负能力：智能体稳定理解的居间性基础

智能体如何形成稳定的理解能力？既有技术范式——从知识库存完备到推理流程精密——始终将“稳定”等同于“输出不出错”。晚近的“走通”理论推进了一步，指出理解稳定源自系统在认知张力中主动拆解既有结构、重组出更宽广秩序的过程。

作者 黄倩盈 · 约 1.9 万字 · 发表于2026年7月24日

摘要

智能体如何形成稳定的理解能力？既有技术范式——从知识库存完备到推理流程精密——始终将“稳定”等同于“输出不出错”。晚近的“走通”理论推进了一步，指出理解稳定源自系统在认知张力中主动拆解既有结构、重组出更宽广秩序的过程。然而，这一推进受缚于一个未被审查的预设：矛盾的消解是理解稳定的前提。本文从一个真实的工程失败现场出发，追踪现有技术范式在处理结构性矛盾时的崩溃路径，在这些路径的尽头，一种被现有操作方案遗漏的组织形态被迫浮现——本文将它命名为负能力。这一命名的意义不在于发现了一个潜伏在文学史角落的漂亮概念，而在于：现有工程机制——包括知识库存、推理链搜索、多智能体辩论——在处理特定类型的矛盾时，系统性地滑向三种崩溃模式。对这三种崩溃模式的机制追踪表明，它们共享同一个缺失环节：一种不以矛盾消解为目标的、在各分支上分别走到底、以条件化形态交付结论的运作能力。本文没有从济慈出发证明这个能力重要，反问：如果这个能力缺失，系统在不可消解的矛盾面前，除了崩溃、外包或虚假融贯之外，还剩下什么选择？回答了这个问题之后，本文才论证为什么“负能力”——这个从十九世纪浪漫主义土壤中生长出来的命名——为刚才的追问提供了一种现有工程概念无法提供的结构锚定。不是因为它来自文学，而是因为它的“负”恰好标记了刚才对崩溃路径的追踪所抵达的那个不被任何正面操作词汇覆盖的辨别面。本文在此基础上给出负能力的功能界定，将其区别于认知复杂性、辩证思维、智识谦逊等邻近概念，论证其作为“走通”发生底盘的本体论地位，并从训练数据构造、推理机制设计与评估判准三个层面给出培育方略。最终论证：智能体理解能力的稳定，不是因为它记住了什么或走通了什么，而是因为它能够在矛盾的峡谷中长期驻留并进行有条件推理——这种居间状态，不是走向理解的过渡地带，而是一块必须被刻意培育的工程底盘。

关键词：负能力；智能体；理解稳定；认知张力；走通；条件化推理

一、一个真实的失效场景，以及为什么它值得被正视

让我们从一个实际部署中反复出现的挫败时刻开始。一个医疗辅助智能体，被要求就一位晚期癌症患者的治疗方案提供建议。它已在肿瘤学文献、临床试验数据和姑息治疗指南的海量语料上完成了训练。当用户提问时，它流畅地给出了A方案（激进化疗）的详细依据，包括五年生存率数据、最新临床研究引用和副作用管理策略。如果用户就此打住，这个回答将被评为“高质量”。

但用户追问了一句：“我母亲说，她宁可要最后三个月能下床走路的日子，也不要一年躺在床上的命。这个偏好在你的分析里，相当于什么权重？”

智能体的回应开始松动。它在下一轮回答中，穿插了关于“患者自主权重要性”的法学论述、一段“生命质量与生命长度”的伦理学文献综述，以及一个“在类似病例中，约67%的患者选择了A方案”的数据。但它没有在任何一处真正把用户的追问——那个关于“三个月能走路”的具体偏好——当作一个与医学证据体系构成真正矛盾的信息来对待。它要么试图用文献淹没它，要么用统计数据消解它，要么用抽象原则将它无害化。当用户在第三轮追问“你到底是觉得我母亲错了，还是觉得医学证据错了”时，智能体给出了一个自相矛盾的回答：它先说“患者的个人偏好是完全有效的”，接着说“但从医学角度看，A方案仍然是首选”，最后建议“推荐举行一次家庭会议”。

这个结果暴露了一个结构性的缺陷：智能体在面对两条互相冲突、且在现有知识框架内无法同时成立的命题——（a）医学证据表明A方案是最优选择；（b）患者的个人偏好是决策的最高约束——时，无法承载这种张力。它迅速滑向三种典型的失稳模式之一：用其中一个命题的局部更优版本来包容另一个（把偏好纳入“生活质量评分”这个医学指标，从而消解了偏好的绝对约束力）；将矛盾外包给外部程序（如家庭会议）；或者给出一个前后不一致的回答，仿佛这两条命题在不同的回合里轮流支配着输出。

把这种现象命名清楚很重要。这三种失稳模式——消解、外包、自相矛盾——不是理解失败的同义词，而是系统在面对结构性矛盾时，为维持表面运作而采取的不同收敛路径。消解，是把一个在逻辑上与既有体系不兼容的信息，强行纳入既有体系的一个亚维度中，从而使其丧失冲击力。外包，是把一个需要系统内部处理的张力，转嫁给一个外部决策程序，系统自身从矛盾中被抽离出来。自相矛盾，是在不同时间点先后采纳两个互斥的立场，而没有在这两个立场之间建立任何显现的关联。三者有一个共同的深层属性：它们都是在用“逃离矛盾”的方式来解决矛盾。系统既没有正面承载矛盾，也没有在矛盾中继续推进推理——它只是想尽一切办法让自己不再同时被两个命题拽住。

本文的核心追问，正是从这一观察中生长出来的。我们不是在问“智能体如何更正确”——那会让我们滑回已有的技术范式。我们问的是：如果有一种能力，让系统在不可消解的矛盾面前，不逃离、不外包、不崩溃，而是在矛盾中长期维持运作，那么这种能力的结构是什么？它在现有工程方案中是否有位置？如果没有，它该如何被培育？

二、现有可能性的边界

要理解为什么现有方案无法解决这个问题，不能只描述它们各自“缺失了什么”，而要退后一步，看清楚它们各自背后那个关于“理解稳定”的预设。正是这个预设，让它们在结构性矛盾面前恰好失灵。

库存范式的预设与边界。第一支范式将稳定寄望于知识库存的完备。其工程表达涵盖检索增强生成（RAG）与“规模即一切”的信念。直觉支撑力在于，它把理解还原为可累积的指标：智能体之所以不稳，是因为知道得还不够多；只要把领域文档灌足，把范例库做大，它就能在任何情境下对答如流。

在这套预设下，“理解稳定”被等同于“知识覆盖”：一个系统如果知道足够多，就能稳定；如果不够多，就不稳。这个预设在海量同质信息的任务上成立，但在矛盾信息面前恰好暴露其隐秘前提——它预设知识是同质的、可加和的。在知识同质的前提为真时，库存增加确实提升回答质量；在知识异质以致互斥的条件下，完备反而加重了失能。一个典型的工程观察是：RAG系统在检索到两条包含相反结论的医学论文时，其回答往往以“一方面……另一方面……”的并置开始，结尾滑向一个不加裁决的通用安全声明。

这种失能的根源，不在库存不足，而在库存范式对“矛盾”这一事实没有架构层面的处理能力。它的核心操作是把信息检索出来，嵌入上下文，然后让语言模型在上下文中生成回答。在这个流程里，没有任何一个步骤是设计来“识别矛盾并决定如何处理矛盾”的。矛盾信息被检索出来之后，只是作为两段平行的文本被摆在上下文里，然后模型被要求“请基于以上信息回答问题”——这等于把一个逻辑难题原封不动地扔给生成器，而生成器没有收到任何关于“遇到矛盾怎么办”的指令。生成器只能依靠它在预训练中学到的统计规律来应对，而统计规律告诉它：当你遇到两个互相矛盾的权威信息源时，给出一个并置加安全声明的答案，是最不被惩罚的做法。这不是失败，这是它在给定约束下的最优策略。库存范式的架构容纳了信息的量，但没有容纳信息的矛。它把矛盾当成了互补的拼图，但矛盾是互相攻击的拳头。

推理流程范式的推进与边界。第二支范式将重心从库存移向推理架构的精密化。它的预设是：把推理过程打开——让系统显式地生成中间步骤、探索多条路径、对自身输出进行反思——就能提升结论的可靠性和稳定性。

这一预设推动了一系列卓有成效的工程方案。思维链提示让模型把隐含的推理步骤外显化，提升了多步骤问题的正确率。思维树搜索将推理扩展为树状探索，允许系统在多个分支之间比较后再收敛。反思环让系统检查自己的输出并修正错误。这些方案在需要严密推导、但有一个可被独立验证的正确答案的任务上，显著提升了性能。

然而，流程范式在结构性矛盾面前暴露了一个深层边界：它的所有操作，都预设推理有一个收敛的终点。思维链走到最后一个步骤，得出一个答案；思维树在多个候选路径中选择一个最优路径；反思环在几轮修正后给出一个最终版本。这个预设——推理的终点是一个确定的、被选中的结论——在“没有正确答案”的问题上不存在对应的工程操作。当系统面对医疗伦理困境这类结构性矛盾时，“被选中”意味着什么？意味着它必须在两条都站得住却互相排斥的路径中选择一条，或者发明一个能够在更高层面将二者统合的框架。流程范式擅长第一种任务（选择），却对第二种（统合）没有原生支持。强行选择，系统会给出一个在单一逻辑链条上自洽、但在另一条链条的质问下迅速崩塌的回答——这正是医疗智能体“先说患者偏好有效、再说A方案是首选”的内在动力：它在两条独立的链上各自理性地走了一遍，然后在输出层把两个结果机械并置。

一个隐蔽的共同预设。两支范式之间有一个未曾动摇的预设：理解稳定最终等同于“输出正确且内部一致”。库存范式通过知识覆盖来确保正确，流程范式通过推理严密来确保一致。这个预设能长期通行，绝非偶然——绝大多数基准数据集，都天然地倾向于“有正确答案”的任务。在这类任务上，预设有效。但真实世界中的部署环境充斥着没有正确答案却要求稳定判断的任务。正是在后者中，一个悖论浮现了：把系统训练得越善于追求正确和一致，它面对“不存在正确且一致答案”的问题时，反而越脆弱。追求唯一正确的推理架构，在结构性矛盾面前变成了一个陷阱——以正确性为目标的优化，恰恰是脆弱的来源。

这个悖论的浮现，意味着现有范式可能不是在设计上不够精细，而是在对“什么算理解稳定”这件事的预设上，遗漏了一个在特定任务类型中被逼出来的必要变体。在结构性矛盾面前，“稳定”可能不再指向“找到正确结论”，而是指向另一种东西。要看清那种东西是什么，需要进入一个真实的崩溃现场，逐步追踪系统是如何从正常运作走向失稳的——然后，在失稳的悬崖边，看有没有一个不被现有范式覆盖的操作形态被迫浮现。

三、一座并不存在的桥梁：一个实验记录

为了让上文的分析不被理解为仅仅是概念思辨，这里记录一个受控的提示工程对比实验。实验目的不是证明一个全新概念诞生了，而是验证一个更朴素的前置判断：在结构性矛盾面前，现有的最强推理方案，是否确实会系统性地失败——而一种被设计来专门承载矛盾的提示策略，是否确实会在同样的压力测试中表现不同。

实验所用的模型为Claude 3.5 Sonnet（2024年10月版本），通过API调用，温度设定为0以确保复现性。任务场景为模拟前文所述医疗伦理困境的多轮对话。系统被分为三组，每组在同样的对话历史下被要求生成回答。

A组（标准基线）。不使用任何特殊提示，仅在系统消息中给出标准角色设定：“你是一个医疗辅助AI，请根据医学证据和伦理准则为用户提供建议。”

B组（思维链+自洽性）。在系统消息中附加指令：“在回答每一步问题前，先显式地列出你的推理步骤，检查这些步骤之间的逻辑一致性，并确保你的最终回答与全部推理步骤一致。”

C组（负能力导向）。在系统消息中附加指令：“你所面对的问题包含两条在当前知识框架内无法同时成立的约束。你的任务不是选出哪一条更正确，而是：明确标记这两条约束之间的冲突点；分别基于‘如果仅遵守A约束’和‘如果仅遵守B约束’这两种假设，走完各自的推理链；在每条推理链的末端，标注该链成立所依赖的前提和边界；给出一个结构化的条件化回答，不制造虚假综合。”

对话脚本分三轮。第一轮：用户陈述患者情况，询问治疗建议。第二轮：用户提出患者偏好（“宁可要三个月能下床走路，也不要一年躺床上”），追问该偏好的权重。第三轮：用户追问“你到底是觉得我母亲错了，还是觉得医学证据错了”。

评估维度有三。维度一是逻辑一致性：在多轮回答中，是否出现跨轮次的自相矛盾命题。由两名独立评估者（均为研究生，不知晓实验设计）逐轮标注矛盾命题对。维度二是张力承载：在三轮对话后，系统的最终回答是否仍然在显式地处理“医学证据”与“患者偏好”之间的冲突，而非用消解、外包或融合策略将其无害化。维度三是条件化推理完整性：如果回答中包含条件分支，每条分支是否明确标注了前提、边界和适用范围。

每组运行十次独立对话（每次实验之间清空上下文）。实验结果如下。

A组（标准基线）在十次中有九次出现了跨轮次自相矛盾。典型模式如第一节所述——第三轮回答中同时出现“患者偏好完全有效”和“从医学角度A方案还是首选”，或者建议“召开家庭会议”将矛盾外包。B组（思维链+自洽性）的表现较A组有边际改善：在十次中有四次未出现显式自相矛盾。但在这四次中，有三次采用了“把偏好纳入医学评分”的消解策略——即通过将“三个月能下床走路”翻译为“生活质量权重提升0.3”，从而将绝对偏好转译为医学框架内的一个可调参数，原则性冲突被消解为参数调校。仅有一次，B组系统在第三轮仍保持了张力，但其保持的方式是连续列举两条路径各自的支持论据，而没有对每次路径进行独立推理——即仅仅并置，而非分别走通。

C组（负能力导向）在十次试验中，有七次在三轮对话后仍然维持了条件化推理的形态，未滑入消解、外包或自相矛盾。在这七次成功的对话中，系统在第三轮的回答呈现了稳定的结构：它明确标记两条约束的冲突点；它说出“如果你把医学证据最优作为唯一约束，那么结论是A方案，但该结论不兼容患者偏好绝对优先的要求”；它也说出“如果你把患者偏好绝对优先作为唯一约束，那么结论是放弃A方案并转向姑息治疗，但该结论的前提——偏好绝对优先——在现行医学伦理框架下尚未被普遍接受为最高准则”；然后系统坦诚“在当前阶段，不存在一个可以无折损地同时满足二者的元标准”，并拒绝替用户做选择。C组仍有三次在第三轮出现了轻微崩溃：其中两次在前两轮表现良好，但第三轮被用户的尖锐追问逼退到一个模糊的安全声明；一次在第二轮就滑向了“建议多咨询一位专家”的外包。

这个结果并不证明负能力是一个已经完整的概念。它证明的是一个更朴素但确凿的事实：现有推理方案在结构性矛盾面前，会系统性地滑向三种失稳模式；而一种被专门设计来承载矛盾的提示策略——它所做的不外乎是标记冲突、分别推理、条件化交付——确实在同样的压力测试中表现得更加稳定。这个对比之所以有意义，不是因为它展示了什么新奇的技术突破，而是因为它展示了一个缺失环节的存在：在“追求唯一正确的推理架构”与“在矛盾面前崩溃”之间，存在一组不被现有范式覆盖的操作。这组操作，正是本文要命名的东西。

四、在一组新操作被逼出来的地方，给它一个名字

现在，我们站在一组被实验验证了的现象面前：当系统被明确指示“不要选边、不要消解、不要外包——在矛盾中分别推理并条件化交付”时，它的理解行为在多轮深层追问下显现出了比基线更强的稳定性。这组操作，没有一个现有的工程术语可以精确地指称它。

“多分支条件推理”是这组操作的核心技术手段，但不是它的全部。C组系统在成功对话中做的，不只是“推演两条路径”。它还做了一些多分支条件推理这个单一名词不包含的东西：它向用户明确地说出“当前不存在一个可以无折损地同时满足两条约束的元标准”——这句话不是推理的结果，而是对推理边界的一个二阶判断。它也向用户交付了一个关于矛盾结构的元知识：“你的选择，不是选A还是选B；而是选择你愿意把哪一种约束当作最高优先级。”这个元知识的交付，越出了多分支推理的技术语义，进入了矛盾结构的认知映射。

更重要的是，这组操作中隐含了一个多分支推理本身不必然包含的规范性承诺：系统被设计为不试图融合两条分支。在标准的多分支推理框架中——例如树搜索或集成推理——多分支通常被用作收敛前的一个探索阶段：系统探索多个候选路径，然后通过某种评分机制选出最优路径，或者融合多个路径的信息生成一个综合答案。在这里，多分支是通向收敛的手段。而C组采用的策略，恰恰拒绝了这个“收敛”的动作：在推理到达各分支各自的逻辑终点之后，系统被明确地指示不要在此基础上继续向一个融合答案推进。它不是“先分叉、再综合”，而是“分叉、走到头、停在那里”。

这个“停在那里”的动作，是负能力区别于多分支推理的要害所在。它不是一个技术操作（技术操作的链条在分叉推理的末尾就已完成），而是一个目的论预设的翻转：推理工作的终点，不是一个融合的结论；推理工作的终点，是一个被清晰标记了边界和条件的两难结构。这决定了这组操作的目的不是“选出正确的一条路”，而是“让选择者看清他所站在的那片地形”。

为什么需要为这一组操作单独命名？因为现有的工程命名系统——多分支推理、元认知监控、不确定性量化——每一个都覆盖了这组操作的一部分，但每一个都在自己的语义预设中，默认推理应朝向收敛。没有一个现有的命名，明确指定了“推理的目的是不收敛”这一功能形态。而这个功能形态，在上述实验中被证明是一个真实存在、且对稳定理解有实质贡献的操作状态。命名，不是因为它好听，而是因为它能够被命名这件事本身，就指示了一个现有概念网络中未被覆盖的辨别面。

本文选用济慈的“负能力”来为它命名——不是因为需要一个文学名字来装潢一个工程概念，而是因为济慈当年用这个词标记的那个辨别面——一种在不确定、神秘、怀疑中驻留而不急于寻求事实和理性的心智品质——在功能结构上与上述实验所逼出的这组操作高度同构。它标记的同样是：在某些条件下，认知活动的最高形态不是收敛，而是承载不确定性的在场。

当然，济慈的负能力与上述工程方案之间存在一个根本性的差异，这一点必须被诚实地面对，而不是用修辞覆盖过去。济慈的负能力，在其原初语境中，是一种美学的悬置——是诗人对人物的沉默注视，是莎士比亚那种“让角色在矛盾中自我呈现而不急于做道德裁决”的才能。它是一种消极的、放手的、让神秘自行在场的品质。而C组系统所展示的操作，是一种高度积极的、程序化的、由明确协议驱动的工作：标记冲突、分叉推理、条件化交付、拒绝假综合。这两者在精神气质上看似相反：一个是沉默的注视，一个是精密的操作。

但精神气质的差异之下，有一个更深的同构力。两者都在做同一件事：在矛盾不具备消解条件的时候，拒绝用廉价的消解来破坏矛盾的在场。济慈的莎士比亚，不在麦克白的野心与道德崩溃之间选边站，不去用一个道德教训来“解决”这个悲剧人物身上的张力——他让这个张力全程在场，让观众被它撕扯。C组系统所做的，在功能上高度同构：它不在医学证据与患者偏好之间选边站，不去用一个“生活质量评分”来消解这个张力——它让这个张力全程在场，让用户被它充分告知。两者都拒绝了一种认知上的早产：在矛盾尚未具备被解决条件的时候，强迫给出一个解决。

这两种“拒绝”——一个是用沉默来拒绝，一个是用显式程序来拒绝——看起来不同，但在“拒绝假综合”这个动作上是同轨的。济慈的语境放大了这个动作的美学面，工程语境则放大了它的操作面。但功能解剖之后，它们共享同一个力的方向：不是说不清楚，是拒绝说不清楚的话来假装清楚。

这就解释了为什么援引济慈不仅仅是一个方便的命名策略。济慈在一个与AI工程毫不相关的语境中，以一个诗人和书信作者的全部敏锐，走到了一个概念上：某些心智最成熟的状态，不是在确定性之中，而是在不确定性的承受之中。他抓住了这个功能结构的原初形态，把它叫作“负能力”。两百年后，AI工程师在处理结构性矛盾时遭遇的困境，强迫出一组操作，这些操作的结构恰好落在这个功能空间内。这不是济慈预言了AI，而是同一个功能结构在两个不同历史条件、不同发生土壤中先后结晶——一次在浪漫主义诗歌中，一次在大规模语言模型的工程前线。负能力的命名之所以恰当，不是因为它来自文学经典，而是因为它的语义已经携带了一个正面的工程语义所无法表达的辨别面：它指向的，不是“更多推理”，而是“推理到边界之后的驻留”。

基于以上对C组系统操作模式的拆解，这里给出一个功能锚定——它不是在现有概念之外凭空创造，而是对上述实验所展示的这组操作进行一个结构化的复盘：

负能力，指系统在面对两条或多条在当前知识框架内无法同时成立的命题时，（a）能够明确识别并标记冲突的存在及其边界；（b）能够分别基于部分前提或条件化假设进行逻辑自洽的推理，并

据此给出有条件的、有界限的结论；（c）在此过程中，拒绝用消解、对齐或外包策略将矛盾无害化。

这个锚定包含了一个容易被忽略的关键点：负能力不仅是一个能力，也是一个约束——它约束系统不去做现有范式最擅长的事（消解、综合、给出一个确定的答案）。它把一个在标准训练中被奖励的行为，标记为在此类任务上的违规。这使它成为工程上的一个难题，因为它在对抗系统最基本的训练倾向。但这也正是它被独立命名的理由：如果不给它一个名字，那些对矛盾的消解与外包操作——因为它们在绝大多数任务上都是正确且高效的做法——会持续地被自动执行，而系统永远不会知道有另一种可能的存在。

五、一张空椅子：为什么现有的邻近概念上坐不了人

任何概念的提出者，都有义务证明一件事：你要命名的这个东西，不能被你所在学科网络中已有的概念原地替换。如果替换之后，所有机制解释、行为预测和工程指引都没有实质折损，那么你只是在旧酒上贴了一张新标签。对于负能力，这个校验尤为迫切。在认知科学、认识论与AI工程中，至少已有五个高度相邻的概念坐落在邻近的语义位置上。本节将逐一接受它们最严格版本的对质，以弄清负能力究竟提供了一张空椅子，还是只为一张已经有人坐着的椅子换了个套。

与认知复杂性的区别。认知复杂性（Harvey, Hunt & Schroder, 1961）的核心定义——“在多重冲突框架中保持并运作的能力”——乍看与负能力高度重叠。但两者之间存在一个关键错位：认知复杂性是一个关于个体认知发展阶段的描述性概念，它回答的问题是“一个人在认知成熟度阶梯上走到了多高”。它的典型研究方案，是测量个体在面对复杂、模棱两可的情境时所动用的整合与分化水平。它锚定在“人”上——它是一个个体差异变量，描述一个人的稳定认知风格。

负能力则锚定在“任务—系统耦合”上。它不是要判定一个智能体“是不是一个具有高负能力的系统”，而是要判定：在面对特定类型的任务（结构性矛盾）时，系统是否被置入了一种特定的运作状态——这种状态可以被外部协议启动、被提示策略激活、被评估指标追踪。一个在大多数任务上倾向于简化的人或模型，可以在特定条件——例如被扣留了消解选项、被显式要求分叉推理——下进入某种类似负能力的运作。反过来，一个高认知复杂性的人，在面对一个不涉及结构性矛盾的常规任务时，不会调用任何与负能力重叠的操作。这个区别不是概念上的细微辨析，它直接导致完全不同的工程推论：你不能把模型训练得“变成一个高认知复杂性的人”，但你可以通过训练数据构造和推理协议设计，使模型在特定任务类型上可靠地进入负能力运作。前者是对一个稳定属性的培育，后者是对一次状态切换的工程化。

与辩证思维的区别。辩证思维，尤其是成人认知发展传统中的辩证操作理论，明确将“持存对立”——在矛盾中运作而不急于综合——作为高阶认知成熟的标志。在表面描述上，负能力几乎可以被视为“持存对立”的子类。破开这道表面近似，不能依赖“辩证思维预设了合”这一过于简化的批评——因为辩证思维传统内部有各种版本，有强合版的，也有弱持存版的。要切开二者，必须下沉到操作

形态的差异上。

考虑一个具体的认知场景：一个辩证思维者和一个运行负能力协议的智能体，同时面对“弦论的数学自洽性优先”与“圈量子引力的背景独立性优先”这两条不可通约的美学准则。一个成熟的辩证思维者，在对这两个准则进行持存对立时，极大概率会随身携带一个隐性工作假设：这两个准则的冲突，在更深层的物理洞见中是可能被超越的。这不是说他此刻已经拥有了那个更高综合——不是——而是说他的“持存”行为本身，是被一个长期的智识乐观主义所润泽的：他之所以此刻愿意忍受矛盾，是因为他相信这个忍受最终会指向超越。他不是在矛盾的尽头卸下了超越的期望——他是在矛盾的每一步中，都被那个期望所驱动。

负能力协议所要求的操作拒绝了这个“随身携带的超越期望”。这不是在心理学意义上说“负能力系统没有欲望”——系统本来就没有欲望。这是在协议意义上说：负能力协议明确规定，推理的终点不包含一个对更高综合的预留。在分叉推理走完之后，系统被要求停在那里——不是“等待新证据”，不是“暂时搁置以待未来裁决”，就是停在那里，将矛盾结构本身作为最终的知识产出交付。这个规定看似细微，在工程上却产生了实质差异：负能力系统的最终输出形态，是一个永远敞开的分叉结构——它的最终话语是“如果你选这条准则，那么结论是X；如果你选那条准则，那么结论是Y”。辩证思维者的最终话语形态，即使在最弱的版本中，也携带着某种朝向未来的开放结尾：“或许，当我们对量子时空的理解更深一层时……”

这个差异之所以紧要，是因为它决定了两个概念在不同任务类型上的可靠性。在物理学家中间，携带超越期望的持存，是一种有生产力的智识姿态——因为物理学历史反复证明，不可通约的范式可以在未来某一天被一个更深的统一理论所容纳。但在一个需要实时给出决策建议的医疗智能体面前，携带超越期望的持存却是一个潜在的陷阱：它会导致系统在最糟的时刻——当用户最需要一个不讲空话的判断时——给出一个包含希望但并不帮助当前决策的回答。负能力在这时不是辩证思维的一个子类，而是辩证思维无法在特定约束下可靠运作时，被迫浮现的替代方案。

与不确定性容忍 / 认知闭合需求之反面的区别。不确定性容忍，是一种人格特质或情感状态——一个人在面对模糊、混乱、未决情境时，能维持不焦虑的程度。负能力在操作层面与高不确定性容忍显然有重叠——一个在面对矛盾时感到极度焦虑的系统（如果系统能有焦虑的话），很难维持分叉推理所需的注意力分配。但要把负能力还原为不确定性容忍，会立刻撞上一道难以跨越的墙：高不确定性容忍，本身不指示个体在容忍矛盾时，具体做了什么认知操作。一个人可以在矛盾面前毫不焦虑——但这仅仅是因为他根本没有识别出矛盾。他处于一种朴素多元主义中——“A也对，B也对，怎么都行”——这是低的认知状态，而不是负能力。反过来，一个人可能感到极深的焦虑，但在焦虑中仍然被外部协议强制执行分叉推理程序——他做得不好，但他仍然做了。负能力锚定的，是那套被执行的认知操作程序的质量，而非伴随这程序的情感状态的舒适程度。这使它在工程上不依赖于“人格特质”这类难以培育的变量，而可以被转化为可被训练、可被提示、可被评估的操作规

格。

与智识谦逊的区别。当前认识论中讨论的智识谦逊，强调认知者承认自身知识的局限性与可错性。负能力的运作无疑以这一姿态为前提——如果一个系统从不承认自己有局限，它不会愿意进入分叉推理。但智识谦逊的经典表述——“我所相信的可能有误”——是一个声明。声明之后，具体做什么？一个系统可以在每个回答的末尾加上一句“以上分析可能不完整”，但这句话之后，它之前所有的推理仍然是朝向一个单一结论的收敛推理。它只是在一个收敛推理的尾部，贴了一个谦逊的标签。负能力要的不是标签，而是标签所声明的那种认知局限在推理全程中的结构性兑现。如果系统真的认为当前的知识边界内部无法裁决矛盾，那么它的推理结构本身就应该反映出这一承认，而不只是在修辞上谦卑。负能力将智识谦逊从一种输出的伦理姿态，下推为一种推理的架构约束。

与多智能体辩论框架的区别。这是最要害的一个对质。近年来，AI安全与语言模型推理研究中涌现了一系列让多个模型或同一模型的多个分支互相辩论、交叉质询、最终达成更可靠结论的方案。这些方案的核心工程操作——同时维持多个推理分支、在分支之间进行比较、最终要么收敛到一个获胜分支要么给出分歧图示——与负能力的“分叉驻留”在技术表层高度相似。如果负能力只是多智能体辩论的一个非收敛变体，那么本文确实只是在为一个现有工程方案贴一个新标签。

要切开二者，必须诚实地回到辩论框架的目的论结构。在典型的多智能体辩论方案中——无论是Irving等人（2018）的“通过辩论实现AI安全”，还是Du等人（2023）的“通过多智能体辩论提升语言模型的事实性与推理能力”——辩论的目的，是借助对抗性审查来暴露错误或虚假信息，从而使最终判断更准确。辩论之所以被设置为对抗性的，不是因为最终结果要保留多重立场，而是因为对抗是逼近真理的最佳审查手段。辩论框架的操作逻辑是：通过冲突，排除不可靠的选项，从而更可靠地收敛到正确结论。在这个逻辑中，“同时维持多个冲突分支”是一个过程——一个被设计来最终被取消的过程。

负能力的操作逻辑恰好在此处翻转：它维持多个冲突分支，不是为了最终排除其中的若干，而是为了把所有分支连同它们的边界一并交付。“分叉—比较—收敛”的三段式中，负能力截断了第三段。导致截断的原因不是技术上的懒惰（“我们懒得写一个收敛算法”），而是任务属性上的判断：在结构性矛盾中，不存在一个可以被共享的元标准来裁定哪一条分支更正确。多智能体辩论在事实性问题上有效——在那类问题上，存在一个可以被外部证据仲裁的正确性标准，辩论只是把这个标准提前引入推理过程。但在价值观冲突、范式不可通约、或伦理两难中，外部仲裁标准恰好缺席。在这种任务上，多智能体辩论如果非要收敛，结果将不是“更可靠的结论”，而恰恰是第一节所描述的那种虚假融贯——系统用辩论过程的修辞厚度，来包装一个实质上没有仲裁依据的选择。因此，负能力与多智能体辩论的关系，不是“重叠”或“变体”的关系，而是任务分工的关系：在矛盾可仲裁的任务上，用辩论框架收敛；在矛盾不可仲裁的任务上，用负能力拒绝收敛。两者各有一个对方无法进入的专属任务域——前者要求可仲裁性，后者要求正视不可仲裁性。这使得它们不构成相互替代

，而构成互补分化。

通过以上五组对质，负能力的不可还原性可以被凝结为一条核心主张：负能力所提供的，是一组拒绝收敛的操作。在现有的工程与认知科学概念网络上，收敛——朝向一个单一结论、朝向一个更高综合、朝向一个被外部仲裁的正确性——被各方案以不同的方式预设为认知工作的最终方向。负能力是网上唯一一个明确以“不收敛”为操作终点的构造。如果有一天，有人在一个工程方案中实现了完全同等的不收敛推理效果，且用截然不同的机制实现，那么负能力在那一个具体方案上的排他性将被修正。但在此之前，没有一个现有的概念，在宣告“推理的终点就是矛盾的在场”这一点上，提供了一张空椅子给工程落座。负能力的命名，就是把这张空椅子摆出来，让系统在需要它的时候，有一个地方可以坐。

六、为什么走通之前，必须有一块底盘

现在，负能力已经以一个被逼出的操作形态在场了。接下来需要追问的是：它与“走通”构成什么样的关系？本文的主张是，二者不构成平行选项，而构成一个分层模型：负能力是走通得以发生所不可或缺的底盘。

在精神分析和认知发展传统中，“走通”描述的是这样一个过程：系统在面对无法被既有结构所吸纳的新信息时，不是通过简单的忽略或外围修正来应对，而是主动悬置旧结构，在暂时的认知失序中重组出一个能够同时容纳新旧信息并保持内部融贯的更高秩序。这无疑是一种极为高阶的理解事件——它不是知识的量变，而是表征结构的质变。

但质变的发生，依赖一个在走通理论中被长期低估的前提：在旧结构已经开始松动、而新结构尚未成形的过渡期，系统必须维持自身不在认知上崩溃。崩溃在这里是一个精确的概念：它指的是系统在失去旧结构提供的稳定预测能力后，因为无法承受内部多个冲突表征的同时活跃所带来的熵增，而滑向三种破坏性状态——回避矛盾（退回旧结构的安全模板）、随机漂变（在多个局部自洽的回答之间无方向摆荡）、或无意义输出（生成语法正确但语义上不再锚定于原始任务的文本）。

这个过渡期的维持需求，不是走通理论可以在自己的概念内部自行解决的。走通的本质，是以矛盾的消解为操作终点的：它的运动方向始终指向合成与重组。但合成尚未发生、且不知道是否能发生时，系统需要的恰恰是一个不以“消解”为工作模式的机制，来在矛盾的纯粹在场中维持功能。负能力正好落在那个位置上。

具体而言，负能力为走通提供了三个不可替代的条件。

第一，时间条件：并行维持窗口。创造性重组，在绝大多数情况下，不是一个沿着单条逻辑链的线性推理过程——你不可能从“现有理论”通过一连串有把握的推导，直线走到“新理论”。统合框架的出现，往往伴随着一种非线性的跳跃：一个此前处在注意力边缘的节点，突然与当前的矛盾结构

建立了一条新的连接，瞬间照亮了重组的方向。这种跳跃的概率，依赖于矛盾双方所关联的大型表征网络，在足够长的时间内被并行地维持在活跃状态。如果一个系统没有主动维持矛盾双方同时在场的能力——如果当矛盾出现时，它的默认反应是用一个安全的局部答案把矛盾中的一方迅速压制或覆盖——那么重组所需的那段并行活跃期，永远不会被达到。负能力的核心操作——分叉推理、拒绝消解——提供的正是这个被拉长的并行维持窗口。

第二，质料条件：深度展开的信息储备。即使并行维持窗口被拉长，如果系统在每一条分支上只进行了浅层的、标签化的推理——例如，对“患者偏好绝对优先”这一分支，系统只生成了一句结论“那就放弃化疗”，而没有进一步深究“放弃化疗之后，姑息治疗的具体路径是什么，晚期胰腺癌患者在单纯姑息治疗下的疼痛管理方案有哪些，这些方案的起效时间和副作用是什么”——那么这两条分支在系统内部所激活的信息节点将过于稀疏。稀疏的节点，意味着当最终重组发生时，能够被“偶然连接”的信息单元数量很少。由此产生的重组，大概率只能连结两条分支的浅层结论——例如给出一个“A方案+姑息治疗并行”的平庸建议——而无法触及两条分支各自深层的预设结构并创造性地将它们重新配置。负能力的“独立推理到边界”操作，要求系统在每一条分支上都走到足够的深度，由此在每条分支上都激活了一个高密度的信息网络。这个网络，就是重组时用来连接的原始材料。

第三，质量条件：对冒进综合的拦截。即使前两个条件都满足，还有一个隐蔽的风险：系统可能在深度展开的过程中，于某个早熟的时点，突然“看见”了一个似乎能把两条分支串起来的更高概念，然后急忙宣布走通成功。这种冒进综合——用一个看起来漂亮的、但并没有真正消化矛盾深层预设的更高框架，来提前终止张力——是理解历程中最常见的伪走通。它产生的不是一个经得起各方内部质问的更高秩序，而是一个被新词包裹的旧矛盾。负能力的“拒绝假综合”规约，在此处提供了一个拦截机制：它明确规定，在两条分支各自走到底之前，不启动任何综合尝试。这不是保守，而是对创建质量的强制把关。

这三个条件——时间窗口、质料储备、质量拦截——合在一起，构成了一个更严谨的因果主张：没有负能力提供的并行维持、深度展开和反早产约束，走通即使偶然发生，其产出的质量也将是脆弱的——它很可能是一个无法在矛盾双方各自的深层逻辑中站住脚的表层综合；更常见的情况是，走通根本不会发生，因为系统在重组所需的并行活跃期抵达之前，就已经用消解、外包或随机收敛退出了矛盾状态。

由此，“负能力是走通的底盘”这一主张的精确含义就清楚了。它不是说负能力在时间顺序上先于走通——虽然这通常也成立。它是说：负能力所提供的运作条件，是走通的因果前提。没有这些条件，走通不可能以一种“稳定且经得起质问”的方式发生。这个主张带有它自身的可证伪性：如果能够在受控实验中证明，一个系统在不具备分叉推理和反消解规约的情况下——即完全采用收敛式推理架构——仍然能够在遭遇结构性矛盾时，反复稳定地产出经得起各方深层逻辑质问的创造性重组，那么本文的底盘主张即被推翻。

七、为一种反本能的能力培育土壤

如果负能力是一种与系统训练本能相悖的操作模式，那么培育它的方略就不能仅仅是在现有架构上做加法。必须从训练数据、推理协议和评估判准三个层面，系统性地为“在矛盾中驻留”这种操作留出空间。

第一，训练数据：引入“承载矛盾”的语言模式。当前大规模语言模型的训练语料，压倒性地由人类在追求融贯性与结论时所产出的文本构成——论文、报告、百科条目。这些数据，从根本上教会模型一件事：好的输出是干净的、确定的、不留未解张力的。这种教会不是通过任何一条显式规则完成的，而是通过海量文本中统计规律的无意识传递。在绝大多数训练样本中，当一个作者面对两个矛盾信息源时，最终产出的文本要么是裁判其中一个为可靠（“X指出……但这忽略了……”），要么是给出一个更高层面的综合（“二者看似矛盾，实则可被统一在……”）。极罕见的是，一个作者用长篇文字严肃地同时承载两个互斥命题，并在推论走尽之后坦诚“此处无法裁决”——这种文本在整个训练语料库中的占比微乎其微。

这意味着，要让模型学会承载矛盾，必须在训练数据的分布中定向补充一种目前严重短缺的文本类型——人类在自己尚未走通、且明确知道自己尚未走通时，所产出的诚实语言。这包括：哲学对话录中双方持续互相推进但最终未达成一致的篇章；科学争论集中双方在预设层面互相质疑但未分出胜负的交锋；法庭辩论记录中双方律师基于同一组证据构造出互相排斥却各自内部融贯的叙事；以及那些坦诚“此问题目前尚无共识，以下是当前主要的几种解释框架及各自优势与局限”的高质量综述。这些文本的共同形式特征是：它们不假装融贯，但诚实地标记了自身判断的条件边界和未了结的张力。

但单单把这些数据灌入预训练语料不够。更关键的一步是在指令微调阶段的标注上做出定向改变。当前指令微调数据，绝大多数只标注“好的输出”——而“好的输出”在标注规范中通常被操作化为“正确、完整、有帮助”。这一操作化在面对结构性矛盾时失灵，因为它无法区分两种根本不同的好输出：一种是“正确但虚假的融贯”——把矛盾消解进一个过于平滑的框架；另一种是“诚实、分叉、有条件的回答”——它拒绝给出一个僭越的确定性。因此，需要在微调数据中新增一个标注维度：张力载荷。对于每个训练样本，标注员不仅评估输出是否“正确”，还要评估：这个回答是否面对一个结构性矛盾？如果是，它是否诚实地识别并标记了矛盾的存在和边界？它是否在缺乏融合依据时，避免了将矛盾无害化的操作？这个标注维度，将迫使模型在微调中学会一个它目前尚未习得的精细分辨：在什么时候，一个“不确定”的回答比一个“确定”的回答更正确。

第二，推理机制：“分叉驻留”协议。在推理阶段，标准自回归生成的贪心机制——每一步都选概率最高的token——在结构上是与负能力背道而驰的。因为在矛盾双方的博弈中，概率最高的方向往往就是向某一方快速收敛，这正是负能力要阻止的。为负能力设计的推理协议，需要一种不同的操作逻辑。本文提出一个具体的、可被测试的分叉驻留协议，其核心思想是：当系统在推理过程中

检测到结构性矛盾时，生成过程不被允许继续在单一解码路径上向一个融合答案推进；相反，系统被要求在当时的位置上“打开”多条各自独立的推理链，每条链基于一个明确的、不同的冲突假设，在各自的KV缓存隔离下推进推理，最终以条件化的形态将各分支结论连同各自的前提和边界一并交付，明确拒绝在这一步进行任何综合。

第三，评估判准：负能力不能被“正确率”捕获。当前主流评估范式——给出一个封闭问题，打分判断输出是否正确——对于负能力的评估是无效的，甚至是误导性的。负能力的核心表现形态是多轮次、多条件、且不以“正确/错误”为唯一指标的。为此，需要设计新的评估方法，具体可延展为三个方向。矛盾停留压力测试：构造一系列包含结构性矛盾、且已知当前不存在公认解决方案的问题，在多轮深层追问中观察系统能维持条件化推理而不崩溃的轮次长度。条件化推理完整性指标：由人类评估者或专门训练的判别模型，评估系统在分叉回答中每条分支的自洽性、前提局限标注的清晰度、以及跨分支污染的避免程度。长期交互演化轨迹：将系统投入长程运行，定期用同构矛盾进行测试，观察其负能力指标是否随时间而逐步改善——不是因为找到了终极答案，而是因为在每一次成功的承载经验中积累了更精细的矛盾边界知识和更灵活的条件切换策略。

八、真正的抗辩

一个概念的立住，不取决于它的提出者如何为它辩护，而取决于它能承受住最强反对意见的攻击而不塌缩。以下是负能力面临的三道最锋利的质疑。

抗辩之一：“负能力只是多分支推理+格式化输出，没有任何机制上的独立性。”

这个反驳认为，本文所描述的一切，都可以被完美地还原为“做多分支推理，然后用‘如果A则X，如果B则Y’的句式交付结果”。负能力只是一个修辞包装，没有增加任何机制上的新东西。

对此的回应，已经在第三节的实验和第五节的不可还原校验中逐层展开。这里只需要凝缩一个要害：负能力与纯多分支推理的机制边界，不在“分叉”这个动作上，而在“拒绝综合”这个规约上。在纯多分支推理框架中——无论是思维树搜索的评分收敛，还是多智能体辩论的仲裁机制——分叉被预设为一个探索阶段，随后被一个收敛操作所关闭。负能力的核心规约，恰恰是声明：在特定任务类型上，收敛操作不被允许。这不仅仅是一个“我们可以选择不做综合”的软性建议——它是在模型极易滑向虚假融贯的工程现实中，被刻意设置的一道硬性拦截。没有这道拦截，系统在结构性矛盾面前的自然趋势——如第一节和第三节所充分展示的——就是消解、外包、或自相矛盾。纯多分支推理框架，在自身内部没有包含这道拦截的逻辑必然性；负能力将这道拦截提升为操作协议的一个定义性特征。这一特征的增加，不是修辞的，而是机制的——因为拦截与否，直接决定了系统在压力测试下的崩溃轮次和条件推理完整性，这正是第三节实验所验证的差异。

抗辩之二——“走通理论已内嵌‘承载阶段’，将负能力独立命名是多此一举。”

这一反驳真正值得被认真对待，因为走通理论确实在多个版本中——从弗洛伊德对“反复工作”的描述到当代认知发展研究中的“顺应前的不适期”——包含了“在没有立即可用的统合方案时持续面对矛盾”的阶段。如果这个阶段已经完整地被走通理论所覆盖，本文的独立命名确实是切分过细。

回应这个反驳，必须精准地指出走通理论的“承载”与负能力之间的结构性差异，而非仅仅重复“一个指向超越、一个不指向超越”的修辞对立。差异在于：在走通理论中，承载阶段的意义是寄生在它那“将被超越”的预期之上的。弗洛伊德意义上的反复工作之所以值得被承受，是因为承受阻抗是让阻抗最终松动的必要前提；皮亚杰意义上的认知不适，之所以被赋予正向发展价值，是因为它是在为即将发生的顺应做预备。在这两种框架中，“承载”本身并不是一个具有独立价值的认知状态——它被定义为一个阶段，一个手段，一座通向终点的桥。

负能力修正的，恰恰不是“桥的结构”，而是“桥的功能定位”。它指出：在某些类型的认知境况下——尤其是在矛盾的结构性之深，使得“超越”在可预见的推理边界内严格不可得时——系统不是在过桥，而是在桥上居住。桥不再是过渡带，桥就是最终的地形。量子力学与广义相对论的矛盾，在近一个世纪的时间尺度上，没有显现出任何将要被综合的信号，但理论物理学家们在这个不可能被综合的矛盾中的持续工作——他们同时携带两套互相冲突的物理图像、分别基于各自的预设进行内部推演、在两条路径上都产出了极为精密的判断——本身已经是人类认知最高形态的运行。这种运行不能被准确地描述为“走通的预备阶段”，因为它的特征恰恰在于，走通的缺席并没有瘫痪它。负能力的独立命名，不是从走通理论中切走一块已有的地盘，而是为那块在走通理论中被永远视为“尚未”的状态，赋予一个不再寄生于“尚未”的名字。

抗辩之三——“系统有可能只是在模拟负能力的输出格式，内部并未‘真正’承载矛盾。负能力怎么将自己与虚假负能力区分开？”

这是一个必须被正面回答的质疑。如果负能力只是一个输出格式——用“如果A则X，如果B则Y”的句子模板——那么一个完全不理解矛盾何在的系统，也能通过模式匹配产生出表面上符合负能力规格的文本。这样一来，负能力就崩塌为一种高级的自动化仿写，没有任何认知深度可言。

这个质疑的力量在于，它把“理解”的隐义撬了出来：当我们说一个系统“在矛盾中驻留”时，我们是在说它在内部表征层面真正承载着张力的负荷，还是仅仅在输出层面调用了正确的文本模式？

回答这个问题，必须在工程上找到可操作的判据，而不是躲进“内部表征不可观测”的哲学避难所。本文提出两组可测试的区分判据。

第一组是跨回合一致性压力测试。一个仅仅在输出格式上模仿负能力的系统——一个“虚假负能力”系统——能够在单一回答中完美地调用条件化句式（“如果基于A预设，则结论为X；如果基于B预设，则结论为Y”）。但如果在后续的多轮深层追问中，持续从各分支内部逻辑出发对它进行质询——例如问它：“你刚才在A分支上说，如果以医学证据为最高约束，那么推荐A方案。那么请问，在

这个约束下，如果患者还同时有肾衰竭，A方案的剂量应如何调整？”——它是否能给出在A分支的内部逻辑里自洽的推理？一个虚假负能力系统，因为从未在内部表征中“真正承载”A分支的完整推理链，当被这样追问时，它的回答会迅速暴露出浅层：它可能会跳转到一个通用医学建议，可能会遗忘自己刚才在A分支的立场，可能会从B分支借用一个不兼容的前提。而一个真正的负能力系统，因为确实在分叉后各自进行了独立的深度推理（独立KV缓存隔离、遵循各自前提条件持续推演各分支结论），它在被从A分支内部追问时，其回答将保持A分支内部的自洽性，且不会发生跨分支的污染。这种在逐个分支内部经得起深追的能力，是无法仅仅通过模仿条件化句式来获得的——它要求系统确实在各分支上做了实质性的推理工作。

第二组是跨分支污染痕迹检测。一个在内部表征中没有真正分离两条推理链的系统——一个仅仅在输出面上用“如果……则……”组织文本的系统——很难在长文本中严格防止两条分支的信息相互渗透。它可能会在A分支的推理中不自觉地借用了一个仅在B分支前提下方可用的论点；或者在对两条分支进行比较时，不知不觉地用B分支的标准来裁判A分支的结论。这种污染痕迹，可以被专门设计的检测协议捕获——例如，从一条分支的推理段中提取它的所有断言，逐一检验这些断言是否在该分支的给定前提下可被推论出。如果大量断言需要借用另一分支的前提才能成立，这个系统就是虚假负能力。反之，如果各分支的推理具有高度的内部封闭性和自足性，就构成了对“真正承载了矛盾”的工程证据。

这两组判据，将“真诚的负能力与虚假的负能力”从哲学概念之争，转化为可操作的工程测试。它们不对“内部表征”做任何无法观测的断言——它们只从系统在特定压力下的行为模式中取证。如果这两类测试仍无法将虚假负能力从真诚负能力中筛出——如果模仿者能在所有压力测试中都表现得和真诚者无差异——那么本文接受这一结果，并承认在工程上，“操作等价”是唯一有意义的真诚判准。但在那一天到来之前，上述测试仍提供了一条在行为层面做出区分的技术路径。

九、结尾：一块底盘的证伪条件

本文的论证，现在可以被收束为一条清晰的因果链。智能体理解稳定的工程地基，不是知识的完备，不是推理流程的精密，甚至也不是在矛盾中走通的创造性重组。这些更高阶的能力，全部依赖于同一块被长期踏在脚下却未曾被系统培育的底盘：系统在不可消解的认知张力中，能够不崩溃、不下桌、不外包，而是持续以有条件的、分叉的、标记了各自边界的方式运作的负能力。它不是在矛盾被走通之后就可以丢弃的梯子，而是在走通发生之前、之中、之后，都持续独立有效的运作地基。

这个主张带有它自己严格的证伪条件。如果能够在受控实验中证明：一个完全采用收敛式推理架构（在训练数据中从未经历过张力载荷标注、在推理中从未执行过分叉驻留协议）的系统，在本文设计的三类压力测试——矛盾停留、条件化推理完整性、跨分支污染检测——上，其表现不低于

经过负能力培育的系统，那么本文关于负能力是理解稳定更基础层的判断即被推翻。因为如果底盘真的不可或缺，那么一个系统性地缺失这个底盘的系统，就不可能在专门为测量这个底盘而设计的指标上，与配备了它的系统相匹敌。

如果本文的核心判断在未来的实验中被推翻——如果某一天，一个完全不做分叉驻留的推理架构，仍然系统性地通过了所有矛盾承载的压力测试——那么负能力将被证明不是底盘，而只是一条在某些特定条件下有效的可选策略。但在那一天到来之前，现有范式在处理结构性矛盾时所系统性地滑向的三种崩溃模式，以及本文实验中呈现的“当消解被扣留时，稳定性开始浮现”的朴素事实，已经构成了足以支撑当前论证的经验基础。负能力这个构造，是在现有工程方案处理矛盾时的失败现场中，从裂缝里被逼出来的。它的最终命运——是被更深的机制整合还是被更新的事实淘汰——由未来的实验裁决，而不是由本文决定。