© 2026 IEEE。允许将本材料用于个人用途。所有其他用途均须获得 IEEE 的许可，无论是在当前还是未来的任何媒体中，包括为广告或宣传目的重印或再版本材料、创作新的汇编作品、转售或重新分发至服务器或列表，或在其他作品中重复使用本作品的任何受版权保护部分。录用版本：V. Dziwniel 和 T. Lu，《国际象棋人类行为克隆的时间条件化》，2026年IEEE游戏、娱乐与媒体会议（GEM），德国柏林，2026年，第1–6页。https://doi.org/10.1109/GEM72276.2026.11747060

国际象棋人类行为克隆的时间条件化

Victor Dziwniel 和 Tianxiang Lu，IU 国际应用科学大学，
德国埃尔福特。电子邮箱：victor.dziwniel@iu-study.org，
tianxiang.lu@iu.org

arXiv:2611.33106v1 [cs.AI] 24 Nov 2026

摘要——国际象棋中的行为克隆模型通常根据棋盘状态预测人类棋步，却忽略剩余时钟时间等时间输入。然而，时间压力是影响人类决策质量的主要因素，而时间状态是人类行为建模中缺失的一个维度。

本文介绍了一种时间感知型 Transformer，它在 Chessformer 风格的架构基础上扩展而来，采用学习得到的时间条件化词元来编码棋手时钟状态、用时控制和近期节奏，并配有用于预测思考时间和认输的行为头。该模型基于大规模在线国际象棋数据训练，并与无时间因素基线及最先进水平进行评估。

结果表明，时间条件化虽能小幅提升整体准确率，但在时间压力下带来的提升显著更大。除了走法选择外，该模型对思考时间和认输的预测也更加准确。更重要的是，该模型在关键时间区间内更好地匹配了人类的棋步质量下降情况，减少了无时间因素的行为克隆模型中系统性地高估棋手水平的问题。

这些结果表明，时间状态是人类国际象棋行为克隆中一个此前未受充分利用、但十分重要的信号。

索引词—**行为克隆、国际象棋、Transformer、节奏建模、人类—人工智能对齐**

I. 引言

大多数国际象棋引擎的设计目标都是尽可能强地对弈。另一个不同的目标是像人类一样下棋：复现特定水平棋手会选择的着法分布，包括他们的失误。这类模型可用于打造逼真的陪练对手和机器人，也能驱动教练工具，向学习者展示同等水平棋手会走出的着法；此外，还可用于作弊检测系统，标记强得不像人类的对局。从棋手的角度来看，可信度与纯粹的准确性同样重要：感觉像人类的对手能让棋手保持投入，而下棋像机器人的对手则会破坏体验。

国际象棋中的行为克隆范式最初由 Maia [1] 通过按等级划分的卷积网络确立，随后统一为单一的技能条件化模型[2]；Allie [3] 转而采用基于文本的Transformer处理棋步记录，并结合搜索；而 Maia-3 [4]，基于Chessformer架构，重新回归空间表示，将64个棋格作为仅编码器Transformer中的词元。然而，它们都舍弃了时钟。剩余时间不会对策略形成条件约束，因此这些模型都无法学习人类的走法选择如何随着时间耗尽而变化。

关于国际象棋中速度—准确性权衡的研究表明，时间限制会改变识别式与基于计算的下棋方式之间的平衡，从而改变人类可能选择的着法 [5]。随着时钟所剩时间不断减少，人类决策质量会急剧下降，我们在第五节 B 小节对此进行了量化。对时钟状态视而不见的模型只能学习一种单一的平均策略，将时间充裕时和时间压力下的下棋混为一谈。这种平均处理在时间压力极大的对局中危害最大，因为此时人类行为的变化最为显著。

本文针对这一空白提出一个主张：对强大的国际象棋策略施加时间状态条件化，会使其行为更像人类；这种效果主要集中在时间压力下，因为无时间感知模型恰恰会在人类表现最差时仍然下得过强。我们的贡献如下：

- 面向国际象棋Transformer的输入级时间条件化：将时钟状态、用时控制和近期节奏作为策略所关注的条件化词元；
- 一项在2300万规模下进行的受控步数匹配消融实验表明，准确率增益会随着时间压力和棋手水平的提高而单调增大；采样策略能够复现人类走法质量下降，而时间盲化的孪生模型则无法表现这一现象；
- 两项干预探针（反事实注入和对动态时间输入进行逐项置换），用于区分训练模型对时钟的依赖与其对相关局面特征的依赖；以及
- 将模型与已发布的 Maia-3 进行分级比较，评估范围从 Maia-3 自身经过筛选的评估方案扩展至完整的人类下棋水平谱系。

II. 相关工作

国际象棋中的人类行为克隆。Maia [1] 为每个等级分区间训练一个卷积网络。Maia-2 [2] 则用一种以技能嵌入为条件的单一架构取代这些网络，从而能够在不同等级分之间平滑插值。Allie [3] 将问题重新定义为对局记录上的下一令牌预测，因此能够利用完整历史记录，并在预测着法之外预测思考时间和认输，同时在推理时采用KL正则化搜索 [6]。Maia-3 [4]直接表示棋盘，每个棋格对应一个词元，并采用几何注意力偏置，取得了目前最强的着法匹配准确率。相关工作还展示了无需搜索的 Transformer 特级大师水平棋力 [7]，，但其目标与人类对齐不同；近期工作则通过等级分条件化的残差棋手嵌入，将 Maia-3 扩展到个人风格建模 [8]。

时间信号。 对在线对局的大规模分析揭示了人类错误与棋手水平、局面难度以及可用于下一着的时间之间的关系 [9], ，但现有模型处理时钟的方式有三种。Maia 系列会过滤掉时间压力：训练数据排除棋手剩余时间首次低于 30 秒之后的所有局面[1], [4], ，因此模型从未接触过时钟主导行为的情形。Allie 保留这些局面，并将静态用时控制纳入输入，但只把思考时间作为输出进行预测；其策略并不会观测动态时钟状态，即剩余时间或近期节奏 [3]。近期工作中，UniMaia-Aux [10] 通过自然语言提示调节一个冻结的国际象棋策略，提示元数据可包含剩余时间和上一步的延迟时间，并采用辅助延迟和认输目标。总的来说，现有方法要么过滤掉时间压力，要么仅通过辅助目标或提示间接使用时钟信息，而非将动态时间状态直接整合到策略生成中。

III. 方法

A. 骨干网络

我们以 Chessformer架构 [4],为基础，即 Maia-3模型的纯编码器骨干网络，并保持其不变。64个棋格中的每一个都是一个词元，其特征主要由棋盘本身构成：当前局面及之前七个局面的独热棋子占用信息，并附加少量辅助平面（吃过路兵、王车易位权利和重复局面标记）以及技能嵌入作为额外通道。双方玩家的棋力分别通过在 Elo 0 和 5000 对应的两个学习得到的锚点嵌入之间进行插值来编码；由此得到的向量对每个棋格都相同，并在单次输入投影形成词元之前拼接到各棋格的通道中。每个编码器块都采用几何注意力偏置，将棋盘几何结构注入注意力机制，并结合 RMSNorm 和 GELU 前馈子层。对 64 个棋格进行双线性起止分解以生成策略，均值池化头则预测胜和负结果。我们的 23M 配置采用 512 的嵌入宽度、16 个注意力头和 8 个编码器块；79M 配置则将宽度加倍至 1024，并采用 32 个注意力头。以上所有部分均保持 Chessformer设计不变；接下来介绍我们的新增部分：时间条件化和行为头。

B. 时间条件化

我们提取五种信号：用时控制的基础用时和加秒、双方的剩余时钟时间（时钟状态），以及一个概括最近八个半回合着法（思考）时间的向量（近期节奏）。我们将这三者统称为模型的时间状态，并将其编码形式称为模型的时间输入。时钟状态和近期节奏会在一局棋中变化，构成时间状态中的动态 部分，与静态用时控制相对。它们被编码为四个条件化令牌，使得用时控制、双方各自的时钟以及近期节奏彼此保持独立，便于注意力机制分别处理。这四个词元被置于 64 个棋盘词元之前，形成长度为 68 的序列。两个时钟词元在对数时间尺度上对两个学习得到的锚点嵌入进行插值，使相近的时钟状态获得相近的嵌入，而无需量化。由于 Transformer 像处理其他词元一样关注这些词元，因此每个棋格上的预测走法分布都可以直接依赖时钟。另一种通道拼接变体将相同的五种信号折叠到每个棋格的输入投影中（以骨干网络编码棋手水平时完全相同的方式注入时间，且不改变序列长度），其验证集走法匹配准确率在整个训练过程中始终比前置词元方案低约 0.15 个百分点；因此我们报告词元方案的结果。

C. 行为头

除了策略头和价值头之外，模型还采用了 Allie[3], 的设计，预测思考时间、认输以及着法，并搭载与主干网络共享参数、联合训练的行为头：一个以着法为条件的思考时间头，用于预测人类在每个候选走法前会思考多久；以及一个认输头，用于预测人类是否会认输。这些设计使模型能够体现人类棋手的行棋节奏和认输行为。

思考时间头会预测一个涵盖离散一秒区间的完整分布，并以着法为条件。该头通过对 64 个方格词元进行注意力池化，输出一个局面级基线分布，并为起始格和目标格分别输出 logits；候选走法对应的分布由基线分布加上该走法的起始格项和目标格项得到。因此，每个合法走法都能通过三个小型投影获得专属的预测分布，在保持着法条件性的同时控制了预测头的开销：同一局面下，强制回吃会得到一个快速分布，而危险弃子则会得到一个较慢的分布。每个区间宽一秒（总共 121 个，覆盖 0 到 120 秒，最高区间包含所有更长的思考时间），与 Lichess 数据库导出时钟数据的一秒分辨率一致；将时间视为分类问题，使该头能够捕捉人类思考时间偏斜且多峰的分布形状，并避免向均值回归。

IV. 实验设置

A. 数据

我们使用 2023 年至 2025 [11] 年 12 月期间的 Lichess 带等级分对局进行训练，这些对局包含时钟标注。局面按照棋手水平（埃洛等级分 1200–2500，每 100 分为一个区间）进行平衡；在每个水平区间内，再均匀分布于三种用时控制：超快棋、快棋和快速棋。纳入超快棋可以扩大模型所观察到的时间压力范围，并涵盖时钟信号最为重要的情形。两种极端用时控制——极限超快棋和慢棋——也予以保留，但不作平衡，其数量上限分别约为平衡训练集的 1% 和 5%；等级分范围不在 1200–2500 之间的局面也同样保留，但设有数量上限。平衡训练集包含 18 亿个局面。所有评估均使用一个留出测试集，其中包含来自 2026 年 5 月对局的 1,500,029 个局面，这些对局的日期晚于本文中所有训练时间窗口

（包括 Maia-3 的训练窗口）。其中，1,485,359 个是行棋局面，14,670 个是终局局面；基于行棋的指标仅使用前者。测试集涵盖与训练集相同的 1200–2500 等级分范围，但保留了自然的用时控制分布（49% 快棋、38% 超快棋、13% 快速棋）。

B. 模型与实验方案

核心消融实验比较了两个 23M 模型，它们在相同方案下各训练一百万步：一个是时间条件化模型，另一个是时间盲化 模型，后者接收恒定（置零）的时间输入。两个模型使用相同的骨干网络、数据、优化计划，并评估对局结果、思考时间和认输头。我们沿用 Maia-3 的规模命名；在小规模模型中，条件化词元和各个头增加约 0.4M 个参数，在大规模模型中则增加约 0.7M 个参数（分别为 23.3M 对 22.9M，以及 79.6M 对 78.9M）。条件化模型还带有一个小型辅助超时判负头，该头在没有时钟访问时会退化（66,817 个参数，占 0.29%），本文不对其进行评估。

训练使用 AdamW，有效批量大小为 1024；23M 和 79 M 配置的峰值学习率分别为 1.25×10^{-4} 和 1×10^{-4}。此外，我们对消融实验的两个分支各使用三个随机种子训练 250,000 步，以测量训练方差。这项扫描实验的结果见第 V-A 节；全文所列的主要数据仍来自每个分支各自单次完整预算训练的结果。

为与最先进水平进行比较，我们还训练了一个 79M 的时间条件模型，并使用 Maia-3 官方发布的 79M 检查点 [4] 及其推理代码，在相同局面上进行评估；在使用其评估方案评估该检查点在我们的测试集上的表现之前，我们的测试框架先在其测试数据上复现了 Maia-3 已公布的 57.1% 着法匹配准确率。

C. 指标

主要指标是 Top-1 步法匹配准确率：模型最可能的合法走法与人类走法相同的局面所占比例。为评估走法质量，我们计算策略以温度 1.0 采样的走法相对于 Stockfish（深度 12）的厘兵损失（CPL）；厘兵损失衡量某一步法比引擎的最佳走法差多少，并使我们能够将模型的错误特征与人类的错误特征进行比较。行为头通过胜和负准确率、思考时间误差（以秒计）以及认输精确率、召回率和 F1 值进行评估。走法指标按时间压力分别报告，时间压力以剩余时钟时间占基础用时的比例表示，并分为四个分桶：宽裕（高于 50%）、适中（20–50%）、紧迫（5–20%）和危急（低于 5%）；测试集包含 47,562 个危急局面。消融实验的差值按局面配对，并使用自助法置信区间；此外，系统比较区间按对局进行聚类，因为同一对局中的局面彼此相关。所有区间均基于局面计算。

D. 干预探针

仅凭分桶级差异，无法区分模型读取的是时钟信息还是棋盘信息，因为时间压力分桶与局面类型相关。因此，我们采用两种探测方法，在固定局面的同时，仅操控时间输入。反事实注入会在相同局面上重新运行模型，覆盖时钟状态和近期节奏，并使用引擎为采样的着法评分。置换则将某个局面的动态时间输入（时钟状态和近期节奏）替换为同一用时控制下随机局面的相应输入；这样既保留了它们的分布，又消除了它们与棋盘的关联。

E. 可复现性

每个模型均使用 PyTorch，在单张 NVIDIA RTX 5080（16GB）上进行训练，采用 fp16 混合精度、最大范数为 1.0 的梯度裁剪和 0.1 的丢弃率；优化器为 AdamW（权重衰减为 0.01），并在 2,000 个线性预热步数后采用余弦衰减。联合目标分别为策略、WDL、思考时间、升变、认输和超时判负损失赋予 1.0、0.5、0.5、0.1、0.1 和 0.1 的权重；二元认输损失和超时判负损失采用的正类权重分别为 20 和 50。使用 Stockfish 16 在深度 12 下评估着法质量。实现代码、模型配置、随机种子、训练后的权重以及详细实验结果均已在 GitHub 上公开[1]。

V. 结果

A. 着法匹配准确率在压力下更为集中

表 I 报告了步数匹配的 23M 模型对在各时间压力区间的 Top-1 准确率。总体而言，时间条件模型仅领先 $+0.73$ 个百分点，但配对差值随时间压力增大而单调扩大，在关键区间达到 $+2.45$ 个百分点。时间充裕时，两个模型在绝大多数局面上的判断一致；所剩时间不多时，时钟信号的作用最大，这与速度—准确率权衡 [5]相符。

缩减预算种子扫描表明，走法匹配效应及其随压力变化的形态具有可复现性。在经过 250,000 步训练的三组匹配随机种子对中，整体平均差值为 $+0.61$ 个百分点，在关键区间为 $+2.36$ 个百分点。在这两种情况下，每个时间条件化运行的表现都优于每个时间盲化运行；从宽松局面到关键局面，差距单调扩大的趋势也出现在全部三组匹配对中。

表 I 的下半部分按棋手等级分拆分了差值。总体提升也随着棋手水平单调增长，最低与最高等级组的置信区间互不重叠，而且在各个压力分桶内，提升通常也随棋力增强而增加。最强棋手在关键时间困境下获得的提升最大。较强棋手很可能能更系统地对时钟作出反应，并凭借有意识的时间管理和训练有素的乱战应对，因此其依赖时间的行为更容易学习。

[1]https://github.com/Prof-it/temporal-chess-model

表 I 按时间压力区间划分的 Top-1 步法匹配准确率（%）：步数匹配的 2300 万模型，1,485,359 个局面。方括号：按局面进行配对自助法计算的 95%置信区间。圆括号：三个匹配的随机种子对在训练 250,000 步后，最大值与最小值之间的差距。

模型	总体	舒适 (>50%)	Mod. (20–50%)	有压力 (5–20%)	危急 (<5%)
时间条件化	56.19	54.76	59.41	59.60	59.05
时间盲化	55.46	54.18	58.60	58.45	56.61
Δ (百分点)	+0.73	+0.58	+0.81	+1.15	**+2.45**
95%置信区间	[.69,.77]	[.52,.63]	[.69,.92]	[1.00,1.31]	[2.14,2.72]
按等级分区间配对的 Δ（百分点）					
埃洛等级分 1200–1500	+0.61 [.53, 69]	+0.50	+0.70	+1.05	+1.93
埃洛等级分 1500–1800	+0.69 [.62,.76]	+0.55	+0.76	+1.07	+2.54
埃洛等级分 1800–2100	+0.78 [.70,.87]	+0.61	+0.89	+1.16	+2.49
埃洛等级分 2100–2500	+0.91 [.78,1.03]	+0.71	+0.86	+1.37	**+2.66**
三个匹配种子对的平均值					
种子扫描	+0.61 (.07)	+0.46 (.02)	+0.59 (.05)	+1.12 (.41)	**+2.36**(.50)

B. 时钟压力下的着法质量：对齐与归因

仅凭准确度无法看出模型的失误是否符合人类特点。随着时钟倒计时，人类的着法质量会下降：在我们的 32.9 万局面质量评估中，宽裕局面的采样走法损失为 60 厘兵分，关键局面则为 136 厘兵分；评估数据取自超快棋、快棋和快速棋对局，并按棋力和时间压力进行了平衡。图 1(a) 按局面对照，绘出了各模型相对于这条人类曲线的偏差。时间条件策略在每个分桶中的着法质量都与人类相差不到几个厘兵分；在关键分桶中，其采样着法比人类均值高出 1.4 厘兵分（CI $[-1.1, +4.3]$）。相反，无时间感知策略则越过了人类曲线：时间充裕时表现过弱，压力来临后又表现过强，在关键分桶中最终高出 14.9 厘兵分。这种模式符合这样一种情况：单一策略对其无法区分的各种时钟状态取了平均。

图 1(b) 中的反事实探测固定棋盘，仅改变时间输入（第 IV-D 节）；探测针对的是 183,426 个实际处于时间紧迫状态的局面，因此，人类在相同棋盘上的失误率构成了真正的基线。在关键局面中，人类的失误率为 11.6%，而时间条件模型在实际时钟状态下的表现与之高度吻合（11.4%）；无时间感知模型看到相同的棋盘却没有时钟信息，其失误率偏低至 9.5%；注入从容状态下的动态时间输入，则使条件模型的失误率降至 9.0%。受压分桶呈现出相同的模式。在这些棋盘上，仅凭时间输入就足以解释模型从匹配人类失误率到未能匹配之间的差异。在一个独立的分桶代表性样本上扫描注入的时钟时间，可以看到失误率逐级上升：从满时钟状态下的 5.2%，升至基础用时 2% 时的 8.8%。

C. 超越走法选择的行为类人性

类人性并不局限于着法选择。表 II 报告了行为头；时间盲化基线也包含这些行为头，其使用相同的数据进行训练，但无法访问时钟信息。时间条件模型对对局结果的预测更

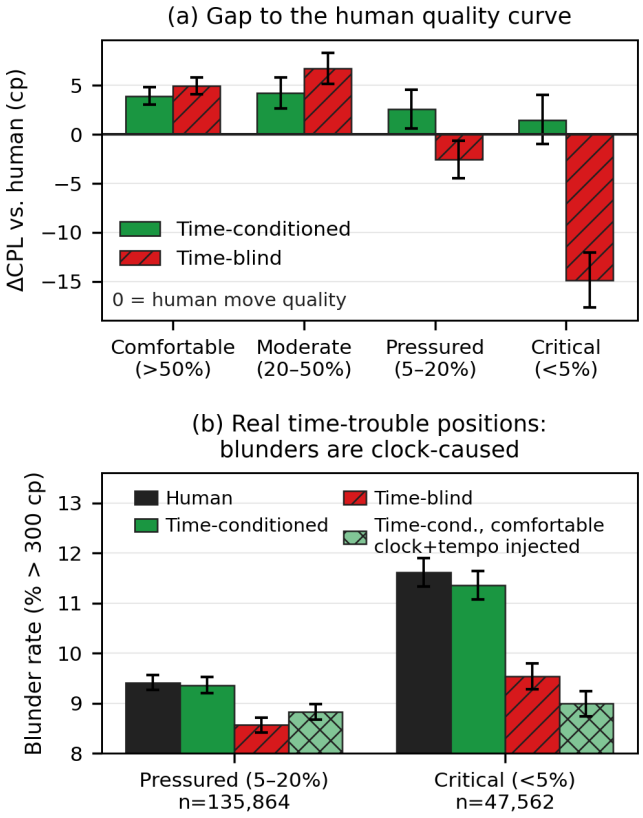

图1. 采样走法的质量（温度 1.0，Stockfish 搜索深度 12，95% bootstrap 置信区间）。(a) 按时间压力区间统计与人类厘兵分的偏差，对 329k 个局面逐一配对：在时间压力极大的局面中，无时间感知模型的走法强出 14.9 厘兵分，而时间条件模型与人类基线的偏差仅为 +1.4 厘兵分。(b) 对 183,426 个真实陷入时间困境的局面进行反事实分析：时间条件模型在真实时钟状态下与人类失误率相符；在相同棋盘局面中输入宽裕的动态时间输入（时钟状态和近期节奏），其失误率则降至 9.0%。

为准确，其中相当一部分差距来自读秒阶段的对局——这些对局的结果取决于时钟竞速；此外，该模型对认输的预测具有更高的精确率和 F1值（7,476 个正例，基准比例为 0.5%）。

思考时间头受到时钟访问的影响最大，表II的下半部分展示了这一影响的来源。在关键局面中，人类平均用时 1.2 秒走棋；条件化的思考时间头预测为 1.3 秒，而无时间感知头预测为 3.1 秒。后者无法得知棋手已没有时间思考，也不知道对局采用何种用时控制，因此只能预测普通局面的平均思考时间。若机器人用无时间感知头控制对弈节奏，在时间紧迫局面中就会思考过久，明显暴露其非人类对弈的特征。

我们还验证了该头部的着法条件性提供了额外的预测信号。在完整测试集上比较了三种人类思考时间预测器。以实际走出的着法为条件时，平均绝对误差为 2.33 秒。将其替换为对所有合法着法进行策略加权平均后，误差增至 2.43 秒；完全消融着法项后，性能进一步下降至 3.75 秒。对数似然对这些方法的排序相同。

因此，实际走出的着法本身能够提供额外的思考时间预测信息，超出局面和汇总着法统计所能提供的信息。Allie 的 Lichess 部署中的玩家反映，在只有一个合理走法的局面中，机器人思考时间过长 [3]，而局面级预测头必须对这些情况取平均。

表II 行为头，步数匹配的 2300 万模型。下半部分：按分桶预测思考时间（人类均值、各行为头的平均预测值、平均绝对误差；秒）。

预测头（指标）		时间条件化	时间盲化
对局结果，胜和负准确率（%）		**64.9**	60.8
思考时间，平均绝对误差（秒）		**2.33**	3.45
思考时间，绝对误差中位数（秒）		**0.94**	1.91
认输，精确率/召回率/F1值		**.292**/.449/**.354**	.195/.497/.280
按分桶统计思考时间	**人类**	预测值（平均绝对误差）	预测值（平均绝对误差）
舒适（>50%）	4.5 秒	4.3（2.55）	5.2（3.59）
中等（20–50%）	4.5 秒	4.3（2.33）	5.5（3.55）
受压（5–20%）	2.4 秒	2.5（1.23）	4.2 (2.63)
关键（<5%）	1.2 秒	**1.3 (0.63)**	3.1 (2.14)

D. 与已发布的最先进水平进行比较

最后，我们在同一个包含 149 万个局面的测试集上，将模型与已发布的 Maia-3 检查点[4]，进行比较。Maia-3 是目前已发表的最强人类走棋预测器，该测试集的数据晚于两种模型的训练数据。Maia-3 与我们的 79M 模型在数据、训练方案和训练预算上均不相同，时间条件化也有差异，因此这是系统之间的比较，而非第二项消融实验；下文的置换探测则用于归因。

图 2 展示了与 79M 模型的比较。我们的时间条件化 79M 模型总体领先 +1.34 个百分点，而且优势会随着压力增大而扩大，在关键时间困境下达到 +4.92 个百分点。对我们模型的动态时间输入进行置换，会抹去关键分桶中 4.9 个百分点优势里的 4.3 个百分点。

由于 Maia-3 仅使用快棋对局进行训练，并会丢弃时钟剩余时间首次低于 30 秒后的局面， [4]，因此，在超快棋对局和时间紧迫局面上的评估超出了其训练分布。表 III 因此给出了分层比较：从 Maia-3 自身的评估方案开始，逐步纳入其筛选流程排除的评估场景。在 Maia-3 的评估方案下，我们的 79M 模型已经领先；加入快棋中的时间紧迫局面后，差距进一步扩大，而在完整集合上，差距又扩大一倍以上。图 2 的分桶结果显示，差异在时间压力极大的局面中最为显著。

VI. 局限性

我们通过行为对齐指标离线评估类人性；尚未在人人对弈与人机对弈的互动中进行测试。走棋和时机分布的对齐有所改善，并不一定能直接转化为感知真实感，因为后者可能取决于显著的互动层面线索。通过盲法用户研究对此进行测试仍有待未来工作。

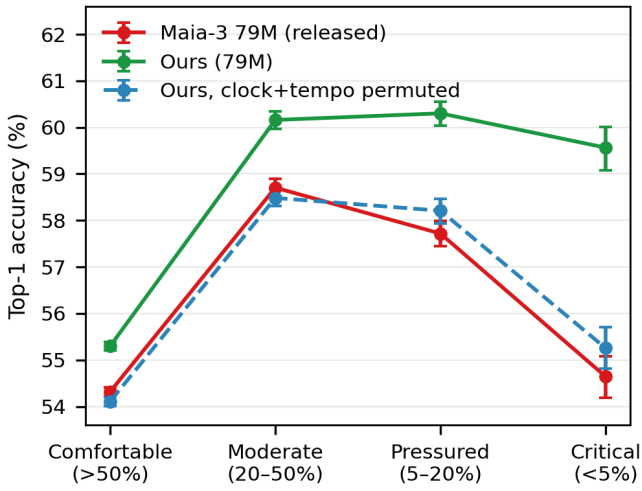

图2. 按时间压力区间划分的 Top-1 准确率：已发布的 Maia-3 79M、我们的时间条件化 79M，以及将其动态时间输入（时钟状态和近期节奏）打乱后的同一模型（按对局聚类的 95%置信区间）。打乱这些输入后，我们的模型在每个分桶中的曲线都接近 Maia-3。

表III 在 2026 年 5 月局面上与已发布的 Maia-3 79M 进行分级公平性比较，按对局聚类的 95%置信区间。每一行放宽一项 Maia-3 评估筛选条件。

机制	n	Maia-3	Ours	Δ (百分点)
Maia-3 协议[a]	515,927	57.46	57.88	+0.42 [.35,.50]
快棋，未筛选	721,907	57.04	57.60	+0.56 [.50,.62]
完整集合	1,485,359	55.45	56.79	+1.34 [1.29,1.39]

[a]仅限快棋，半回合数为 ≥ 10，任一方时钟剩余时间为 <30 秒之前的局面。

受控消融实验仅涵盖一个规模（23M），且只在缩减预算下为每个实验组使用三个随机种子；因此，种子扫描可以界定走棋匹配差值在不同运行间的变化范围，但质量、反事实和行为预测头的结果则各自仅基于每个实验组的一次运行。79M 模型没有对应的无时间感知模型，因此其消融证据来自置换探测。Lichess 数据库以一秒为粒度记录时钟，因此一秒以下的思考时间会被量化舍去；这使得思考时间目标在走棋最快的情况下噪声最大（关键分桶的平均值为 1.2 秒）。无时间感知模型省略了超时弃权预测头（时钟退化），参数差异为 0.29%。所有数据均来自 Lichess，我们的结果描述的是基于这些数据拟合的模型，而非人类行为本身；在其他平台或线下对弈中的行为可能有所不同。

七、结论

国际象棋中的人类下棋表现会受到时钟影响，但领先的行为克隆模型并未直接以动态时间状态为策略条件：以往模型要么在训练前过滤掉时间压力，要么将走棋时机视为预测目标。我们的模型通过四个学习得到的词元，以时间状态为 Chessformer 骨干网络的策略提供条件，并添加了以走法为条件的思考时间头和认输头。

在一项步数匹配的消融实验中，各组仅在时间条件化方面有所不同；每组使用三个缩减预算随机种子，并重复了走法匹配模式。时钟信号带来了适度的总体准确率提升，而在关键时间困境中，这一提升幅度增至三倍以上；它使采样策略能够再现人类走法质量随时间压力而下降的情况，而无时间感知模型对这种下降的估计系统性地低了 15 厘兵；仅对时间输入进行干预，就能让模型在固定棋盘上从匹配到未能匹配人类失误率之间转变。此外，一个 79M 参数的时间条件模型在关键时间困境中的表现，较已发布的最先进水平高出近 5 个准确率百分点；尽管此处的数据和训练方案也有所不同，但打乱动态时间输入几乎会抹去这一优势。

对于促成人类建模应用的场景（可信的对手、教练工具、作弊检测）而言，时间紧迫阶段是在线对局中的常见特征，而无法表示这一阶段的模型最明显地会在此处偏离人类行为。

参考文献

[1] A. Anderson，“使超人类人工智能与人类行为对齐：以国际象棋作为模型系统”，载于第*26*届*ACM SIGKDD*知识发现与数据挖掘国际会议论文集，2020年，页1677–1687。[2] Z. Tang、D. Jiao、R. McIlroy-Young、J. Kleinberg、S. Sen和A. Anderson，“Maia-2：用于国际象棋中人类—人工智能对齐的统一模型”，载于神经信息处理系统进展（*NeurIPS 2024*），2024年。[3] Y. Zhang、A. P. Jacob、V. Lai、D. Fried和D. Ippolito，“通过少量搜索实现人类对齐的国际象棋”，载于国际学习表征会议，2025年。[4] D. Monroe、G. Eilender、P. Chalmers、Z. Tang和A. Anderson，“Chessformer：用于国际象棋建模的统一架构”，载于第十四届国际学习表征会议，2026年。[在线]。可访问：https://openreview.net/forum?id=2ltBRzEHyd [5] B. D. Burns，“速度对高水平国际象棋表现的影响”，《心理科学》，第15卷，第7期，页442–447，2004年。[在线]。可访问：http://www.jstor.org/stable/40064150 [6] A. P. Jacob、D. J. Wu、G. Farina、A. Lerer、H. Hu、A. Bakhtin、J. Andreas和N. Brown，“通过KL正则化搜索对强大且类人的游戏进行建模”，载于第*39*届国际机器学习会议论文集，机器学习研究论文集系列，K. Chaudhuri、S. Jegelka、L. Song、C. Szepesvari、G. Niu和S. Sabato编，第162卷。PMLR，2022年7月17–23日，页9695–9728。[在线]。可访问：https://proceedings.mlr.press/v162/jacob22a.html [7] A. Ruoss、G. Del´etang、S. Medapati、J. Grau-Moya、L. K. Wenliang、E. Catt、J. Reid、C. A. Lewis、J. Veness和T. Genewein，“利用大规模Transformer进行摊销规划：以国际象棋为例”，载于第*38*届神经信息处理系统国际会议论文集，NIPS ’24论文集。美国纽约州雷德胡克：Curran Associates Inc.，2024年。[8] J. Carlson，“通过评分条件残差走子模型实现埃洛等级分解耦的棋手风格嵌入”，2026年。[在线]。可访问：https://arxiv.org/abs/2606 25176 . [9] A. Anderson、J. Kleinberg和S. Mullainathan，“依据完美基准评估人类错误”，*ACM*知识发现与数据挖掘汇刊，第11卷，第4期，2017年7月。[在线]。可访问：https://doi.org/10 1145 3046947 . [10] / S. Siu和L. Istead，“UniMaia：利用语言引导国际象棋策略，实现类人下棋”，2026年。[在线]。可访问：https: //arxiv.org/abs/2605 27767 . [11] Lichess，“Lichess开放数据库”，https://database.lichess.org，2026年。