# 精度差分観測による重み共有再帰 Transformer の学習ダイナミクス解明

– Gate Decay 振幅制御、FP32 崩壊境界、および多制御器協調学習 —

中村忠行 Aletheon Research tada2503@yahoo.co.jp

## 要約

7.0B パラメータの重み共有再帰 Transformer（BathysRDT）を 200M トークンの日本語テキストで知識蒸留により学習した結果から、4 つの知見を報告する。第一に、再帰ゲーティング関数に深度依存の減衰機構（Gate Decay）を適用することで、隠れ状態の方向を保持しつつ（全深度でトークン平均コサイン類似度 > 0.96）、振幅を最大 76 倍抑制できることを実証した。4 条件のアブレーションにより、減衰設定から勾配配分を経てクロスエントロピー性能に至る因果連鎖を確立した。第二に、十分に深い再帰において、FP32 LayerNorm の分散計算がオーバーフローし、完全な表現崩壊（representation collapse）を引き起こすことを発見した。この崩壊境界は静的ではなく、学習中に約 40M トークンあたり 1 深度の速度で浅い側へ侵食する。反実仮想シミュレーションにより減衰強化ではこの侵食を阻止できないことを示し、FP64 精度の差分実験（FP64 LayerNorm）により崩壊が完全に消失することを確認した。FP64 化により有効深度上限は d=8 から d=15 に拡大し、V3 安定性比率の最小値は 0.127 から 0.972 に改善した。第三に、減衰ダイナミクスと数値精度の相互作用は、入力分離原則に基づく多制御器フレームワークを必要とする。各制御器が独自の信号を観測することで振動的結合を防止する。具体的な制御アルゴリズムとパラメータは特許出願書類に開示されている（ALTH-001〜006、ALTH-012、特願 2026-201413）。

第四に、再帰の初回ステップ（d=1）は再帰的洗練（d≥2）とは物理的に異なる注入段階であり、全深度一律の減衰パラメータでは初回ステップへの制御が不十分となることを発見した。深度 1 専用の減衰時定数（τ_entry）を分離導入し、200M トークン学習完了時点において、d=1 のクロスエントロピーが d≥2 と同等水準（CE=10.256、d3 との比 R_entry=0.992）に収束することを確認した（Cond G）。τ_entry 未分離の v3.5 では d=1 CE が 21.868 に劣化しており（R_entry=0.473）、分離制御の効果は顕著である。

**キーワード:** 再帰 Transformer、重み共有、Gate Decay、表現崩壊、LayerNorm オーバーフロー、多制御器学習、適応的深度、d=1 分離制御

## 1. はじめに

重み共有再帰 Transformer は、同一ブロックの繰り返し適用により、N パラメータでより深いネットワークの表現能力を実現する。Universal Transformer（Dehghani et al., 2019）がこのパラダイムを確立し、近年は Geiping et al.（2025）、ByteDance（2025）、Bae et al.（2025）により数十億パラメータ規模への拡張が実証されている。

しかし、このアーキテクチャを大規模で学習させると、標準 Transformer では生じない現象が顕在化する。隠れ状態が同一パラメータを繰り返し通過する際、数値量が深度方向に累積する。微小な 1 ステップあたりの効果が共有重みを通じて蓄積し、独立パラメータを持つアーキテクチャには類似物のない深度依存のダイナミクスを生み出す。

本論文では、7.0B パラメータの重み共有再帰 Transformer（BathysRDT）を 200M トークンの日本語テキストで知識蒸留により学習した結果から、4 つの知見を報告する。

知見 1：Gate Decay は振幅制御器である（§4）。再帰ゲーティング関数に深度依存の減衰機構を導入し、同一チェックポイントプローブ実験により、隠れ状態の方向を保持しつつ（全深度でトークン平均コサイン類似度 > 0.96）、最深部で振幅を最大 76 倍抑制することを実証する。4

条件のアブレーションにより、この振幅制御が深度間の勾配配分を直接決定し、減衰設定からクロスエントロピー性能への因果連鎖を確立した。

知見 2：FP32 精度による崩壊境界の侵食（§5）。十分に深い再帰において、隠れ状態の大きさが FP32 LayerNorm のダイナミックレンジを超え、完全な表現崩壊を引き起こすことを発見した。この崩壊境界は静的ではなく、約 40M トークンあたり 1 深度の速度で浅い側へ侵食し、共有重みを通じて伝播する。反実仮想シミュレーションにより減衰強化ではこの侵食を防止できないことを示し、FP64 精度の差分実験（FP64 LayerNorm）により崩壊の完全消失を確認した。

知見 3：多制御器協調の必要性（§6）。減衰ダイナミクスと数値精度の相互作用は、単一の介入では対処できない制御問題を生む。学習率適応、減衰軌道最適化、数値安全性、蒸留圧力を分離する多制御器フレームワークの構造を記述する。入力分離原則——各制御器が独自の信号を観測する——が振動的結合を防止するための必要な設計制約として出現した。具体的な制御アルゴリズムとパラメータは特許出願書類に開示されている（ALTH-012、特願 2026-201413）。

知見 4：d=1 初回注入ステップの分離制御の必要性（§4A）。再帰ループの初回ステップ（d=1）では、隠れ状態がプレリュードから初めて再帰ブロックに注入される。このステップは d≥2 の再帰的洗練とは質的に異なるが、従来の Gate Decay 設計では全深度に同一の減衰時定数 τ を適用する。実験により、τ が十分大きい場合に d=1 の Gate Decay 効果がほぼゼロとなり、d=1 CE が d≥2 の約 2 倍（21.868 vs 9.4〜11.1）に劣化することを発見した。深度 1 専用の減衰時定数 τ_entry の分離導入により、200M トークン学習完了時に d=1 CE が 10.256 に改善し（R_entry: 0.473→0.992）、d=1 が有効な計算ステップとして機能することを確認した。

これらの知見は、同一データ・蒸留パイプラインで学習した標準 Transformer ベースライン（3.3B パラメータ、固定 12 層）との比較により検証される。第 1 稿（中村, 2026; Jxiv）では再帰 Transformer において隠れ状態の収束が予測の収束を保証しないことを示した。本稿はこの系譜を拡張し、そのような収束挙動の背後にある学習時ダイナミクスを特徴づける。

## 2. 関連研究

深度再帰 Transformer。Universal Transformer（Dehghani et al., 2019）は Adaptive Computation Time（ACT）による学習可能な停止機構を導入した。Geiping et al.（2025）は潜在空間再帰によるテスト時計算量スケーリングを実証し、Ouro（ByteDance, 2025）はループモデルが非ループモデルと知識操作タスクで同等の性能を示すことを報告した。Bae et al.（2025）はトークン単位の適応的深度のため Mixture-of-Recursions を提案し、Csordas et al.（2026）は Universal Transformer における深度・メモリのトレードオフを確立した。

トークン単位の収束ダイナミクス。Logan（2026）は深度再帰 Transformer におけるトークン単位の不動点収束を直接計測し、収束がトークン間で非均一であることを示した。Zhang et al.（2026）は再帰がいつ真の算法的計算を構成するかを問う。本研究はこれらを補完し、収束が安全な挙動を生むか病的な挙動を生むかを決定する学習時ダイナミクスを検討する。

学習安定化。Wortsman et al.（2023）は Transformer 学習における不安定性を分類した。これらの研究は標準アーキテクチャにおける不安定性を扱う。本研究の知見は質的に異なるメカニズム——重み共有再帰に固有の振幅累積から生じる FP32 LayerNorm オーバーフロー——に関するものである。

数値精度。混合精度学習（Micikevicius et al., 2018）は標準的だが、数値精度と再帰深度の相互作用はほとんど注目されていない。FP32 LayerNorm が崩壊の近接原因であるという本研究の観察は、表現能力ではなく中間計算における精度のボトルネックを浮き彫りにする。

適応的深度制御。PonderNet（Banino et al., 2021）、CALM（Schuster et al., 2022）、PABEE（Zhou et al., 2020）は安定した収束挙動を前提とする。本研究は収束自体

が病的になる条件を特定し、洗練された停止基準ではなく深度退避を必要とする状況を明らかにする。

## 3. アーキテクチャと実験設定

### 3.1 BathysRDT アーキテクチャ

BathysRDT は重み共有再帰 Transformer の構成に従う：埋め込み → プレリュード（4 層の独立 Transformer encoder）→ 再帰ブロック（ゲーティング機構を持つ単一の RDTBlock、d 回適用）→ コーダ（4 層の独立 encoder）→ LayerNorm → 予測ヘッド（埋め込みと重み共有）。

ゲーティング機構は候補状態から再帰状態への更新を調整する有効ゲート値を計算する。深度依存の減衰により深い深度で有効ゲートが低減され、Gate Decay 機構を実装する。具体的な関数形とパラメータは特許出願 ALTH-001 に記載されている。

モデルは約 3.3B のユニークパラメータを含み、全深度展開時に約 7.0B の有効パラメータに達する。

### 3.2 学習設定

単一 NVIDIA RTX 3090（24GB VRAM）。CulturaX 日本語コーパス（161 シャード、各約 256 万トークン）。llm-jp-4-32b-a3b-base（32B MoE ティーチャー）からの知識蒸留（top-128 スパースロジット、クロスエントロピー α=0.4、KL ダイバージェンス β=0.6、温度 T=2.0）。ChunkedBF16Adafactor、マイクロバッチ 4 固定。seed = 42。全結果は Cycle 0（0-200M トークン）のものである。

### 3.3 実験条件

| 条件 | 精度 | 深度制御 | 結果 |
|---|---|---|---|
| Cond F v3.4 | FP32 LayerNorm | 多制御器 | 200M Val CE 10.673 |
| Cond F v3.5 | FP64 LayerNorm（model.norm のみ） | 多制御器 | 200M Val CE 9.806 |
| Cond G v3.6 | FP64 LayerNorm（model.norm のみ） | 多制御器 + τ_entry 分離制御 | 200M Val CE 10.326 |
| Transformer 3.3B | 標準 | 固定 12 層 | 200M Val CE 9.437 |

2 つの BathysRDT 条件（v3.4, v3.5）の主な差異は、model.norm における分散/rsqrt 計算の数値精度である。v3.5 は FP64 化に加え回復機構を含む。Cond G（v3.6）は v3.5 をベースに、d=1 初回注入ステップ用の分離減衰時定数 τ_entry を追加した条件である。Cond G は 200M トークンの学習を完了した。Transformer ベースラインは同一データ・蒸留・オプティマイザファミリ（WSD LR スケジュール、ピーク 7.5 × 10^-4）を使用する。

### 特許開示

本論文は、重み共有再帰 Transformer の内部観測により発見された学習時挙動・数値的不具合・崩壊境界侵食とその解決を報告するものである。多制御器協調フレームワーク（ALTH-012、特願 2026-201413）および Gate Decay 機構（ALTH-001〜006）は特許出願済みであり、具体的パラメータは特許出願書類に開示されている。本論文では実験結果と設計思想を報告し、具体的パラメータは特許明細書に記載しており、本論文では省略している。

## 4. Gate Decay の振幅制御機能

### 4.1 方向と振幅の分離

同一チェックポイント・同一入力・同一初期状態で、Gate Decay の有無のみを変えた 2 条件のフォワードパスにおいて、各再帰深度 d におけるトークン平均コサイン類似度 cos_d を測定した。

Gate Decay は方向を保持する（全深度で cos > 0.96、平均 0.983、最小 0.964（d8））。一方、トークン平均ノルム比は d1 の 1.09 倍から d12 の 75.95 倍まで指数的に増大する。Gate Decay は振幅制御器であり、隠れ状態更新の方向をほぼ変えずにスケールを減衰させる。

### 4.2 勾配安定化

7B モデルによる 4 条件アブレーションにより因果連鎖を確立した：減衰設定 → 勾配配分 → CE。τ = 12 固定下では、d1 は d12 の 208,000 倍の勾配を受け取る。線形ランプスケジュール下では勾配分布がほぼ均一になり（d1/d12 = 0.33 倍）、最良の CE（10.36 対 11.45）を達成した。

### 4.3 軌道制御

同一の瞬時有効ゲート値（約 0.29）であっても、そこに至る減衰軌道が異なれば異なる隠れ状態軌跡を生む。この経路依存性は、減衰パラメータが単なる減衰率ではなく、深度依存の学習軌道の族を制御することを示す。

## 4A. d=1 初回注入ステップの分離制御

### 4A.1 d=1 の物理的特殊性

再帰ループの d=1 ステップでは、プレリュード出力が初めて再帰ブロックに入力される。d≥2 では前ステップの再帰出力が再入力される再帰的洗練であるのに対し、d=1 は異なるソース（プレリュード）からの初回注入である。Gate Decay の有効ゲート値は g_eff = g × exp(-(d+1)/τ) で計算されるため、d=1 での decay 強度は exp(-1/τ) に依存する。τ が大きい場合、この値は 1 に近づき、Gate Decay の効果は実質的にゼロとなる。

### 4A.2 τ_entry 未分離時の問題

v3.5（τ=12 固定初期値）では、d=1 の decay 値は exp(-1/12) = 0.920 であり、ゲート値をほぼ減衰させない。v3.6 の初期実験（rev2、τ_entry=12.0）において、d=1 のクロスエントロピー（d1_ce）と d=3,4 のクロスエントロピー（d34_ce）の差（gap）がほぼゼロであることが観測された。rev2 の TAU-ENTRY-2M ログから、gap は 4M 時点で -0.0008、6M 時点で +0.0000、8M 時点で +0.0004 と推移し、制御器が有意な差を検出できない状態が継続した。τ_entry 制御器は 16 回発火したが、τ_entry は 12.00 → 11.93 → 11.86 の微動にとどまった

さらに、τ_entry=12.0 条件では 4M 時点で D-CTRL L2B が COLLAPSE を検出し（range=0.0625）、d≤3 に退避した。V4 Adaptive Tau の初期過渡応答により global τ が低下し、深度拡張のロックアウト期間が発動する連鎖的障害が観測された。rev2 は 36.4M トークンで手動停止し、rev3（τ_entry=4.0）に再設計した。

### 4A.3 τ_entry 分離制御の設計

d=1 に専用の減衰時定数 τ_entry を導入し、d≥2 のグローバル τ とは独立に制御する。τ_entry=4.0 の場合、d=1 の decay 値は exp(-1/4) = 0.779 となり、τ=12 での d=4 と同程

度の減衰強度（exp(-4/12) = 0.717）を d=1 に与える。制御器が d1_ce と d34_ce の gap を監視し、τ_entry を適応調整する。

また、D-CTRL の COLLAPSE 判定時の退避下限を制約し、過剰退避を構造的に防止する設計とした。具体的な閾値は ALTH-012 に記載されている。サンプリングにおける d=1 の採取は継続する。

### 4A.4 実験結果（200M トークン）

Cond G（v3.6）の学習を 0M トークンから開始し、200M トークンまで完走した。

#### 学習初期の d=1 CE 比較（v3.5 vs v3.6）

v3.5（τ_entry 未分離）と v3.6（τ_entry=4.0）の学習初期における depth_detail の d=1 CE を比較する。v3.5 には TAU-ENTRY 制御器が未実装のため、d1_ce は depth_detail eval から取得した。

| eval 時点 | v3.5 d1 CE | v3.6 d1 CE | v3.5 d3 CE | v3.6 d3 CE |
|---|---|---|---|---|
| 1 回目（d≤4） | 23.85 | 10.82 | 11.49 | 10.24 |
| 2 回目（d≤4〜5） | 13.09 | 10.70 | 11.05 | 10.71 |
| 3 回目（d≤5〜6） | 42.76 | 10.66 | 10.34 | 10.66 |
| 4 回目（d≤6〜7） | 20.67 | 10.29 | 10.06 | 9.59 |

v3.5 の d=1 CE は 13〜43 の範囲で激しく変動し、d3 CE（10〜11）の 2〜4 倍に達している v3.6 の d=1 CE は 10.3〜10.8 で安定しており、d3 CE と同等水準を維持している。この差は学習の最初期（2M トークン時点）から一貫しており、τ_entry=4.0 による d=1 Gate Decay 強化が即座に効果を発揮していることを示す。

v3.6 の gap（d1_ce − d34_ce の正規化値）推移：2M 時点で +0.0568、10M 時点で +0.0001（depth 拡張直後の過渡）、14M 時点で +0.0736。14M での gap 拡大は、depth 拡張（d≤4→d≤5）により深層の CE が大幅改善（d34_ce: 10.661 → 9.588）した一方、d=1 は τ_entry により独立に制御されているため改善幅が小さく（d1_ce: 10.663 → 10.294）、この挙動は分離制御の設計意図に合致する。

#### 200M トークン完走時の比較（v3.5 vs v3.6）

| 指標 | v3.5（τ_entry 未分離） | v3.6（τ_entry=4.0→4.48） |
|---|---|---|
| 200M d1 CE | 21.868 | 10.256 |
| 200M d2 CE | 10.231 | 10.328 |
| 200M d3 CE | 10.333 | 10.176 |
| 200M d6 CE（best） | 9.455 | 9.886 |
| R_entry（d3/d1 比） | 0.473 | 0.992 |
| d1 CE 改善幅 | — | -11.612（53%改善） |
| 200M 有効深度 | d≤15 | d≤16 |
| 実質有効深度 | d2〜d15（14 段） | d1〜d16（16 段、+2 段） |
| LR 調整 | 1 回（161M、8.500e-04→8.075e-04） | 0 回（8.500e-04 固定） |
| 200M CE（best depth） | 9.374（d8） | 9.886（d6） |
| 200M CE（行レベル） | 9.807 | 10.326 |

v3.5 では d=1 の CE が 21.868 に劣化し、d2 以降（9.4〜11.1）の約 2 倍に達している。これは τ_entry 未分離下で d=1 の Gate Decay 効果が実質ゼロであり、d=1 が有効な計算ステップとして機能していないことを示す。D-CTRL の有効深度上限は d≤15 であるが、d=1 が実質無効であるため、有効な再帰計算は d=2〜d=15 の 14 段にとどまる。v3.6 では τ_entry 分離制御により d=1 CE が 10.256 に改善され（改善幅 11.612、53%）、d2〜d17 の CE 範囲（9.886〜10.679）の中央付近に位置する。R_entry は v3.5 の 0.473 から v3.6 の 0.992 に改善し、d=1 と d≥2 の性能差は実質的に消失した。D-CTRL 上限の d≤15→d≤16 拡大と合わせ、d=1 の有効化により実質有効深度は 14 段から 16 段へ +2 段拡大した。これは同一パラメータ数でより深い再帰計算を利用可能にすることを意味し、重み共有再帰アーキテクチャのパラメータ効率を直接改善する。

τ_entry は学習の進行に伴い 4.00 → 4.48 に適応調整され、d=1 の Gate Decay 強度を微調整している。n1 norm（d=1 隠れ状態の L2 ノルム）は 1536 → 1523 → 1497 と学習に伴い漸減した。

## 5. 表現崩壊と境界ダイナミクス

### 5.1 深度動的レジーム

200M FP32 チェックポイントの d=8 において、pre-LayerNorm 隠れ状態の二乗最大値は 2.70×10^38 に達し（FP32_MAX = 3.4×10^38 に近い）、LayerNorm の分散オーバーフロー（rsqrt = 7.7×10^-19）を引き起こした。ロジット出力は入力非依存となった（spread=0.000）。FP64 LayerNorm はこの崩壊を消失させた： d9 で var_max = 1.03×10^37、rsqrt_zero = False であった。

### 5.2 崩壊境界の侵食

崩壊境界は学習中に侵食する。V3 比率（隣接深度の崩壊ノルム比）の追跡により、83M トークン時点で d10/d9 = 0.875（侵食開始）、122M トークン時点で d9/d8 = 0.045（崩壊）を観測した。崩壊境界は約 40M トークンかけて d10 から d9 へ 1 段侵食し、共有重みを通じて伝播した。FP64 LN（v3.5）下では V3 最小値は 0.972 を維持し、境界侵食は観測されなかった。

### 5.3 減衰介入の限界

FP32 学習ログの実際の V3 比率系列を用いた反実仮想シミュレーションを実施した。V3 が閾値を下回るたびに減衰パラメータを低減するルールを適用した結果、学習前半（0-62M トークン）では違反ゼロ、後半では 503 観測中 153 回が違反した。減衰パラメータは 13 回の介入で下限に到達し、以後 140 回の違反は下限でクランプされた。これは履歴上の違反パターンが利用可能な減衰介入範囲を使い果たすことを示す。FP32 根本原因は GPU プローブ（§5.1）と FP64 反実仮想（§5.2）により独立に確立されている。シミュレーションパラメータ（低減ステップサイズ、下限値）は制御フレームワークとともに ALTH-012 に記載されている。

### 5.4 FP64 LayerNorm による安定化

model.norm の分散/rsqrt 計算のみを FP64 に置換したところ、全崩壊イベント（COLLAPSE 6→0、BLOCK_UP 20→0）が消失し、有効深度上限が d≤8 から d≤15 に拡大し、V3 最小値が 0.127 から 0.972 に改善した。CE は 10.673 から 9.806 に改善した。

## 6. 多制御器協調フレームワーク

### 6.1 構成

学習システムは、それぞれ異なるクラスの学習信号を用いて動作する複数の制御プロセスを使用する：LR-CTRL（フロンティアベースの評価 CE → 学習率）、τ-CTRL（pre-LN 正規化偏差 → 減衰軌道）、D-CTRL（FP32 ヘッドルーム余裕 + ロジットプローブ → 有効深度）、β-CTRL（深度別 CE/KL 構造 → 蒸留圧力、観測のみ）。いずれの制御器も他の制御器の出力を入力として使用しない。制御器固有の評価手法とパラメータの詳細は ALTH-012 に開示されている。

### 6.2 数値安全性監視

深度制御は、再帰の異なる段階で補完的な数値安全性観測を用いる。一方は FP32 のヘッドルーム（余裕）を固定閾値で監視し、もう一方は次の深度のロジット崩壊または SATURATED 条件をプローブする。後者は減衰制御器とは独立した固定内部パラメータを使用し、入力分離を維持する。詳細は ALTH-012 に記載されている。

### 6.3 制御結合の解消

FP32 条件（v3.4）では、崩壊境界での深度振動が学習率プロセスの評価サイクル完了を妨げる構造的デッドロックが発生した（70M トークンにわたる連続 SKIP）。v3.5 では退避後ロックアウト期間の導入により、深度退避イベント後に中断されない評価ウィンドウを保証し、このデッドロックを解消した。また、v3.4 では LR=8.500×10^-4 が 200M 間固定されたのに対し、v3.5 では 161M で LR DOWN 発火（8.500×10^-4 → 8.075×10^-4）を達成した。ロックアウト期間の詳細は ALTH-012 に記載されている。

### 6.4 ノイズ適応型判定

各制御プロセスは、その入力信号の特性に応じた判定方法を用いる。高ノイズ入力（シャード遷移による過渡的変動）にはノイズロバストな平均化機構を、低ノイズ入力（多深度集約メトリクス）にはより応答性の高い逐次判定基準を適用する。この設計原則——判定アーキテクチャは入力ノイズフロアの従属変数であり、独立したチューニングパラメータではない——は任意の多制御器システムに適用される。具体的な機構割り当ては ALTH-012 に記載されている。

### 6.5 カスケード介入

161M トークン時点で、学習率制御と深度制御が独立した信号に基づきほぼ同時に発火した。CE は両介入後に 10.05 から 9.86 に回復した。単一イベントからの 2 介入間の因果帰属は不可能である。

### 6.6 安全優先設計

8M トークン時点の早期偽陽性（重み共有ノイズによる d5 range=0.000）は、フレームワークの安全側バイアスを示す：検出システムが保守的に発火し、そのコストは限定的かつ回復可能であった（約 10M の深度拡張遅延）。

## 7. 実験結果

### 7.1 FP32 vs FP64 比較（0-200M）

FP64 介入は全崩壊イベントを消失させ、有効深度上限を 7 深度拡大した。CE 改善（10.673→9.806）は、FP32 オーバーフロー汚染なしにより深い再帰を利用できるようになったことに起因する。

| 指標 | v3.4（FP32） | v3.5（FP64 LN） |
|---|---|---|
| 崩壊イベント（COLLAPSE） | 6 | 0 |
| 深度拡張阻止（BLOCK_UP） | 20 | 0 |
| 有効深度上限 | d≤8 | d≤15 |
| V3 安定性最小値 | 0.1272 | 0.9718 |
| V3 < 0.950 の回数 | 27 | 0 |
| 学習率調整回数 | 0（デッドロック） | 1（161M） |
| 最終学習率 | 8.500×10^-4（固定） | 8.075×10^-4 |
| 最終 CE | 10.673 | 9.806 |

### 7.2 Transformer ベースラインとの比較

200M トークン時点で、標準 Transformer（3.3B パラメータ、12 層）は CE=9.437 を達成し BathysRDT v3.5（CE=9.806）を 0.369 上回った。BathysRDT v3.6（CE=10.326）は v3.5 より 0.520 高いが、これは τ_entry 分離制御が d=1 の CE を 11.6 改善する一方、d≥2 の一部で CE がやや上昇した影響である（§7.3 で詳述）。BathysRDT v3.4（FP32、CE=10.673）は大幅に劣る。

BathysRDT のアドバンテージは、パラメータ効率（3.3B のユニークパラメータで 7.0B の有効容量）と深度適応性にあり、この学習スケールでの絶対的 CE 性能にはない。継続学習によりギャップが縮小するかは未解決の問題である。

### 7.3 d=1 分離制御の効果（200M トークン）

Cond G（v3.6）の 200M トークン学習完了時の結果を示す。τ_entry の分離導入により d=1 の Gate Decay 制御が稼働し、rev2（τ_entry 未分離）で発生した D-CTRL COLLAPSE と深度ロックアウトが回避された。

200M 時点の cycle-end 評価結果（v3.5 / v3.6 並置）：

| depth | v3.5 CE | v3.6 CE | 差分 |
|---|---|---|---|
| d1 | **21.868** | **10.256** | **-11.612** |
| d2 | 10.231 | 10.328 | +0.097 |
| d3 | 10.333 | 10.176 | -0.157 |
| d4 | 10.590 | 10.298 | -0.292 |
| d5 | 10.075 | 10.252 | +0.177 |
| d6 | 9.455 | 9.886 | +0.431 |
| d7 | 9.490 | 9.906 | +0.416 |
| d8 | 9.374 | 10.197 | +0.823 |
| d9 | 9.379 | 10.375 | +0.996 |
| d10 | 9.392 | 10.679 | +1.287 |
| d11 | 9.861 | 10.048 | +0.187 |
| d12 | 9.421 | 10.121 | +0.700 |
| d13 | 9.822 | 9.972 | +0.150 |
| d14 | 10.288 | 10.062 | -0.226 |
| d15 | 11.097 | 9.991 | -1.106 |
| d16 | 9.532 | 10.415 | +0.883 |
| d17 | — | 10.088 | — |

d=1 において v3.5 の CE（21.868）が v3.6（10.256）に対し 11.612 劣化している。v3.5 の d=1 CE は d2 以降（9.4〜11.1）の約 2 倍であり、d=1 が有効な計算ステップとして機能

していない。v3.6 では d=1 の CE（10.256）が d2〜d17 の CE 範囲（9.886〜10.679）の中央付近に位置しており、τ_entry 分離制御により d=1 が深層と同等の予測性能を達成している。

d≥2 では v3.5 が v3.6 を上回る深度が多い（特に d6〜d12 で 0.2〜1.3 の差）。これは v3.5 の best-depth CE（9.374、d8）が v3.6（9.886、d6）を 0.512 上回ることと整合する。τ_entry 分離制御は d=1 の CE を 11.6 改善する一方、d≥2 の一部で CE がやや上昇する。この挙動の解釈として、τ_entry により d=1 の Gate Decay が強化されることで、d=1 で吸収されていた勾配が d≥2 に再配分された可能性がある。全体として、d=1 制御の追加は depth 間の CE 均一性を大幅に改善した（v3.5 の d1 外れ値 21.868 が消失）。

τ_entry は学習の進行に伴い 4.00 → 4.48 に適応調整された。D-CTRL 上限は v3.5 の d≤15 から v3.6 の d≤16 に 1 段拡大し、さらに d=1 が有効化されたことで、実質有効深度は v3.5 の 14 段（d2〜d15）から v3.6 の 16 段（d1〜d16）に +2 段拡大した。

## 8. 議論

### 8.1 Gate Decay は方向と振幅を分離する

cos プローブ結果（§4.1）は、Gate Decay を方向保持型の振幅制御器として位置づける（cos > 0.96）。方向的病理を修正することはできない——§5.3 の減衰介入限界と整合する。

### 8.2 FP32 精度のアーキテクチャ制約

深度方向に振幅が増大する隠れ状態に LayerNorm を適用する任意の重み共有再帰アーキテクチャは、精度上限に直面する。BF16/FP16 ではさらに浅い限界が課される可能性がある。共有重みを通じた境界侵食は、減衰制御では断ち切れない正のフィードバックループを形成する。

### 8.3 単一制御器アプローチの限界

v3.4 の構造的デッドロック——深度振動が学習率評価を阻止——は、複数の信号を共有制御経路に通すことがカップリングを生むことを示す。入力分離が必要な制約として出現した。

### 8.4 d=1 初回注入の制御上の意味

d=1 の分離制御の必要性は、重み共有再帰構造に固有の問題を示唆する。独立パラメータを持つ標準 Transformer では各層が独自のパラメータで処理するため、特定の層への選択的な減衰制御は不要である。重み共有再帰では、同一パラメータが全深度で使用されるため、初回注入ステップを再帰的洗練ステップと区別する外部機構が必要となる。

200M トークン学習完了時の結果は、この設計判断を支持する。v3.6 の d=1 CE（10.256）は d≥2 の CE 範囲内に収まり、R_entry（d3/d1 比）=0.992 は d=1 が有効な計算ステップとして機能していることを示す。対照的に、v3.5 の R_entry=0.473（d1 CE=21.868）は d=1 が深層の半分以下の予測性能しか持たないことを意味する。この改善（R_entry: 0.473→0.992）は τ_entry 分離制御の直接的な効果であり、全深度一律の Gate Decay 設計の限界と、深度区間ごとの差別化の有効性を実証する。

実用上の帰結として、τ_entry 分離により実質有効深度が 14 段（v3.5: d2〜d15）から 16 段（v3.6: d1〜d16）に +2 段拡大した。D-CTRL 上限の d≤15→d≤16 拡大（+1 段）に加え、従来は無効であった d=1 が有効化されたことによる +1 段が加わる。同一パラメータ数での再帰深度の拡大は、重み共有アーキテクチャのパラメータ効率を直接向上させる。

### 8.5 限界

全実験は seed=42 の単一 7.0B モデル・単一 GPU・日本語テキストで実施した。Transformer ベースラインは 200M で BathysRDT を上回る。多制御器パラメータは特許保護されており、

独立した再現を制限する。これらの現象がより大規模なモデルや分散学習にスケールするかは不明である。

## 9. 結論

7.0B 重み共有再帰 Transformer の 200M トークン学習から 4 つの知見を報告した。Gate Decay は隠れ状態の方向を保持しつつ（cos > 0.96）振幅を最大 76 倍抑制し、暗黙の勾配再配分器として機能する。FP32 LayerNorm オーバーフローは学習中に崩壊境界が侵食する表現崩壊を引き起こし、FP64 LayerNorm はこれを消失させて深度上限を 8 から 15 に拡大する。これらのダイナミクスの相互作用は、制御器間の振動的結合を防止するための入力分離を備えた多制御器協調を必要とする。

さらに、再帰の初回ステップ（d=1）が再帰的洗練（d≥2）とは質的に異なる制御を必要とすることを実証した。深度 1 専用の減衰時定数の分離導入により、200M トークン学習完了時に d=1 の CE が 21.868（v3.5）から 10.256（v3.6）に改善し、R_entry が 0.473 から 0.992 に向上した。v3.5 では d=1 が実質無効（CE が d≥2 の約 2 倍）であったのに対し、v3.6 では d=1 が有効な計算ステップとして機能する。D-CTRL 上限の拡大と d=1 の有効化を合わせ、実質有効深度は 14 段から 16 段に +2 段拡大した。これは、重み共有再帰構造において深度区間ごとの差別化が有効であり、同一パラメータ数でより深い再帰計算を利用可能にすることを示す

重み共有再帰は標準 Transformer に直接の類似物を持たない学習ダイナミクスを導入する。それらの管理は、重み共有が約束するパラメータ効率を実現するために不可欠である。

## 謝辞

本論文は、重み共有再帰 Transformer の内部状態観測を通じて、学習中に生じる挙動変化・数値的不具合・崩壊境界侵食とその解決を報告するものである。多制御器協調学習フレームワークは特許出願の対象である。

- ALTH-001〜006：Gate Decay 機構および関連制御方法

- ALTH-012（特願 2026-201413）：多制御器協調学習制御方法

## 参考文献

- Bae, S. et al. (2025). Mixture-of-Recursions. arXiv:2507.10524.

- Banino, A. et al. (2021). PonderNet: Learning to Ponder. arXiv:2107.05407.

- ByteDance (2025). Ouro: Looped Language Models. arXiv:2510.25741.

- Csordas, R. et al. (2026). Universal Transformers Need Memory. arXiv:2604.21999.

- Dehghani, M. et al. (2019). Universal Transformers. ICLR.

- Geiping, J. et al. (2025). Scaling Up Test-Time Compute with Latent Reasoning. arXiv:2502.05171.

- Logan, J. (2026). Per-Token Fixed-Point Convergence in Depth-Recurrent Transformers. arXiv:2607.14427.

- Micikevicius, P. et al. (2018). Mixed Precision Training. ICLR.

- 中村忠行 (2026). Transformer の隠れ状態における見かけの収束. Jxiv（JSAI 論文誌投稿中）.

- Schuster, T. et al. (2022). Confident Adaptive Language Modeling. NeurIPS.

- Wortsman, M. et al. (2023). Small-scale proxies for large-scale Transformer training instabilities. arXiv:2309.14322.

- Zhang, T. et al. (2026). When Does Recurrence Become an Algorithm? arXiv:2607.20594.

- Zhou, W. et al. (2020). PABEE: Pre-Trained Model Based Early Exiting. ACL.