真正在实时互校中改变自己立场的情形极为罕见——因为展露“被对方改变”的行为，在评论区的公关式辩论中被默认等同于“认输”。AI辅助界面的普及，进一步平滑了这条轨道：用户被鼓励在对话框中输入清晰、结构化的指令，以获得高质量的输出。那些最需要被场逼出来的东西——心里一团还不清楚是什么的混沌，一个只是在咬合处隐隐发疼却还没找到病灶的困惑——因为无法被输入为一个有效prompt，被排除在界面的交互之外。它们没有被禁止，只是被静默地“不做处理”。而那些能够被处理的、清晰的、成型的需求，则在几秒钟内获得了越来越精美的回应。这种“即时被服务”的快感，在与另一个人的笨拙对话中完全得不到——于是人类自己引导自己的认知交往，在偏好的无意识累积中，更多地流向“位先于场”的轨道。

把这三步叠加起来看，一条发生学路径就清晰了：教育的可测量化、职场的效率化、界面的平滑化，分别在不同的历史时刻、由不同的动力推动，但三者在一个关键点上形成了收敛——它们共同把“先在内部成为一个完整的意义拥有者”，从一种制度性的工作要求，变成了一种认知主体对自己内部状态的默认管理方式。个体不再只是在外部的考试或考核中被要求“先想清楚再说”；他在自我的内部认知习性中也逐渐学会了：不确定是一种需要被克服的缺陷，而不是一个正常发生的场的前奏。

当代认知交往的默认稳态，就是这条路径的收敛产物。把它称为“稳态”而非“危机”，是本文刻意采用的严格发生学语言：它不是一个病，不是在偏离某个健康的理想状态；它是在特定的历史约束条件下（公平、效率、大规模协调），被正当选择并不断强化的一套认知交往模式。它的正当性不因为场的缺席而被取消——大规模法律程序、公共行政、教育评估这些制度之所以能运转，恰恰因为它们不需要“场”。但它的成本是：当这套模式从制度领域溢出，侵入了那些原本需要场才能完成的认知交往——深层人际理解、创造性协作、代际间的情感传承——它就把那些交往的场地，从“可能发生”压到了“几乎不发生”。

这就是“位先于位”稳态的发生学故事。它不是一个需要被医治的病，而是一个被特定历史路径逼出来的结构成熟态。指出这一点，不是为了呼唤“回归”——发生学不承认任何回归——而是为了

在下一节中指明：如果我们要在现有的稳态中重新凿出场，我们就必须从承认这个稳态的历史正当性出发，而不是从把它病理化出发。

三、分切："场先于位"与七个邻近传统的硬对话

上一节从案例中逼出了四组条件，并揭示了它们如何在当代认知生态中被系统性地压制。但一套新概念最危险的时刻，不是它还在襁褓中的时候，而是它刚刚成型、容易被邻近大概念"好心收编"的时候。一个概念如果没有与最像它的既有传统完成不可还原性的硬切割，它就不是一个新概念——它只是一套换皮说法。本节将使出全篇最严苛的解剖刀，逐一切开"场先于位"与七个最可能收编它的传统之间的那条分界线。每一刀的任务只一个：指出那个传统装下了什么，漏掉了什么；而"场先于位"恰好切开了它装不下的那笔。

3.1 与巴赫金对话性：声音先在，还是位置后被逼出？

巴赫金的对话性（dialogism）是本文必须面对的第一个——也可能是最难缠的——对手。在他的《陀思妥耶夫斯基诗学问题》与更晚期的言语体裁论著中，巴赫金极其锐利地论证了：意义从不单独栖居在说话者内部，也不单独栖居在听者内部，而在二人对话的边界上生成；话语总是"双声性的"（double-voiced），总在应答着前面的某个人并等待着被后面的某个人应答；自我意识本身，也是在对话中被建构的——"只有通过他人，通过他人的眼睛，我才能看见自己。"

这听起来与"场先于位"高度共振。确实，巴赫金把意义的生成地点从独白主体内部搬到了对话的边界上，在这一点上，他比本文所批评的"单方动作"模型远为深邃。本文对这条传统感到的不是竞争焦虑，而是由衷的敬意——正是因为它如此迫近，才必须切开。

我的切入口是：巴赫金的对话性里，那些参与对话的"声音"，是从哪来的？

在巴赫金的体系中——尤其在《言语体裁问题》相关论述中——每一个说话者进入对话时，都不是空着手进场的。他携带着特定的"社会性言语体裁"（speech genre）——这是一种被特定社会活动领域反复使用的、相对稳定的话语类型。一个军官在军队中的话语、一个父亲在家庭晚餐桌上的话语、一个恋人在情书里的话语——这些声音在进入具体对话之前，就已经被它们各自所属的社会活动领域塑好了。巴赫金会说：是的，意义在对话的边界上生成；但他紧接着会说：那些参与生成的声音本身，是带着各自的社会性"腔调"（accent）进场的，它们不是被对话从零逼出来的。

这与"场先于位"之间存在一道硬性分界线。巴赫金的对话性，描述的是已经携带了社会性腔调的声音如何在对话中遭遇、冲突、杂交、彼此改变；而"场先于位"追问的是一个比这更早的时刻：在这场对话发生之前，那个"被理解的我自己"的位置——我对自己某段经验的那种可以被语言命名的清晰感受——并不存在。它不是被我的社会性言语体裁预先给定了一个"腔调"然后带进场里的，它是在这场对话里第一次被逼出了一个腔调。换句话说：巴赫金的对话性回答的是"先在的声音如何

被对话改变”；“场先于位”回答的是“声音本身的位置——谁是有话要说的人，什么是那个话的内容——如何被对话第一次生出来”。前者处理声音的变调，后者处理声音的创生。

这不是在说巴赫金错了。巴赫金要解释的是社会话语的多样性与对话本质，他不需要去追问“位置第一次从哪里来”——因为在他的分析层上，社会位置先于任何一次具体对话而存在，是很合理的预设。但本文要追问的，恰恰是当这个预设本身在特定交互中被暂时悬置时，会发生什么。那个深夜阳台上的说话者，在开口之前并不是以一个“已经被社会位置定义了腔调的主体”的身份进场的——他是以一个“连自己有什么要说的都还不知道”的混沌状态进场的。巴赫金的对话性装不下这个状态，因为他的理论需要的正是声音已经可以被识别为归属于某个社会领域的那一时刻；而本文的场恰切发生在那一时刻之前。

这就是不可还原的那一刀：“场先于位”切开的，是巴赫金对话性里“发声位置已在社会语言的编织中先行给定”这个隐性预设所覆盖不到的那道缝——那个连发声位置本身都还需要被对话生出来的原初交互情境。

3.2 与维特根斯坦语言游戏：规则先在，还是规则的生成？

维特根斯坦在《哲学研究》中对“私人语言”的摧毁性论证，是二十世纪哲学中与单方动作模型最彻底的决裂。他的核心论证是：不可能存在一种原则上只能被说话者自己理解的“私人语言”，因为任何语言的意义都必须依赖公共的可遵循规则——而“遵循规则”这件事，不可能由单一个体在内心私密地完成。它需要在语言游戏的公共实践中、在“做”（doing）中、在共同体的纠正与确认中，被成就。理解，不是一个孤立的心理事件，而是一种公共实践中的能力。

这个判断在宏观方向上与本文高度一致。但问题出在同一个地方：规则从哪来。维特根斯坦的语言游戏理论，描述的是已经存在的语言游戏，以及新加入者如何被训练进这些游戏。他极其深刻地展示了：你不可能一个人给自己私人地发明“理解”这个词的用法；这个词的用法，是你在一个语言共同体中被反复纠正、在无数次“对、继续”“不对、不是这样”的反馈中被训练出来的。这个共同体拥有语言游戏的规则，而你的“理解”被这些规则所定义。

“场先于位”要追问的，不是这个——不是“已被规则定义的理解如何在共同体中被训练”，而是：在两个人都还找不到现成规则可用来处理当下情境的时候，规则本身是怎么被逼出来的？积木片段中，小D和小C在做的，不是在遵循一个既有的“飞机游戏”的语言游戏规则——他们是在现场一起制造那个规则。没有一个人事先知道“碰积木边缘表示接入”“弧线轨迹意味着在找落点”“把三角积木放在弧线下方可以让对方接受这是一个机场”这些规则。这些规则，是他们在交互中被共同逼出来的，而且在逼出来之前，连他们自己都不知道会有这样的规则。

维特根斯坦当然可以回应说：他们仍然在运用更底层的语言游戏规则——比如“命名一个东西”的规则、“遵循一个邀请”的规则、“回应一个微笑”的规则。这个回应在哲学上是成立的，但它恰恰暴

露了语言游戏理论的边界：它擅长解释复杂的实践活动如何可以被层层还原为更底层的既有规则，但它不去处理规则在交互中第一次被创生的那个时刻。把那个时刻说成“只是在运用更底层的规则”，只是把问题的起点往下挪了一层，并没有回答问题本身：那些更底层的规则，又是怎么来的？维特根斯坦不需要回答这个问题，因为他的任务不是发生学——他不需要追问规则的发生，他只需要论证规则的公共性。但本文的任务恰是发生学：追问那第一次把规则逼出来的交互事件，它的发生结构是什么。“场先于位”切割的，正是语言游戏理论里“既有规则总在交互之前已经存在”这个预设所遮蔽的创生地带。

3.3 与常人方法学和对话分析：穷尽了交互机制，还是漏掉了位置的发生？

这是本文必须面对的最邻近的技术对手。常人方法学（ethnomethodology）的奠基人Garfinkel以对“秩序的局部成就性”的精密分析闻名——社会秩序不是高高悬在个体头上的铁笼，而是个体在每一刻的具体交互中、用他们的“常人方法”（ethnomethods）持续地、局部地、即兴地“做成”的。同样，以Sacks、Schegloff、Jefferson为代表的对话分析（Conversation Analysis, CA），把话轮转换、修补、沉默、偏好组织等交互机制拆解到了令人叹为观止的微秒级精度——Schegloff对电话对话开头的研究、Jefferson对“重叠话语”和“沉默时长”的精微测量，都展示了人类何以在交互中完成极其复杂的实时协调。

本文2.2节的四组条件——空白停留、实时互校、相互容忍——在字面上看起来极像CA术语的哲学改写版。如果我不能证明“场先于位”切开了CA已经做的事情，那本文的命名就只是对CA已有成果的现象学复述，不配称为新概念。

我的切入口是这样：CA穷尽地描述了交互如何被“做成”，但它系统性地不追问——谁在做？

CA的方法论基础，是对录音/录像转写材料的逐行分析。它的分析单元，是话轮（turn）、话轮构建单位（TCU）、序列（sequence）、修补（repair）——这些单元都预设了一件事：有一个说话者和一个听者，在轮流交换话轮。CA当然会描述“说话者如何通过话轮构建来分配下一个说话者”这样的精细机制，但它从不追问“说话者和听者这两个位置本身，是如何在对话中被第一次分配的”。因为这个追问，对于CA来说是一个方法论上不可操作的问题：你永远不可能在录音转写中“听到”位置在被分配之前的那个混沌状态——因为录音机能录下的，只能是已经被说出来的话；当第一个词被说出来时，说话者已经作为一个说话者而立在那了。CA的分析时间起始于第一个话轮开始的那一刻。但本文的场，发生于那之前——发生于一个人还在犹豫要不要开口、颤颤巍巍试探性地放出第一声“飞机”的那个前话轮时刻。那个时刻在CA的转写里只会被标注为“沉默”或“话轮内停顿”，不会被分析为一个正在发生位置分配的过程，因为CA没有这个分析范畴。

换句话说：CA穷尽了“交互机制”的描述，但它没有——也不需要——追问这个机制本身的产物是什么。“说话者”和“听者”这两个位置，在CA的理论中是被当作机制的输入而非输出来处理的。本

文的“场先于位”，恰恰把这两个位置当作机制的输出来处理：它们不是对话的起点，而是对话的产物。CA的分析粒度是毫秒级的，但只要它不把“位置的发生”当作一个变量来追问，它就收编不了本文的核心判断。这一刀，是粒度对粒度的硬切。

3.4 与克拉克的联合行动：共同基础是输入还是输出？

Herbert Clark在《Using Language》（1996）中发展出了一套极有影响力的“联合行动”（joint action）与“共同基础”（common ground）理论。Clark的核心洞见是：语言使用不是单方编码—解码，而是两个人共同参与的“联合行动”——每一次话轮交替都叠加在前一次话轮的基础上，逐步累积出一个只有这两个人能共享的“共同基础”。共同基础不是事先摆在那里的，而是在交互中被两个人一层一层共同搭建起来的。

这听起来与本文的“场”极为相似。但仔细看Clark的论证结构，会发现一个微妙的但决定性的倾斜。在Clark的模型中，“共同基础”以“共同感知”“共同知识”和“联合行动本身”为来源——两个人如果同在一个房间，他们都看到了一盏灯亮了，那么“这盏灯亮了”就自动进入了他们的共同基础。在这个过程中，这两个人仍然是被当作已经有了各自认知系统的积分器的：他们各自接收外部输入（看到灯亮），各自进行推理（他知道我看到了，我知道他知道我看到了……），然后以交互为渠道来验证和扩展这个已经建立的基础。

这个模型在处置场的发生时，有一个它自己永远无法反身的盲点：Clark的联合行动和共同基础的框架，始终把交互当作已有认知系统之间交换和累积信息的渠道，而没有把交互当作逼出任何一个认知系统内部都不曾存在过的新认知形态的引擎。联合行动可以解释两个人如何协作搬运一件重物（目标已知、角色可分配、步骤可协调）——但无法解释两个人如何在协作开始之前连“我们要搬什么”都还不知道的情况下，一起摸出了那个要搬的东西的形状。前者是功能耦合，后者是形态创生。

这一刀的切法：Clark的联合行动，其分析的起点是两个已经有各自意向和认知内容的行动者，他们通过交互来协调这些意向并累积共同内容；本文的场，其分析的起点是两个连自己的意向和认知内容都还处于前命题混沌态的主体，他们通过交互来第一次逼出那个意向和内容的形状。Clark的共同基础是交互的输出，但它是以“两个先在认知系统的输入”为前提的；场的位置分配同样是交互的输出，但它在交互发生之前，连谁的认知系统里将要生成什么样的内容都是变量。Clark的回答是“共同基础是被搭建出来的”；本文的回答是“搭建者是谁也是在搭建中被决定出来的”。前一个回答覆盖了联合行动的全部常规情形，后一个回答切开了一个它装不下的极限情形。

3.5 与胡塞尔-舒茨-哈贝马斯的主体间性传统：存在论描述，还是发生学条件追问？

审稿人准确地指出，源自胡塞尔、经舒茨到哈贝马斯“交往行动理论”的“主体间性”（Intersubjectivity）传统，是本文必须正面对手的另一个巨兽。这个传统同样强调理解是主体间共同建构的，而非单方投射。哈贝马斯尤其把“交往理性”建立在对“理想言谈情境”的预设之上——在不受权力扭曲的交往中，参与者通过论证来达成相互理解。

这一传统与本文的同路之处不必赘述。切入口在同一个老地方：传统主体间性所做的工作，要么是存在论的（揭示自我与他者如何在本体上是相互缠绕的，如胡塞尔对“共现”的分析、梅洛-庞蒂对“肉身间性”的分析），要么是规范论的（给出理想交往的条件，如哈贝马斯的有效性宣称与理想言谈情境），而非发生学条件论的——它不追问：在什么具体的、历史的、制度的、认知生态的微条件下，一次具体的交互会或不会实际发生那种“共同建构”的过程？

胡塞尔的现象学描述了自我如何通过“同感”（Einfühlung）在原初的自我学领域中构成他者——这是对主体间性的先验结构分析。在它的框架中，主体间性的可能性是意识活动的先验常量，不是历史变量。舒茨把它社会化了——日常生活世界中的他者理解，总是发生在特定的“视角互易性”和“知识的社会分配”条件下。但即便是舒茨，其分析焦点仍落在“类型化”（typification）和“被认为理所当然的”（taken-for-granted）社会知识结构上，而非一次交互从混沌到结晶的动力学。哈贝马斯的交往行动理论，则把主体间性设定为一种在理想言谈情境中可被实现的规范——它追问的是“真正的共识需要什么条件”，而非“一次真实的相互理解的发生需要什么微条件”。

本文的“场先于位”正是在这个缺口里长出它的不可还原性：它不做存在论描述，不做规范论设定，它追问的是——在给定的历史—制度—技术条件下，一次具体的双向可塑的交互能否实际发生？如果不能，是哪个微条件被卡住了？这组追问，主体间性传统不曾系统处理，不是因为它不屑，而是因为它的理论任务在另一个层。本文的层，是发生学的条件层——把理解从“意识结构”“社会知识分配结构”或“理想交往规范”里拽出来，拽回到一次具体交互中，那些微小的、具身的、被制度与技术所压制的发生条件上。

3.6 与格莱斯的意图识别模型：意义在说话者内部还是在交互的边界上？

格莱斯（H.P. Grice）的会话含义理论，是分析哲学传统中对“理解”这个题目最精致、最系统的一套理论操作。它的核心骨架很简单：说话者S以话语U意谓某件事p；听者H通过识别S的交际意图——包括S说U的意图以及S希望H识别出这个意图的意图——来理解U。理解，就是意图的成功识别与传递。

严格来说，格莱斯的模型不是一个关于人际深层理解的完整理论——它的典型素材是“你能把盐递给我吗”这样的日常会话含义推断，而非“深夜阳台上两个人共同逼出一段从未被说清过的经验”那样的发生事件。但正因其在分析哲学中的统治地位，本文必须清晰地指出：为何格莱斯的“意图—识

别”框架，不是本文所论的理解的发生学基底？

格莱斯框架的核心特征，是它把“意义”的所在地预设为说话者个人的内心。说话者在发话之前，先有一个“我意谓p”的意向状态；发话就是这个意向状态的公开化；听者的任务就是从公开化的信号中逆向推断出那个先在的意向内容。在格莱斯的模型里，意义在交互开始之前就已经在说话者的心灵内部成型了——它的形状是p，它的边界是确定的，它等待的只是一个足够好的编码和一份足够细心的解码。

“场先于位”追问的，是意义还没成型的那种情形——说话者自己在开口时心里并没有一个相当于p的清晰命题，有的只是一团模糊的、尚未被语言驯化的、连他自己都在交互中第一次逼近的东西。这种情形下，没有先在的“意图”可以被逆向识别——因为意图本身，是在交互推进的过程中才逐步凝出形状的。格莱斯的基础认知架构（先有意图，后编码，再传递，最后解码）在原理上无法处理这种情形，因为它预设了那个最要紧的东西——意义的形状——在对话开始之前已经完成。

当然，你可以把格莱斯的模型扩展：说话者的意图不是一个固定的p，而是“我想要和这个人一起逼近某个我还说不清的X”——但这样一来，意图识别就不再是识别一个确定的内容，而是识别一个开放性邀请。而“识别开放性邀请”这件事，本身已经超出了格莱斯的原始框架——它不再是一个对命题态度的二阶识别，而是一个需要在交互中实时参与、相互试探、共同推进的过程。在这个过程中，意图不再是那个推动理解的动力源，而是被交互不断地重新生成和修改的派生品。因此，意图不在解释的底层——场在更底层。

3.7 与预测加工：双向耦合还是交互创生？

预测加工理论（Predictive Processing），尤其是Friston的自由能原理与Clark近年的行动导向预测加工模型，在当代认知科学中构成了一个极有吸引力的替代框架：大脑被理解为一个层级化的预测机器，它不断地生成自上而下的预测并试图最小化感官输入的预测误差；对他者的理解，就是对自己将要从他者那里接收到的感官信号建立最优预测。在某些版本中——尤其是那些强调“双向耦合”的版本——两个预测系统可以通过反复的相互预测来达到某种同步。

这个框架看似也处理了“双向”和“互调”，但与本文的场之间有不可通约的硬分歧。在预测加工框架中，两个进行双向预测的系统，各自都先验地拥有一个生成模型——那是用来产生预测的内部结构。这个生成模型可以在交互中通过预测误差被更新（学习），但更新的目标永远是最小化预测误差——系统的优化方向是“让自己的预测更接近对方实际发出的信号”。把对方变成自己可预测的。

场的发生，则走了一条完全相反的路。在积木片段中，小C把小D的弧线轨迹的可能落点预读出来，并放了一块积木在那里——这个动作，不是为了“最小化她自己对小D的预测误差”，而是为了给小D的下一步创造一个新的约束，让小D在那个约束里做出他自己本来做不出的动作。小C的介入不是让系统更准确地预测小D，而是让小D在那一刻变成了另一个样子——变成了一个愿意落在“机场”

上的飞行员。这是干预而不是预测，是创生而不是收敛到最优预测。

这中间的底层差别在于：预测加工的最小化逻辑，内在地偏向稳定和收敛——系统通过消除预测误差来维持自身的稳态，即使引入了“主动推理”（active inference），也是通过行动来采样符合预测的感官输入，以减少意外。而场的双向可塑性，需要的恰恰是系统能够容忍甚至追求“意外”——一种“你刚才做的这件事不在我的预测空间之内，但它让我打开了一个我本来没有的下一步”的交互事件。这种事件在预测加工框架中，被信号为“高预测误差需要被消除”；在本文的框架中，它被识别为“新意义正在发生的唯一入口”。一个是把不确定性当噪声，一个是把不确定性当燃料。这不只是立意不同，这是结构性的不可通约。

四、再入场：在什么条件下场能被重新逼出？

前面三节完成了拆解与切割，这一节该回到建构上来。如果“位先于场”是当代认知交往的默认稳态——是在效率、公平、大规模协调的正当压力下被历史地选择出来的收敛产物——那么我们就不应该把它当作需要被消灭的敌人。它不需要被消灭，也无法被消灭。真正的问题不是“如何用场替换位”，而是：在明确知道位的结构占据主导地位的当下，在什么样的局部条件、制度缝隙和交互微选择中，场仍然可以被逼出来？本节的任务，不是给出一份生活建议清单，而是用与第二节同等硬度的机制分析，去识别那些能够让场的四组条件在当代认知生态中重新被激活的具体情境。

4.1 条件反推：什么样的情境可以重新打开场

从第二节的四组条件出发，可以反推出：场的重新发生，要求一组与当下主流认知生态的默认值正好反向的设置。

第一：一方先放弃“位”的安全。双向可塑性不能空等，它需要一个初始动作。在不对等的交往中（师—生、医—患、管理者—被管理者，乃至亲密关系中那些默认“我更冷静”的一方），拥有更多安全空间的一方，必须先做那个放弃者——先把自己的预设暴露出来，先说出“这一步我也不确定”，先让对方看到自己正在被对方影响。这不是道德呼吁，而是发生学上的必要条件：如果两个人都在等对方先拆下防御，场永远无法开始。在权力不对等中，这个初始动作必须是上位者做出的，因为在不对等关系中，下位者的“不确定”一旦暴露，立即会被权力结构解读为缺陷；只有上位者用自己的不确定去顶开那个缝隙，下位者才能安全地跟进。这不是理想主义的“平等呼吁”，而是一个精确的、可操作的条件识别：场在不对等关系中的发生顺序必定是单向先动的，而这个先动者只能是承受更少风险的一方。

第二：有意识地延迟“命题封装”。场需要信息空白处的共同停留，而在当代交往中，空白是被“效率冲动”第一个挤掉的东西。要重新打开空白，就不能把它当成尴尬的冷场去填充，而要把它当作还在孕育的东西去守。具体来说，就是在对话中的某些时刻，当对方刚说出一个试探性的、还带着

很多不确定的词时，不立刻用善意的“你是不是想说X？”去帮他封装——封装是位的典型动作：帮对方把还不成型的东西赶紧做成一个成品。延迟封装，等于给对方留出那个他自己去逼近自己意思的时间，也给自己留出那个被他的混沌牵引到陌生方向的机会。这种延迟不是被动的“不说话”，而是一种主动的承载姿态——用注视、用倾听的姿势、用不用力去催的身体语言来告诉对方：不急，它还在长。

第三：在微观节奏中恢复不可预测的互调机会。实时互校需要一个对话节奏，这个节奏的敌人不是“话多”，而是“话太完整”。当代成人的对话，太习惯于把每一轮发言都做成一段完整的、有开头有结尾有论点的微型文章——这种“完整发言”的规范，从教育到职场被反复训练，最后变成了认知交往的默认格式。但它恰恰封死了实时互校的入口：如果每个人都是在自己完全说完之后再交给对方，那么互校就只能发生在话轮之间，永远进不了话轮之内。场的恢复，需要有意识地容许自己的发言出现断裂、犹豫、半句话、临时改口——这些不完整的东西，不是“表达缺陷”，而是唯一能邀请对方进入你的语流的信号。一段被打断的、还在摸索的、尚未成型的语流，比一段精雕细琢的完整发言，更有机会激发出对方的实时介入。

第四：共同保护那片“我们不知道会走到哪”的前方。这可能是最难操作的一个条件，因为它对抗的不仅是外部的时间压力，还有参与双方内部的认知习性。在每一次重要对话开始之前，双方都需要做一个心理上的“预放弃”：放弃对对话走向的控制，放弃“今天必须得出一个结论”的预设任务，放弃把对话当成一个需要被高效完成的项目。这个心理动作不需要被大声宣告出来——宣告本身就太郑重、太像另一个任务了——但它需要被双方的姿态所传达：对话的起点，是我们都不知道答案；这趟路的价值，不在终点那个结论，而在走的过程中逼出了什么。

把以上四点并置，可以看见一个清晰的共同结构：场的重新发生，在任何给定的当下情境中，都不是一个“自然”的自发过程；它要求至少有一方做出反向于周遭认知交往默认值的微选择——选择慢于效率、选择不确定于清晰、选择暴露于防御、选择过程于产出。这些选择之所以是“微”的，是因为它们不需要制度变革、不需要技术许可、不需要全社会的节奏转变；它们可以在两个人之间的任何一次对话中被立即执行，只需要其中一个人先做出那个反向动作。

4.2 制度缝隙：哪些已有空间中场的发生概率更高

以上论述停留在个体互动的微层面。但本文对“位先于场”的分析，从一开始就是制度性的——教育的可测量化、职场的效率化、界面的平滑化，都是制度选择沉淀成的认知生态。因此，场的恢复问题也不能只停留在个体选择层面，还必须追问：在现有的制度结构中，是否存在某些缝隙，由于那些缝隙自身的运作逻辑尚未被完全“位化”，使得场的发生概率天然更高？

本文识别出两类制度裂缝。

第一类是“非产出导向的陪伴性制度空间”——例如心理咨询中的特定流派（非结构化长程动力性治疗）、临终关怀中的陪护、某些形式的导师制指导关系（而非绩效导向的辅导）。这些空间的共同特征是：它们所追求的效果，在操作层面无法被预先量化为清晰的目标—手段链。动力性治疗师的工作，不是给来访者一个诊断然后按照标准化流程去“修复”；而是在长达数月甚至数年的每周五十分钟里，持续地和来访者一起待在他的混沌里，让那个混沌在关系的长期容纳中自己找到形状。临终陪护的志工，根本不抱“解决问题”的目标——他们唯一的任务就是“在那里”。在这些空间中，“效率”的逻辑内在地失效了，因为效率需要一个可被预先定义好的产出，而这些空间存在的全部原因，恰恰是那个产出在发生之前无法被预知。因此，这些空间天然地为场的发生保留着比普通制度交往更高的概率。本文不主张把所有认知交往都变成长程治疗，但本文指出：这些制度裂缝的存在，证明了场的微条件在当今社会并非全无立锥之地。它们立在那，只是被更主流的效率话语挤到了社会的边缘。

第二类是“目标开放性”的创意协作形式——例如某些不以产品上市为直接驱动的艺术集体创作、自由软件社区中某些没有截止日期的底层架构讨论、学术领域中少数仍在抵抗“论文工厂”节奏的跨学科长期研读小组。这些形式的共同特征是：它们的参与者在一开始并不完全清楚协作的最终产出是什么，并在制度上被容许（有时甚至被鼓励）保持这种不清楚。这就给场的条件三和条件四留出了存活空间。当然，这些空间在当代学术与创意产业的评价体系下面临着巨大的生存压力——它们不断地被绩效指标追问“你的产出是什么”。但这恰恰是本文想要指出的：场的存亡，并不取决于“人性是否丧失”，而取决于制度结构是否还留有一些不追问效率的缝隙。只要缝隙还在，场的可能性就没有从社会中蒸发，只是被压到了边角。

4.3 “反向微选择”的训练可能吗？

最后一个需要回应的问题：本文描述的那些反向微选择——延迟封装、容忍断裂、暴露不确定——听起来像是某种需要特殊天赋或个人修为的交往艺术。如果它们确实是场的必要条件，那是否意味着场的发生只能是少数天生“善解人意”者的专属？

本文的回答是：不需要特殊天赋，但需要有意识的反训练。场的能力——如果我们可以把它叫作一种“能力”的话——不是某种少数人拥有的神秘天赋，而是每个人都曾在童年游戏和亲密对话中自然行使过的一种交互方式。它之所以在成年后变“难”，不是因为它本身难，而是因为我们的认知栖息环境在过去二三十年中，被系统性地修剪成了不利于它发生的形状。我们的教育经历、职业训练和数字交往习惯，都是在做同一件事：训练我们把不确定封装为确定，把过程的摸索压缩为结果的呈现，把交互的开放性收拢为任务的可控性。这个训练是有效的，是正当的，是让我们在这个社会里高效运转的前提。但它同时也是让场的微条件从我们日常交往中消失的那个推手。

因此，重获场的能力，不需要额外的天赋，但它需要一种有意识的“反训练”：在那些你选择要进入场的交往中，刻意地做与你的日常工作习惯相反的事——不封装、不总结、不抢答、不要求对

方把话说完整。这不是学习新技能，这是暂时卸下旧技能。它不需要额外的认知资源，只需要一个暂时的“停工”——暂时停止让效率逻辑指示你该说什么。它不是“成长”，不是“修炼”，而是一个有意识的“不治”——不去修理那些本来就不需要被修理的沉默、断裂和不确定。

诚然，这种反训练是否有普适效果，不是本文能够下结论的——这需要实证研究的跟进。本文能做的，是精确指出它与场的发生之间的机制链接：如果不做有意识的反向选择，位先于位的默认值就会在每一次交往中自动启动；而每一次自动启动，都意味着场被再一次静默跳过。选择不去跳过的那一下，就是场唯一的入口。

五、反驳与回应

一个命题被推到这一步，必须在交卷前直面那些会被同时逼到喉口的最有力反驳。本节选出四个最尖锐的反对意见，逐一给予正面回应。

追问一：如果“位先于场”在制度层面运转良好（法律程序、科学同行评审、公共行政、大规模教育评估），它们并不产生“理解危机”，那么“场”是否只是对一种低效率、高情感消耗的交往方式的偏好？为什么“共同生成”比“精确映射”更配得上“理解”这个名字？

这是一个本文必须正面接住而不能闪躲的追问。本文的回应分三步。

第一步，限定论域。本文所讨论的“理解”，从第一节起就严格限定在“人际间的一次发生性的、在实时交互中共同生成新意义的过程”。法律程序中的理解——法官对法条的解读、律师对判例的援引——本质上是对既成文本在公共标准下的精确适用，完全适用“位先于场”模型。本文不主张用“场”替代这些制度性理解，因为法律的可预期性恰恰建立在位先结构之上——你不能要求法官在每次判案时都和当事人“共同生成”一个新的法律含义。同样，科学同行评审中评审者对投稿的理解，核心是对既成论证的检验，而非与作者在实时交互中一起逼出一个双方都不知道的新论证。本文在这些领域不主张“场”的优越性，因为“场”本来就不是为这些领域而设定的分析框架。

第二步，否定“更配得上”的提法。本文不主张“共同生成优于精确映射”。两者解决的是完全不同类型的认知任务。一位急诊医生在接诊时对症状的快速分类与处置，必须依赖位先于位的精确映射模型——他没有时间、也不应该和患者一起进入场去慢慢逼出诊断。而在同一位患者于数月后与心理治疗师的长期工作中，同样一组症状背后的那个“说不清的痛苦”，却只能在场的微条件中慢慢被逼出形状。两种情境适用两种不同的理解结构，非但不是互斥的，反而是互补的——前者负责紧急处置，后者负责深层成形。问题不在于哪种更优越，而在于：当位先于场的制度模版在当代认知生态中成为唯一可见的范本时，那些本该由场来完成的理解，被不当地塞进了位的框架里，然后在那个框架中“理所当然地”失效。本文做的，不是拔高场而贬低位，而是把两者放回各自适用的不同认知情境中，指出其中一种在当代被过度扩张了。

第三步，回应“效率”之忧。场确实比位需要更多时间、更多情感承受力、更多对不确定性的容忍。在资源匮乏或时间紧迫的条件下，位是更合理的默认选择。本文完全承认这一点。场的发生从来不应该被当作普遍规范——它是一种奢侈的、只应该在特定重要交往中被有意识地选择的模式。这个“奢侈”不应被读作道德上的优越，而应被读作认知生态学上的事实：场的发生需要一组不容易满足的微条件，我们不能要求每次交往都具备它们；但我们也不能因为它们在大多数交往中不满足，就否认在那些对我们最重要的交往中，它们的存在与否是决定理解能否真实发生的关键变量。

追问二：在场发生的不平等问题上，“场先于位”如何不至于沦为一句空洞的乐观？

审稿人指出：在权力不对等的交往中（医—患、师—生、上下级），较弱一方的“不确定”与“还没想清楚”往往无法被安全地展露，因为展露会立即在权力结构中被诠释为“不专业”“不用心”“没准备好”，并承受实质性损失。在这种情况下，“位先于场”不是双方自愿的选择，而是权力已经提前封死了进场的缝隙。

本文完全接受这一事实。本文的回应不采取“收缩论域至平等交往”的智性逃避。恰恰相反：在权力不对等的场景中，场的结构性缺席，应被直接诊断为权力运作的一个沉默后果，而不是交往技巧的失败。

这个诊断的实践意义有两层。第一，它把“理解不能发生”的责任从较弱一方身上卸下来——不是你不擅长表达，而是你被结构性地阻止进入那个唯一能让你的意思被共同逼出来的交互状态。第二，它让那些在权力上位的一方，如果确实在意理解的发生，就必须不使用自己的位置优势去要求对方“先把话说清楚”，而是主动用自己手中更多的安全空间，去为对方凿出一片能进场的微现场——先暴露自己的不确定，先停下来不说那个准备好的结论，先让对方看见你也在忍受混沌。在权力不对等中，进场的顺序不可能是对称的；但这不等于场不可能发生。它只意味着，拥有更多不被惩罚的自由的那一方，有结构性的先行义务。这不是天真的平等呼吁，而是一个精确的、可操作的要求。

追问三：如果未来AI可以在行为上完美模拟场的全部可感信号——适时沉默、语音迟疑、模拟被改变——那么“真实的场”与“被模拟的场”的区分还有意义吗？

这个追问触及了一个更深的本体论问题。本文不从“机器没有意识”出发去回避它，而是尝试给出一个在本文的机制框架内部站得住的技术回应。

区别在于“拟态”与“发生”。模拟是：系统预先拥有一种行为的模式（如“被改变”的话语模式、表示停顿的省略号），然后在某个触发条件下输出这个模式。发生是：系统在交互开始之前并不拥有那个最终状态，那个状态是在交互的实时推进中，被对方的具体行为第一次逼出来的。当前AI的架构，可以在行为表面模拟“被对方的追问改变了想法”的语言文本——如“你说得有道理，我重新想了一下……”——但这个输出的产生机制，并非来自对方那句话在实时中对AI的认知拓扑造成了不可预知

的重塑。它来自AI在训练语料中学到的一个模式：“当对方提出强有力质疑时，表示被改变是合适的回应策略。”这个模式是在训练中被固化到权重里的，而非在当下的交互中第一次被凿出。这就是拟态：行为上不可区分，发生学上截然不同。

那么，如果未来出现一种新架构——它在单次交互中确实发生了实时权重更新，且更新方向无法被设计者预先指定——这时，区分的立足点就变了。本文的核心判断并不建立在“AI永远不可能进入场”的永恒禁令上；它建立在一组可检验的发生学条件上。第七节将明确列出那个将使本文核心判断被修正的具体技术情景。在那种情景出现之前，本文坚持区分：当前的“被改变”是预先存储的模式的触发，而不是交互中认知拓扑的重组。

追问四：格莱斯的意图识别模型是否足以解释“场”的案例——听者识别说者“想要一起探索”的意图，不就完成理解了吗？

这个反驳值得认真对待，因为它来自一个极其严谨的分析传统。本文的回应是：在格莱斯的模型中，识别一个意图，要求该意图的内容在发话者内部已经以某种命题形式存在——“我意谓p”。即使在高度复杂的元意图层级（“我意谓我意谓……”）中，这个基本结构仍然保持：每一个层级的意图内容，在它被识别之前，已经完成了。但在场的原始情形中，发话者所谓的“意图”——如果我们硬要用这个词的话——并不是一个内容确定的命题，而是一个还在形成中的、连发话者自身都未能命名的趋向。

这是一个范畴错误，不是复杂程度问题。你可以把它套进格莱斯的框架说“对方的意图就是‘和我一起逼近某个说不清的X’”——但这样一来，意图的内容本身（“某说不清的X”）在其被识别时仍然是不确定的。在格莱斯的逻辑中，不确定的内容不能构成一个意图——因为意图的逻辑结构要求其内容在意图被持有时就已被确定（即使只是部分确定）。场中的那种“我倾向某个方向但我说不清那是什么”的状态，不是意图内容的部分确定，而是那个内容本身正在以交互为条件逐步生成——它在意图的逻辑结构外部。

因此，格莱斯的意图模型解释不了场，不是因为场太复杂，而是因为场的起点落在了意图逻辑的边界之外。意图是场的产物——在交互逼出那个形状之后，说话者回头去看，才会说“是的，那就是我一直想说的”——而不是场的前提。

六、结语：在场还能发生的地方，为场凿出缝隙

这篇论文从头到尾在做一件事，而且只做一件事：把“AI能否理解人”这个一度被技术争辩所垄断的议题，搬回到一层更基本的发问上——不是“理解这个动作做得对不对”，而是“理解这件事到底发生在哪里”。分析推到这里，答案已经不能用一句话概括，但它可以被凝结为以下三点同时成立的判断：（1）在当前默认的认知交往模型中，理解被预设为理解者对既成意义的单向映射。（2）真正发生的理解，不是这个映射动作，而是一次交互过程把“理解者”和“被理解者”这两个位置一同逼出的发生事件。（3）那个使位置得以被逼出的交互条件（本文称为“场”），在当代认知生态中不是被主动禁止的，而是在效率逻辑的正当改进中被逐项排挤的；它没有被消灭，只是不再有足够深的缝隙容它发生。

这三点同时成立，但它们不构成一个需要被“解决”的危机。它更像一种发生学上的稳态诊断：不是谁犯了错，而是所有的人都在做他们该做的事——教师该评分，管理者该追进度，设计师该优化交互流畅度——这些该做的事累加起来，把另一种可能的认知交往方式挤得无地自容。那些最需要场来完成的深层人际理解，因为不再有好客的条件，就不声不响地从经验库存中稀薄了。

因此，本文最后一笔的判断不是保护，不是回归，而是在现有的稳态中主动识别和制造缝隙。这个缝隙不必大。它可以在一次深夜对话里，可以在一次不被量规追着跑的期末面谈中，也可以在一个主动按下手机静音的伴侣争执中。它的共同特征是：至少有一个人，在那一刻决定不再追求把话说清楚、把问题赶紧解决，而是选择停留在那个谁都不确定下一步会去到哪里的混沌里，并用自己的在场向对方传递“我不走，你慢慢说”。那个动作，不是任何人格修为或天赋异性，它就是场的那个初始倾斜——一个人先站不稳，另一个人就不会还站得那么稳。

这套东西能不能被训练、被推广、被制度化，本文从始至终未予承诺。本文只承诺一件事：只要有人在某次具体的对话中做了一次那个反向的微选择，场就还会发生。而只要场还在发生，“你懂我”这三个字就还没有失去它原本的重量。

本文的核心命题在什么情况下需要被修正甚至放弃？

这不是修辞上的谦虚，是发生学判断的硬纪律：一个真正的发生学命题，必须能够指出那个将使它失效的具体情景。本文的核心判断是——在当前的AI架构下，由于四组必要条件（双向可塑性、空白停留、实时互校、相互容忍）在结构上无法满足，AI无法进入“场先于位”所指的那种理解发生过程。这个判断在下述情景中需要被修正甚至放弃：

（1）出现一种新交互架构，它能在两次用户输入之间的所有非生成间歇中，维持一个与用户进行前命题层共振的模拟身体（不只是输出省略号，而是用实时变化的内在状态去承载沉默并持续被用户的非语言信号所改变）。

（2）该架构允许在单次交互过程中，被用户的试探实时改变其生成路径的底层组织方式——不是参数更新意义上的微调，而是交互本身逼出了一个在交互之前任何设计者都未能预见的、不可逆的认知拓扑改变。

（3）在满足（1）和（2）的前提下，人类参与者在与这种架构的多轮交互后，独立地、可重复地报告：他们经历了与人类“场”在现象学上不可区分的发生过程——即他们在交互中第一次形成了此前在自己内部从未成型的意义，且他们确认那个意义的形成无法还原为他们在交互之前已经内部完成的表达。

如果上述三项条件同时被满足，“当前AI架构在结构上无法进入场”这一核心判断就必须被修正。本文不把这一天的到来视为自己的失败，而是把它视为它所追问的那个问题在更高层面被重新打开。但在那一天被技术现实所证成之前，本文拒绝用行为表面上的“像”来预支发生学上的“是”。

参考文献

Bakhtin, M. (1984). Problems of Dostoe与ky's Poetics. Trans. C. Emerson. Minneapolis: University of Minnesota Press.

Bakhtin, M. (1986). Speech Genres and Other Late Essays. Trans. V.W. McGee. Austin: University of Texas Press.

Bender, E.M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623.

Bourdieu, P. (1980). Le Sens pratique. Paris: Les Éditions de Minuit.

Clark, A. (2013). Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science. Behavioral and Brain Sciences, 36(3), 181–204.

Clark, A., & Chalmers, D. (1998). The Extended Mind. Analysis, 58(1), 7–19.

Clark, H.H. (1996). Using Language. Cambridge: Cambridge University Press.

Friston, K. (2010). The Free-Energy Principle: A Unified Brain Theory? Nature Reviews Neuroscience, 11(2), 127–138.

Garfinkel, H. (1967). Studies in Ethnomethodology. Englewood Cliffs, NJ: Prentice-Hall.

Grice, H.P. (1989). Studies in the Way of Words. Cambridge, MA: Harvard University Press.

Habermas, J. (1984). The Theory of Communicative Action, Vol. 1: Reason and the Rationalization of Society. Trans. T. McCarthy. Boston: Beacon Press.

Hutchins, E. (1995). Cognition in the Wild. Cambridge, MA: MIT Press.

Merleau-Ponty, M. (1945). Phénoménologie de la perception. Paris: Gallimard.

Merleau-Ponty, M. (1964). Le Visible et l’invisible. Paris: Gallimard.

Rogoff, B. (2003). The Cultural Nature of Human Development. Oxford: Oxford University Press.

Sacks, H., Schegloff, E.A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. Language, 50(4), 696–735.

Schutz, A. (1967). The Phenomenology of the Social World. Trans. G. Walsh & F. Lehnert. Evanston, IL: Northwestern University Press.

Trevarthen, C. (1979). Communication and Cooperation in Early Infancy: A Description of Primary Intersubjectivity. In M. Bullowa (ed.), Before Speech: The Beginning of Interpersonal Communication (pp. 321–347). Cambridge: Cambridge

University Press.

Wittgenstein, L. (1953). Philosophical Investigations. Trans. G.E.M. Anscombe. Oxford: Basil Blackwell.