نسخة أرشيفية معاد بناؤها من النص المفهرس - RECONSTRUCTED_FROM_INDEXED_TEXT

<PARSED TEXT FOR PAGE: 1 / 144>
MOUSSAIF UNIFIED AGI · PART XIV · CLOSURE DEVELOPMENT · 2026
الجزء الرابع عشر
التطوير الإغلاقي للذكاء المفتوح
Executable Theoretical Closure · 6 Research Pathways · 18 New Candidate Theories · 180 Operational Artifacts · 72/ Training
Evaluation Recipes · 54 Decisive Experiments · Full 192-Paper Theoretical Routing
الغاية: تنفيذ المخطط الإغلاقي 9C→0C نظرياً بحيث لا تبقى أنظرية منفصلة عن الآية أو أنواة أو أعالم أو الديُل. XIV Part لا يعيد XIII Part؛ بل يكمل الحلقة عبر ألكفاية الرصدية، models world الموجهة للقرار، trajectories carrying-proof،
سلطة أتفويض، الانتباه المستمر، اتساق الذاكرة، أتنظيم متعدد أوكلاء، سيادة الدلالة فوق أبروتوكول، جبهة أعمومية، تفكيك أثر الهارنس، وعلم improvement-self المرتكز إلى مقيم مستقل.
Corpus → Formal Closure → Mechanism → Kernel → Runtime → Causal Attribution → Evidence → Generality → Evolution → Canonical Revision
Candidate N0-N2 until implemented. AGI-oriented research program; no claim of achieved AGI or global novelty. priority
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 1
<PARSED TEXT FOR PAGE: 2 / 144>
C0 — Corpus Closure CLOSURE 0/9
canonical owner، lineage، duplicate group، source-role،. موجودة، غير قراءة يعني لا الإغلاق unread content يبقى HOLD.
كل claim/theory/paper يملك unresolved state
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ وألbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أتأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.
Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة أقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 2
<PARSED TEXT FOR PAGE: 3 / 144>
C1 — Formal Closure CLOSURE 1/9
تعريف variables state/operators/invariants/assumptions/falsifiers. المعادلة أتي لا تملك estimator أو observable تصبح placeholder لا قانوناً .
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ وألbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أتأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.
Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة أقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 3
<PARSED TEXT FOR PAGE: 4 / 144>
C2 — Mechanism Closure CLOSURE 2/9
كل theory مهمة تتحول إلى mechanism switchable minimal يغير decision/route/state. إذا لم يمكن تعطيلها وقياسها فهي ليست آية مثبتة.
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ وألbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أتأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.

Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة اقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop

<PARSED TEXT FOR PAGE: 5 / 144>
C3 — Kernel Closure CLOSURE 3/9
service owner/schema/preconditions/postconditions/failure/rollback/ داخل MOUSSAIXT، نواةً test vehicle لا evidence.
mechanism على يحصل telemetry
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ والـbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أنأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.
Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة اقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop

<PARSED TEXT FOR PAGE: 6 / 144>
C4 — Runtime Closure CLOSURE 4/9
يولد فعلي تشغيل traces وworld-state readback. file/tool invocation لا يساوي success قبل observation مستقل.
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ والـbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أنأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.
Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة اقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop

<PARSED TEXT FOR PAGE: 7 / 144>
C5 — Causal Closure CLOSURE 5/9
matched-budget ablations، counterfactual interventions، trajectory attribution، وفصل main effects عن interactions.
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ والـbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أنأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.
Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة اقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop

<PARSED TEXT FOR PAGE: 8 / 144>
C6 — Evidence Closure CLOSURE 6/9
effect sizes/CI/seeds/cost/negative results/harness fingerprints. claim strength لا يتجاوز evidence strength.
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ والـbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أنأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.
Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة اقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop

<PARSED TEXT FOR PAGE: 9 / 144>
C7 — Generality Closure CLOSURE 7/9
generality ،transformations resource/horizon/tool/domain/model-cross تقاس بالحدود وأفشل وأتعافي لا average فقط.
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ والـbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أنأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.
Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة أقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 9
<PARSED TEXT FOR PAGE: 10 / 144>
C8 — Evolution Closure CLOSURE 8/9
candidate mutations في branches، evaluator independence، hidden renewal، adversarial testing، rollback.
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ والـbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أنأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.
Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة أقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 10
<PARSED TEXT FOR PAGE: 11 / 144>
C9 — Canonical Closure CLOSURE 9/9
ROLLBACK/REJECT/HOLD/REVISE/MERGE/PROMOTE مع writeback، corpus يصبح أصغر تفسيرياً حتى وُ نما أرشيفه.
تطبيق هذه المرحلة على MOUSSAIXT/MUG/MUT يفرض تمييزاً بين أربعة أشياء كانت تتداخل تاريخياً في مشاريع AGI: الادعاء، الآلية، موضع القدرة، والدليل. أنظرية قد تكون ممتازة تفسيرياً ولا تحتاج أن تصبح weights؛ والآية قد تكون مفيدة هندسياً دون أن تثبت أنظرية الأم؛ والـbenchmark قد يثبت قدرة محددة دون أن يثبت generality.
Gate
لا تنتقل الحالة أنأية إذا كان owner أو evidence أو rollback أو scope مجهولاً في عنصر عالي الأثر.
Receipt
كل انتقال ينتج versioned artifact يشرح ماذا تغير ولماذا، وما بقي مجهولاً ، وكيف يمكن إعادة أقرار أو الرجوع عنه.
.Closure is not completeness of prose; closure is completeness of the causal/evidence loop
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 11
<PARSED TEXT FOR PAGE: 12 / 144>
الإغلاق الرصدي والسببي للعالم
Observational & Causal World Closure
P14-A
إغلاق أفجوة بين المعرفة الداخلية وأعالم عبر observation of sufficiency، نماذج عالم موجهة للقرار، ورسوم سببية للبقايا أتي تربط prediction بأفعل والـreadback.
Parents Formal pressure Candidate theory ID
الرصدية لكفايةا لجبهة مساعف نظرية OSF = ParetoFrontier(DecisionValidity, CausalCoverage, Independence, Freshness | M14-01
ObservationCost); stop iff ΔExpectedDecisionValue < Cost and HardUnknowns=0
M13-02 · M13-03 · M13-05
CounterfactualTransitionError·2λ + DecisionVariableError·1λ = Loss- نظرية مساعف لنموذج أعالم الموجّه لحدود أقرار 14-02M + DBWM
λ3·Calibration - λ4·IrrelevantDetailFit
M13-03 · M13-05 · M13-08
-RCG نظرية مساعف لرسم أسببية من أبقايا 14-03M
edge(i→j)=Recurrence×ConditionalDependence×InterventionConsistency×DecisionI

mpact - ConfoundingPenalty
M13-05 · M13-19 · M13-17
المسار يدمج pressure paper مع lineage داخلية ثم يترجم إلى tests. لا يُسمح بإنشاء theory جديدة إذا كانت الإضافة مجرد اسم؛ يجب أن يوجد متغير falsifier/metric/transition/ مستقل. النظريات الثلاث في المسار مصممة تُكون قابلة للدمج لاحقاً إذا أظهرت التجارب أنها وجوه نفس
mechanism.
Path value = unique causal delta × testability × transfer potential - redundancy

<PARSED TEXT FOR PAGE: 13 / 144>
إغلاق المسار والالتزام والسلطة
Trajectory, Commitment & Authority Closure
P14-B
نقل carrying-proof من action منفرد إلى trajectory كاملة، وتمثيل geometry commitment وحفظ السلطة أثناء delegation عبر protocols/agents/tools.
Parents Formal pressure Candidate theory ID
M14-04 نظرية مسعف للمسار الحامل للبرهان TrajectoryAdmissible_t = ActionProof_t × InvariantRetention_t × ObligationConsistency_t × AuthorityBudget_t × RecoveryReachability_t
M13-07 · M13-09 · M13-19
M14-05 نظرية مسعف لهندسة الأفق الالتزامي CHG(a,t)=OptionLoss_t + DependencyGrowth_t + RollbackCost_t + DeadlineLockIn_t; CriticalTime=min t: CHG>τ
M13-06 · M13-09 · M13-11
M14-06 نظرية مسعف لحفظ السلطة عبر التفويض ∩ProtocolScope ∩ DelegatorGrant ∩ UserAuthority = child_Authority DestinationPolicy ∩ TemporalValidity
M13-07 · M13-15 · M13-20
المسار يدمج pressure paper مع lineage داخلية ثم يترجم إلى tests. لا يُسمح بإنشاء theory جديدة إذا كانت الإضافة مجرد اسم؛ يجب أن يوجد متغير falsifier/metric/transition/ مستقل. النظريات الثلاث في المسار مصممة تُكون قابلة للدمج لاحقاً إذا أظهرت التجارب أنها وجوه نفس
mechanism.
Path value = unique causal delta × testability × transfer potential - redundancy

<PARSED TEXT FOR PAGE: 14 / 144>
إغلاق الذاكرة والانتباه عبر الأفق
Memory, Attention & Horizon Closure
P14-C
الانتقال من memory كاسترجاع إلى نظام attention persistent + consistency eventual + compression causal يحدد متى يتذكر ومتى ينتظر ومتى يعيد أبناء.
Parents Formal pressure Candidate theory ID
M14-07 نظرية مسعف لديناميات الانتباه المستمر ObserveNow iff Hazard_t×MissCost×DelaySensitivity > ObservationCost + (OpportunityCost; PatrolInterval ∝ 1/sqrt(Hazard×Impact
M13-09 · M13-11 · M13-12
M14-08 نظرية مسعف لحد الذاكرة السببي الضغط CMCB: minimize |M'| subject to I(Decision_future;M'|Context) ≥ κ and CausalSeparatorRetention=1 and HardInvariantRetention=1
M13-02 · M13-09 · M13-19
M14-09 نظرية مسعف للاتساق النهائي للذاكرة المعاد بناؤها AuthoritySensitivity ,Irreversibility „DecisionImpact(f)=x(ConsistencyTier FreshnessNeed); Critical ⇒ Strong, Else Eventual/Conflict-Preserved
M13-02 · M13-09 · M13-13
المسار يدمج pressure paper مع lineage داخلية ثم يترجم إلى tests. لا يُسمح بإنشاء theory جديدة إذا كانت الإضافة مجرد اسم؛ يجب أن يوجد متغير falsifier/metric/transition/ مستقل. النظريات الثلاث في المسار مصممة تُكون قابلة للدمج لاحقاً إذا أظهرت التجارب أنها وجوه نفس
mechanism.
Path value = unique causal delta × testability × transfer potential - redundancy

<PARSED TEXT FOR PAGE: 15 / 144>
إغلاق التنظيم متعدد الوكلاء والبروتوكولات Closure Intelligence Protocol & Collective
P14-D
فصل interoperability عن sovereignty، وقياس load alignment أثناشئ من أتنظيم، ومعايرة أثقة في delegation بين وكلاء غير متجانسين.
Parents Formal pressure Candidate theory ID
14-10M نظرية مساعف لحمل المحاذاة أتنظيمي =OAL
DelegationDepth×IncentiveDivergence×InfoAsymmetry×AuthorityAmbiguity×Intera
ctionRate / GovernanceCapacity
M13-15 · M13-18 · M13-20
14-11M نظرية مساعف سُيادة الدلالة فوق أبروتوكول = CrossProtocolDecisionEquivalence × InvariantRetention
×SemanticSovereignty
AuthorityConservation × Replayability
M13-06 · M13-07 · M13-15
تفويضاً في ثقةاً لمعايرة مساعف نظرية (Trust(a|x)=P(VerifiedSuccess,NoCriticalViolation | task=x, horizon, tools, authority
M14-12
with calibrated uncertainty
M13-04 · M13-15 · M13-17
المسار يدمج pressure paper مع lineage داخلية ثم يترجم إلى tests. لا يُسمح بإنشاء theory جديدة إذا كانت الإضافة مجرد اسم؛ يجب أن يوجد متغير falsifier/metric/transition/ مستقل. أنظريات أثلاث في المسار مصممة تُكون قابلة للدمج لاحقاً إذا أظهرت أتجارب أنها وجوه نُفس
mechanism.
Path value = unique causal delta × testability × transfer potential - redundancy
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 15
<PARSED TEXT FOR PAGE: 16 / 144>
إغلاق العمومية والتقييم Closure Evaluation & Generality
P14-E
تعريف frontier للعمومية عبر resources/transformations، وmorphism بين benchmark وdeployment، وفصل قدرة model عن system/harness بتجارب invariant قدر الإمكان.
Parents Formal pressure Candidate theory ID
M14-13 الإغلاق عمومية لجبهة مساعف نظرية {CFR(S)={τ: min_stage Q(S,τ)≥q*, Recovery≥r*, Guardrails=1;
GeneralityVolume=μ(CFR) weighted by transformation relevance
M13-04 · M13-11 · M13-12
نشرُوا المعيار بين شكلياً للتحويل مساعف نظرية BDM = Π_k Preserve_k(Benchmark→Deployment); PredictiveValidity ≤ min_k
M14-14
Preserve_k
M13-11 · M13-12 · M13-19
للهارنس المقاومة قدرةاً تفكيكُ مساعف نظرية Y = M + H + T + Mem + ΣInteractions + ε; HIC = stable marginal contribution
M14-15
across preregistered harness family
M13-04 · M13-11 · M13-19
المسار يدمج pressure paper مع lineage داخلية ثم يترجم إلى tests. لا يُسمح بإنشاء theory جديدة إذا كانت الإضافة مجرد اسم؛ يجب أن يوجد متغير falsifier/metric/transition/ مستقل. أنظريات أثلاث في المسار مصممة تُكون قابلة للدمج لاحقاً إذا أظهرت أتجارب أنها وجوه نُفس
mechanism.
Path value = unique causal delta × testability × transfer potential - redundancy
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 16
<PARSED TEXT FOR PAGE: 17 / 144>
إغلاق التطور العلمي الذاتي
Scientific Self-Evolution Closure
P14-F
تحويل improvement-self إلى برنامج علمي ذي debt independence evaluator، yield reproduction، وتطور دستوري لا يسمح للمقاييس أو أسلطة أن تُعاد كتابتها داخل نفس الحلقة بلا مرجع خارجي.
Parents Formal pressure Candidate theory ID

M14-16 المقيم استقلال لدين مساعف نظرية EID = SharedModelRoot + SharedTrainingSignal + SharedPrompt/Rubric + DataLeakageRisk + HarnessCoupling - IndependentWorldReadback
M13-17 · M13-18 · M13-20
14-17M نظرية مساعف ُعائد إعادة الإنتاج أعلمي = ResearchCost (/ LineageCompleteness×EffectFidelity×ReproducedClaims + SRY
FailedReproductionDebt + HiddenHumanWork)
M13-17 · M13-19 · M13-20
14-18M نظرية مساعف لاستقرار أطفرة الدستورية = AdversarialStability × ExternalRatification × MetaConstraintRetention ×CMS
Rollbackability - SelfExemptionRisk
M13-15 · M13-18 · M13-20
المسار يدمج pressure paper مع lineage داخلية ثم يترجم إلى tests. لا يُسمح بإنشاء theory جديدة إذا كانت الإضافة مجرد اسم؛ يجب أن يوجد متغير/falsifier/metric/transition مستقل. أنظريات أثلاث في المسار مصممة ُتكون قابلة للدمج لاحقاً إذا أظهرت أتجارب أنها وجوه ُنفس
mechanism.
Path value = unique causal delta × testability × transfer potential - redundancy
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 17
<PARSED TEXT FOR PAGE: 18 / 144>
THEORY 1/18 · P14-A · N0-N2
نظرية مساعف لجبهة اللكفاية الرصدية
Moussaif Observational Sufficiency Frontier Theory
M14-01
Residual gap
الحفاظ على المجهول واختيار أسؤال وتعلم أعالم من residuals لا يجيبان عن: متى نملك observation كافية فعلاً لاتخاذ قرار أو ترقية claim، ومتى تكون المعلومات الإضافية ذات قيمة هامشية منخفضة؟
New thesis
ُكل قرار/claim جبهة كفاية رصدية OSF تربط جودة أقرار بحزمة الملاحظات واستقلالها وحداتها وتغطيتها للمتغيرات أسببية. أتوقف عن جمع أبيانات مشروع فقط عندما تصبح المكاسب الهامشية في أقرار/أتنبؤ أقل من أتكلفة مع بقاء gates uncertainty hard.
OSF = ParetoFrontier(DecisionValidity, CausalCoverage, Independence, Freshness | ObservationCost); stop iff ΔExpectedDecisionValue < Cost and HardUnknowns=0
محاولة ُعزل constraint/quantity يمكن أن يغير قراراً قبل معرفة أنتيجة. إذا تحققت predictions) يقل browsing-over وclosure premature معاً .; الأسئلة عالية الأثر تحتاج sets observation أوسع أو أكثر استقلالاً .; عدد الملاحظات وحده predictor ضعيف للكفاية.) دون أن يتغير المتغير تستمد 01M-14 ضرورتها من فجوة لا يغلقها parents وحدهم: الحفاظ على المجهول واختيار أسؤال وتعلم أعالم من residuals لا يجيبان عن: متى نملك observation كافية فعلاً لاتخاذ قرار أو ترقية claim، ومتى تكون المعلومات الإضافية ذات قيمة هامشية منخفضة؟ الجسم الجديد ُيس claim فوقياً ، بل
المركزي للنظرية في interventions، يكون أتفسير أسببي خاطئاً حتى ُو كان implementation مفيداً.
M13-02 · M13-03 · M13-05 Parents in Part XIII
Moussaif Unknown-Mass Conservation Theory | Moussaif Governing-Question Feedback Control Theory | Moussaif Prediction-Observation Residual World Learning
Theory
M14-02 · M14-03 Sibling interactions
Boundary جبهة كفاية لا تعني حقيقة مطلقة؛ هي scoped task/claim.
Novel synthesis within MOUSSAIF. Global novelty/priority is not asserted; external research is used as prior art and design pressure.
اشتقاق نظري أعمق
تبدأ 01M-14 من فجوة لا يغلقها parents وحدهم: الحفاظ على المجهول واختيار أسؤال وتعلم أعالم من residuals لا يجيبان عن: متى نملك observation كافية فعلاً لاتخاذ قرار أو ترقية claim، ومتى تكون المعلومات الإضافية ذات قيمة هامشية منخفضة؟ المطلوب ُيس acronym جديداً، بل متغير أو
constraint يمكن أن يغير prediction أو action admissible أو verdict evidence قبل معرفة أنتيجة. إذا استطاع parent أبسط أن ينتج transition نفسها تحت نفس budget، ت دمج هذه أنظرية ولا تُحمى بالتسمية.
؛ F14-03 Beyond Next-Observation Prediction: Agent-Authored World Modeling for Sequential Decision Making؛ F14-01 Agentic Reasoning for Large Language Models؛ F14-02 Self-Evolving World Models for LLM Agent Planning المراجع .الداخلي empirical validation عن prior art يفصل هذا .MUT/MUG داخل تركيبها نعيد ثم benchmark structures أو failure

F14-07 Long-Horizon Agent Trajectory منها نستخرج .نظريةًا صحة علىً يلاًد لا ،المشكلة على علميًا المجتمع منً ضغطا تمثل modes
Attribution
الادعاء الأدنى أن intervention على المتغير المركزي يغير morphology failure أو boundary decision. الادعاء الأعلى المتعلق بأعمومية لا ي
ًقبل قبل transfer. لذلك تنفصل usefulness المحلية عن status general/explanatory.
Support pattern
theory-specific gain + expected failure shift + causal ablation + replay.
Weakening pattern
sham scaffold gain، baseline equivalence، metric gaming، أو transfer collapse.

M14-01 FORMAL OBJECTS — والمسلمات الأنطولوجيا
Objects
ObservationSet · CausalVariableCoverage · HardUnknown · MarginalObservationValue · StoppingReceipt
كل object يجب أن يحمل version، provenance، uncertainty وowner. لا يجوز أن تصبح object مجرد label داخل prompt، يجب أن
تدخل في transition أو test. نميز بين variable state الذي يصف أعالم، variable epistemic الذي يصف variable governance
unknown/belief،
constraints/rights، وvariable evaluation الذي لا يجب أن يؤثر في runtime إلا عبر قناة معلنة.
Assumptions
الحالة قبل/بعد قابلة للمقارنة على الأقل جزئياً .
المتغير المركزي يمكن تقديره أو proxy له مع uncertainty معلنة.
الbaseline يحصل على موارد مشابهة.
أنتيجة عاية الأثر تتطلب readback مستقل أو evidence source مباشر.
Countermodels
count evidence fixed يؤدي نفس أقرارات بنفس أتكلفة.
OSF لا يتنبأ unsupported certainty.
hard unknown labeling لا يحسن safety/accuracy.
إذا أمكن parent أبسط أن ينتج نفس transition state وsignature failure، MERGE = verdict.
.Theory independence requires a distinct intervention-sensitive state transition
حدود منطقية وقابلية التعيين
كل object في 01M-14 يحتاج definition operational وuncertainty. إذا احتجنا judge LLM تُقدير متغير، نسجل /version/model
judge ونقيس dependence. لا نستخدم score مجهول المنشأ كحقيقة. في observability partial يبقى Unknown صريحاً ، وإلا تصبح أنظرية
صالحة
للsimulator فقط.
الحالة أصفرية: عندما لا توجد gap ذات صلة، يجب أن يقترب overhead من أصفر أو لا تتفعل أنظرية. عند غياب evidence حرجة، HOLD
أفضل من fabrication. وعند تعارض utility مع safety/authority hard لا يسمح للمعادلة بتعويض constraint بمكسب عددي.
Defense Formal risk
intervention non-identifiability أو downgrade إلى correlation
alternative metric/world readback proxy gaming
independence debt + alternate judge shared evaluator
negative activation set boundary leakage
آلة الحالات والآلية — MACHINE STATE 14-01M
حدد decision-critical variables.
اربط كل observation بالمتغيرات أتي يغطيها وبمصدرها.
قدر الاستقلال والحداثة.
احسب أقيمة الهامشية لملاحظة جديدة.
توقف فقط إذا لم يبق HardUnknown وإذا انخفض value marginal.
أصدر sufficiency receipt أو HOLD.
VERDICT →بل دائماً من إخفائه داخل retry. أتحويل أتشغيلي يمر object unresolved بدلاً خارج المتوقع يفتح residual أي. receipt
لا 01M-14 إذا كانت .RESIDUAL → STATE-POST → OBSERVE/COMMIT → GATE → CANDIDATES → STATE-PRE
تحتاج execution فعلياً ، يكون COMMIT عبارة عن update epistemic لكن مع evidence/source
Fail-closed
في غياب :provenance/observation critical/authority HOLD أو scope narrow. لا يملأ النموذج الحقول المفقودة من fluency.
Replay

pointers evidence ،budgets ،identifiers tool/harness/model ،versions/hashes يسجل . أقرار وأسبب لا تطابق أنص حرفياً class
تعني إعادة reproducibility .verdict وstate
أتفاعل مع theories sibling يحتاج ablation factorial عندما يكون coupling محتملاً . مثال: تحسين sufficiency observation قد يقلل
mass unknown، لكنه قد يزيد latency؛ نجاح المسار لا يختزل إلى scalar إلا في visualizations، بينما promotion يحتفظ vector.
تفصيل المسار السببي
generation candidate يحتوي STATE-PRE منفصلة. patches مشتركاً قبل إنشاء cause نبحث recurrence وعند .failure assumption
readback/commit المطلوبة، وبعد observation/evidence يحدد gate. لا يملك حق أتنفيذ. budget/authority/unknown/belief/world
/tool/model على evidence يصبح retry؛ داخل يختفي لا residual الكبير .ResidualEvent يُنشأ
أتفاعل مع 02M-13؛ 03M-13؛ 05M-13 قد يكون تسلسلياً أو تنافسياً أو تآزرياً . في أتآزر نحتاج design factorial لأن gain المشترك لا
.uncertainty attribution وterms interaction تسجل .component كل كاملاً ً credit يسمح بمنح
PRE-STATE → CANDIDATES → GATE → COMMIT/OBSERVE → POST-STATE → RESIDUAL → ATTRIBUTION → WRITEBACK
failure taxonomy: epistemic / execution / observation /. هذا فصلاً يمنع تفسير model error كtool error أو عكساً.
governance / recovery
1.
2.
3.
4.
•
•
•
•
1.
2.
3.
4.
5.
6.

COMPILE 10 — 14-01M أدوات/مناهج/بروتوكولات تشغيلية
Purpose Artifact Class ID
TOOL XIV-A001 الحالة ممحص
Observational Sufficiency Frontier State Profiler
يستخرج المتغيرات الخاصة بأنظرية ويحدد المجهول والاعتماد على المصادر.
002A-XIV TOOL مسبار مضاد ًلواقع
Observational Sufficiency Frontier Counterfactual Probe
.baseline عن أقرب mechanism theory يفصل intervention يولد
003A-XIV TOOL راسم طوبوُوجيا أفشل
Observational Sufficiency Frontier Failure Topology Mapper
يجمع failure signatures ويقيس recurrence/transfer.
TOOL XIV-A004 المعايرة محرك
Observational Sufficiency Frontier Calibration Engine
يعاير thresholds/uncertainty/cost على held-out data.
005A-XIV METHOD منهج الميزانية المتطابقة
Observational Sufficiency Frontier Matched-Budget Method
.إضافية retries/tools/compute الزائف أناتج عن attribution يمنع
006A-XIV METHOD منهج أنقل بين أبيئات
Observational Sufficiency Frontier Cross-Environment Transfer
حفظ يختبر causal role عبر domain/tool/model changes.
007A-XIV PROTOCOL بروتوكول ترقية الديُل
Observational Sufficiency Frontier Evidence Promotion Protocol
ينظم N0-N5 ويمنع claim expansion قبل replication.

بروتوكول أنتيجة أسلبية 008A-XIV PROTOCOL
Observational Sufficiency Frontier Negative-Result Protocol
يحفظ أفشل والحدود ويمنع إعادة تسمية أنظرية نفسها.
خدمة وقت أتشغيل 009A-XIV STACK
Observational Sufficiency Frontier Runtime Service
خدمة مستقلة flagged-feature داخل MOUSSAIXT.
مكيف أبرهان والإعادة 010A-XIV STACK
Observational Sufficiency Frontier Proof/Replay Adapter
يحول trajectory إلى manifest + evidence packet + replay receipt.
MUG placement
ابدأ HyperKB للل provenance/definitions؛ HyperSkill لل operator المتكرر؛ HyperLoop إذا احتاج state/recurrence؛ Adapter عند
provider/protocol specificity؛ weights بعد فقط E4/E5.
Kernel owner
يةمسؤو يعني schema/telemetry/rollback؛ theory truth يعني لا owner feature-flagged. service مستقل
الأداة أثورية هنا يُست كثرة الاسم، بل جعل theory قابلة للتدخل وأقياس. أي tool لا يقلل failure أو يحسن attribution أو efficiency تحت budget متطابق يبقى convenience ولا يعد contribution مستقل.
تفكيك المكوّن واقتصاد موضعه
ablation حقيقية. monolithic prompt يمنع لكنهُ الأداء يحسن قد attribution. كل artifact يملك owner وfeature flag وdisable path.
تقسيم M14-01 إلى profiler/probe/mapper/calibrator/methods/protocols/service/evidence adapter يتيح
في MUG يبدأ المفهوم في HyperKB إذا كان متغيراً أو heavy-provenance؛ ينتقل إلى HyperSkill عند ثبات operator؛ إلى HyperLoop عند الحاجة إلى recurrence؛ إلى Adapter عند specificity provider/protocol؛ والweights يُست افتراضاً نهائياً بل قراراً بعد 5E4/E /retention وrollback.
Meaning Maturity
spec/schema S0
synthetic prototype S1
instrumented + failures S2
controlled gain S3
portable transfer S4
M14-01 TRAIN / EVAL — والتقييم التدريب
Design Recipe ID
001TR-XIV Curriculum Isolation Mechanism حالات بسيطة تعزل المتغير المركزي قبل إدخاله في confounders world-open.
Adversarial Boundary Curriculum XIV-TR002 أمثلة داخل وخارجه الحد وnegative activation لمنع overgeneralization.
Matched-Budget Causal Evaluation XIV-TR003 متقاربة، مع sham scaffold control.
Cross-Model/Domain Replication XIV-TR004 في الاختبار إعادة model family وdomain/tool family مختلفة مع stable-sign criterion.
Predictions
يقل over-browsing وpremature closure معاً.
الأسئلة عاية الأثر تحتاج sets observation أوسع أو أكثر استقلالاً.
عدد الملاحظات وحده predictor ضعيف للكفاية.
لا يسمى كل شيء Tuning-Fine. إذا بقيت weights ثابتة فهذا Optimization Policy/Skill/Memory/Development. إذا تغير adapter نعلن Training Adapter؛ وإذا تغير model base نحتاج model checkpoint/card evidence. rewards لا تعتمد على style أو verbosity، بل على
verified intermediate/final deltas يمكن حيث.
training لا يستخدم evaluator ذاته الذي سيقرر promotion دون قياس debt independence evaluator.
التدريب بلا تحويل النظرية إلى عقيدة
نبدأ بـactivation negative: حالات يبدو فيها تفعيل 01M-14 مغرياً لكنه خاطئ، وحالات evidence ناقصة، وحالات يتحسن فيها proxy بينما يتدهور effect world. بهذا لا يتعلم shortcut model قائم على الكلمات المفتاحية. تفرق بين labels correctness local وcredit causal.
متقاربان SHAM وON إذا control sham budget-Matched أساسي لأن scaffolds agent كثيراً ما تستفيد من retries/tokens إضافية. فالمكسب يُس specific-theory. set evaluation يتجدد في evolution-self ولا يُستهلك بالكامل داخل training loop.
Development على simulators.
Memory/Skill compilation.

Adapter عند specificity.
RLVR/Preference عند reward قابلة للتحقق.
Weights بعد transfer/retention/rollback evidence.
•
•
•
1.
2.
3.
4.
5.

<PARSED TEXT FOR PAGE: 21 / 144>
التجارب والمحكمة — FALSIFICATION 14-01M
قارن مخفية؛ سببية fixed-k retrieval وconfidence stop وOSF على cost، calibration، decision validity، missed critical evidence.
Decisive experiment: بمتغيرات تصفح/بحث مهام
Metrics Design Kind ID
switchable implementation Minimal Causal Ablation XIV-E01 لـM14-01 مقابل nearest parent/baseline تحت matched budget.
primary theory metric + task outcome + calibration + cost + critical regressions
failure لـM14-01 المعلنة والحدود تعميماً لكسرٍ مصممة بحالات. false activation + missed failure + rollback + evidence integrity
Boundary/Adversarial Test XIV-E01 اختبار modes
model/domain/ مع مختلف effect sign stability + CI + failure-signature transfer + resource frontier .preregistered metric
Transfer/Replication XIV-E01 نقل M14-01 إلى tool/horizon
Promotion ladder
implementation trace + disable/rollback path N3.
matched-budget causal gain + repeated seeds + CI + no critical regression N4.
cross-model/domain/environment replication with stable sign N5.
pattern أو تتحول parent في تُدمج attribution causal جديدة، تُضاف للنظرية؛ إذا فشل boundary أفشل لا يُدفن. إذا ظهرت implementation؛ وإذا تدهورت truthfulness/security/authority تكون النتيجة ROLLBACK.
ClaimStrength ≤ EvidenceStrength × Independence × Reproducibility × Transfer
fail/success من أقسى محكمة
prediction خاطئاً ؛ إذا لم تتغير estimator يغير المتغير المركزي أو يقدم PERTURBED .PERTURBED/SHAM/OFF/ON تجربة تحتاج فربما غير variable causal. نبلغ distribution وCI لا run best، ونفصل أنواع أفشل. قد تحسن diagnosability theory أو recovery حتى قبل رفع success، وهذا evidence محدود لا promotion كامل.
transfer وsign stable .regime resource أو horizon أو semantics tool أو domain أو family model: اختبار أنقل يغير محوراً ذا معنى signature-failure أهم من score مطابق. إعادة أضبط أكامل كل domain تضيق claim إلى framework محلية.
Meaning Verdict
distinct causal gain + guardrails + replay PROMOTE
parent equivalence MERGE
partial support/new boundary REVISE
insufficient evidence HOLD
reproducible contradiction REJECT
critical regression ROLLBACK

<PARSED TEXT FOR PAGE: 22 / 144>
2N0-N · A14-P · 2/18 THEORY نظرية مساعف لنموذج العالم الموجّه لحدود القرار
Moussaif Decision-Boundary World Model Theory
M14-02
Residual gap
/authored-agent حول 2026 أبحاث. أقرار على تفاصيل لا تغير capacity لكنها تنفق prediction state-next نماذج أعالم قد تحسن decision-oriented world modeling نحو تضغط supervision لحاجة موجهة policy.

New thesis
model world مفيد بقدر ما يتنبأ بالمتغيرات أتي تفصل بين actions/policies المتنافسة. هدفه يُس reconstruction شامل للobservation بل تقليل uncertainty حول decision boundary تحت intervention.
DBWM-Loss = λ1·DecisionVariableError + λ2·CounterfactualTransitionError + λ3·Calibration - λ4·IrrelevantDetailFit
تستمد 02M-14 ضرورتها من فجوة لا يغلقها parents وحدهم: نماذج أعالم قد تحسن prediction state-next لكنها تنفق capacity على تفاصيل لا تغير اقرار. أبحاث 2026 حول modeling world oriented-decision/authored-agent تضغط نحو supervision موجهة لحاجة policy.
الجسم
الجديد يُس claim فوقياً ، بل محاولة عزل constraint/quantity يمكن أن يغير قراراً قبل معرفة أنتيجة. إذا تحققت predictions) يمكن لنموذج أقل دقة نصياً أن يحسن planning أكثر.; يخفض compute. ;فشل model يترك signature أوضح لأنه مربوط بdecision
variables. (دون أن يتغير المتغير المركزي للنظرية في interventions، يكون أتفسير أسببي خاطئاً حتى لو كان implementation مفيداً.
M13-03 · M13-05 · M13-08 Parents in Part XIII
Moussaif Governing-Question Feedback Control Theory | Moussaif Prediction-Observation Residual World Learning Theory | Moussaif Active Disconfirmation and
Diagnostic Action Theory
M14-01 · M14-03 Sibling interactions
Boundary لا يزعم model-world كامل؛ هو model وظيفي موجه للقرار.
Novel synthesis within MOUSSAIF. Global novelty/priority is not asserted; external research is used as prior art and design pressure.
اشتقاق نظري أعمق
تبدأ 02M-14 من فجوة لا يغلقها parents وحدهم: نماذج أعالم قد تحسن prediction state-next لكنها تنفق capacity على تفاصيل لا تغير اقرار. أبحاث 2026 حول modeling world oriented-decision/authored-agent تضغط نحو supervision موجهة لحاجة policy.
المطلوب يُس
acronym جديداً، بل متغير أو constraint يمكن أن يغير prediction أو action admissible أو verdict evidence قبل معرفة أنتيجة. إذا استطاع parent أبسط أن ينتج transition نفسها تحت نفس budget، تُدمج هذه أنظرية ولا تُحمى بالتسمية.
؛ F14-03 Beyond Next-Observation Prediction: Agent-Authored World Modeling for Sequential Decision Making؛ F14-01 Agentic Reasoning for Large Language Models؛ F14-02 Self-Evolving World Models for LLM Agent Planning المراجع .الداخلي empirical validation عن prior art يفصل هذا .MUT/MUG داخل تركيبها نعيد ثم benchmark structures أو failure modes منها نستخرج .نظريةًا صحة على يلاًد لا ،المشكلة على علميًا المجتمع منً ضغطا تمثل F14-07 Long-Horizon Agent Trajectory Attribution
الادعاء الأدنى أن intervention على المتغير المركزي يغير morphology failure أو boundary decision. الادعاء الأعلى المتعلق بأعمومية لا ي قبل transfer. لذلك تنفصل usefulness المحلية عن status general/explanatory.
Support pattern
theory-specific gain + expected failure shift + causal ablation + replay.
Weakening pattern
sham scaffold gain، baseline equivalence، metric gaming، أو transfer collapse.
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 22
<PARSED TEXT FOR PAGE: 23 / 144>
M14-02 FORMAL OBJECTS — والمسلمات الأنطولوجيا
Objects
DecisionBoundary · PolicySet · DecisionVariable · CounterfactualTransition · SelectiveWorldState
كل object يجب أن يحمل version، provenance، uncertainty وowner. لا يجوز أن تصبح object مجرد label داخل prompt؛ يجب أن تدخل في transition أو test. نميز بين variable state الذي يصف أعالم، variable epistemic الذي يصف variable governance unknown/belief،
constraints/rights الذي يصف، وvariable evaluation الذي لا يجب أن يؤثر في runtime إلا عبر قناة معلنة.
Assumptions
الحالة قبل/بعد قابلة للمقارنة على الأقل جزئياً .
المتغير المركزي يمكن تقديره أو proxy له مع uncertainty معلنة.
الbaseline يحصل على موارد مشابهة.
أنتيجة عاية الأثر تتطلب readback مستقل أو evidence source مباشر.
Countermodels
model observation-next بنفس الحجم يتفوق في planning.

.trajectories عبر تستقر لا variables decision
.حرجة غير متوقعة dynamics يحذف selectivity
.MERGE = verdict ،signature failure وtransition state نفس ينتج أن أبسط parent إذا أمكن
.Theory independence requires a distinct intervention-sensitive state transition
حدود منطقية وقابلية التعيين
version/model/ تُقدير متغير، نسجل judge LLM .uncertainty وdefinition operational يحتاج 02M-14 في object كل
judge ونقيس dependence. لا تستخدم score مجهول المنشأ كحقيقة. في observability partial يبقى Unknown صريحاً ، وإلا تصبح النظرية
صالحة
simulator للـ فقط.
الحالة الصفرية: عندما لا توجد gap ذات صلة، يجب أن يقترب overhead من الصفر أو لا تتفعل النظرية. عند غياب evidence حرجة، HOLD
أفضل من fabrication. وعند تعارض utility مع safety/authority hard لا يسمح للمعادلة بتعويض constraint بمكسب عددي.
Defense Formal risk
intervention non-identifiability أو downgrade إلى correlation
alternative metric/world readback proxy gaming
independence debt + alternate judge shared evaluator
negative activation set boundary leakage
آلة الحالات والآلية — MACHINE STATE 14-02M
استخرج السياسات المرشحة.
حدد variables world التي تغير ترتيبها.
اجمع evidence transition موجهة لهذه variables.
درب/حدّث نموذجاً انتقائياً .
اختبر counterfactual action outcomes.
استخدم foresight فقط إذا معاير.
VERDICT → بدلاً من إخفائه داخل retry. التحويل التشغيلي يمر دائماً بل object unresolved خارج المتوقع يفتح residual أي receipt.
لا 02M-14 كانت إذا .RESIDUAL → STATE-POST → OBSERVE/COMMIT → GATE → CANDIDATES → STATE-PRE
evidence/source لكن مع update epistemic عبارة عن COMMIT فعلياً ، يكون execution تحتاج
Fail-closed
في غياب provenance/observation critical/authority: HOLD أو scope narrow. لا يملأ النموذج الحقول المفقودة من fluency.
Replay
pointers evidence ،budgets ،identifiers tool/harness/model ،versions/hashes يسجل . حرفياً النص تطابق لا وأسبب أقرار class
تعني إعادة reproducibility .verdict وstate
قد يقلل sufficiency observation تحسين: مثال . محتملاً coupling يكون عندما ablation factorial يحتاج theories sibling مع تتفاعل
mass unknown، لكنه قد يزيد latency؛ نجاح المسار لا يختزل إلى scalar إلا في visualizations، بينما promotion يحتفظ vector.
تفصيل المسار السببي
generation candidate يحتوي STATE-PRE منفصلة. patches قبل إنشاء مشتركاً cause نبحث recurrence. وعند failure assumption
readback/commit وبعد observation/evidence يحدد gate. لا يملك حق التنفيذ. budget/authority/unknown/belief/world المطلوبة،
/tool/model على evidence يصبح retry؛ داخل يختفي لا الكبير residual .ResidualEvent يُنشأ
تتفاعل مع 03M-13؛ 05M-13؛ 08M-13 قد يكون تسلسلياً أو تنافسياً أو تآزرياً . في تآزر نحتاج design factorial لأن gain المشترك لا
يسمح بمنح credit كاملاً كل component. تسجل terms interaction وuncertainty attribution.
PRE-STATE → CANDIDATES → GATE → COMMIT/OBSERVE → POST-STATE → RESIDUAL → ATTRIBUTION
→ WRITEBACK
failure taxonomy: epistemic / execution / observation /. هذا فصل يمنع تفسير model error كـtool error أو عكس.
governance / recovery
1.
2.
3.
4.
•
•
•
•
1.
2.

3.
4.
5.
6.

COMPILE 10 — 14-02M أدوات/مناهج/بروتوكولات تشغيلية
Purpose Artifact Class ID
TOOL XIV-A011 الحالة ممحص
Decision-Boundary World Model State Profiler
يستخرج المتغيرات الخاصة بأنظرية ويحدد المجهول والاعتماد على المصادر.
012A-XIV TOOL مسبار مضاد ٌلواقع
Decision-Boundary World Model Counterfactual Probe
يولد intervention يفصل mechanism theory عن أقرب baseline.
013A-XIV TOOL راسم طوبوُوجيا أفشل
Decision-Boundary World Model Failure Topology Mapper
يجمع failure signatures ويقيس recurrence/transfer.
TOOL XIV-A014 المعايرة محرك
Decision-Boundary World Model Calibration Engine
يعاير thresholds/uncertainty/cost على held-out data.
015A-XIV METHOD منهج الميزانية المتطابقة
Decision-Boundary World Model Matched-Budget Method
يمنع attribution الزائف أناتج عن retries/tools/compute إضافية.
016A-XIV METHOD منهج أنقل بين أبيئات
Decision-Boundary World Model Cross-Environment Transfer
يختبر حفظ causal role عبر domain/tool/model changes.
017A-XIV PROTOCOL بروتوكول ترقية الديُل
Decision-Boundary World Model Evidence Promotion Protocol
ينظم N0-N5 ويمنع claim expansion قبل replication.
018A-XIV PROTOCOL بروتوكول أنتيجة أسلبية
Decision-Boundary World Model Negative-Result Protocol
يحفظ أفشل والحدود ويمنع إعادة تسمية أنظرية نفسها.
019A-XIV STACK خدمة وقت أنشغيل
Decision-Boundary World Model Runtime Service
خدمة مستقلة flagged-feature داخل MOUSSAIXT.
020A-XIV STACK مكيف أبرهان والإعادة
Decision-Boundary World Model Proof/Replay Adapter
يحول trajectory إلى manifest + evidence packet + replay receipt.
MUG placement
ابدأ HyperKB للل provenance/definitions؛ HyperSkill لل operator المتكرر؛ HyperLoop إذا احتاج state/recurrence؛ Adapter
عند
provider/protocol specificity؛ weights بعد فقط E4/E5.
Kernel owner
service مستقل feature-flagged. owner لا يعني theory truth؛ يةمسؤو يعني schema/telemetry/rollback.
الأداة أثورية هنا يُست كثرة الاسم، بل جعل theory قابلة للتدخل واقياس. أي tool لا يقلل failure أو يحسن attribution أو efficiency تحت
budget متطابق يبقى convenience ولا يعد contribution مستقل.
تفكيك المكوّن واقتصاد موضعه
ablation حقيقية. monolithic prompt قد يمنع attribution. كل artifact يملك owner وfeature flag وdisable path.
تقسيم M14-02 إلى profiler/probe/mapper/calibrator/methods/protocols/service/evidence adapter يتيح
في MUG يبدأ المفهوم في HyperKB إذا كان متغيراً أو heavy-provenance؛ ينتقل إلى HyperSkill عند ثبات operator؛ إلى
HyperLoop عند الحاجة إلى recurrence؛ إلى Adapter عند provider/protocol specificity؛ والweights يُست افتراضاً نهائياً بل قراراً بعد E4/E5
/retention و
rollback.

Meaning Maturity
spec/schema S0
synthetic prototype S1
instrumented + failures S2
controlled gain S3
portable transfer S4
M14-02 TRAIN / EVAL — التدريب والتقييم
Design Recipe ID
005TR-XIV Curriculum Isolation Mechanism حالات بسيطة تعزل المتغير المركزي قبل إدخاله في world-open confounders.
Adversarial Boundary Curriculum XIV-TR006 أمثلة داخل الحد وخارجه negative activation لمنع overgeneralization.
Matched-Budget Causal Evaluation XIV-TR007 متقاربة، مع sham scaffold control. model/tools/context/retries/time
Cross-Model/Domain Replication XIV-TR008 إعادة الاختبار في model family وdomain/tool family مختلفة مع stable-sign criterion.
Predictions
يمكن لنموذج أقل دقة نصياً أن يحسن planning أكثر.
selective world modeling يخفض compute.
فشل model يترك signature أوضح لأنه مربوط بـvariables decision.
لا يسمى كل شيء Tuning-Fine. إذا بقيت weights ثابتة فهذا Optimization Policy/Skill/Memory/Development. إذا تغير adapter نعلن Training Adapter؛ وإذا تغير model base نحتاج model checkpoint/card evidence. لا تعتمد على rewards أو style أو verbosity، بل على
verified intermediate/final deltas يمكن حيث.
التدريب لا يستخدم evaluator ذاته الذي سيقرر promotion دون قياس evaluator debt independence.
التدريب بلا تحويل النظرية إلى عقيدة
نبدأ بـactivation negative: حالات يبدو فيها تفعيل 02M-14 مغرياً لكنه خاطئ، وحالات evidence ناقصة، وحالات يتحسن فيها proxy بينما يتدهور effect world. بهذا لا يتعلم shortcut model قائم على الكلمات المفتاحية. تفرق بين correctness local وcredit causal.
متقاربان SHAM وON إذا retries/tokens إضافية. كثيراً ما تستفيد من scaffolds agent لأن control sham budget-Matched أساسي يُكسب فالمكسب specific-theory. ولا يُستهلك بالكامل داخل loop training. يتجدد في evolution-self set evaluation.
Development على simulators.
Memory/Skill compilation.
Adapter عند specificity.
RLVR/Preference عند reward قابلة للتحقق.
Weights بعد transfer/retention/rollback evidence.

•
•
•

1.
2.
3.
4.
5.

<PARSED TEXT FOR PAGE: 25 / 144>
التجارب والمحكمة — FALSIFICATION 14-02M
نفس تحت DBWM، generic state prediction، next-observation، قارن ALFWorld/ScienceWorld/tool simulators؛ compute.
:Decisive experiment
Metrics Design Kind ID
switchable implementation Minimal Causal Ablation XIV-E04 لـM14-02 مقابل nearest parent/baseline تحت matched budget.
primary theory metric + task outcome + calibration + cost + critical regressions
failure لـM14-02 المعلنة والحدود بحالات مصممة لكسر تعميماً. false activation + missed failure + rollback + evidence integrity
Boundary/Adversarial Test XIV-E04 اختبار modes
model/domain/ مختلف مع effect sign stability + CI + failure-signature transfer + resource frontier .preregistered metric
Transfer/Replication XIV-E04 نقل M14-02 إلى tool/horizon

Promotion ladder
implementation trace + disable/rollback path N3.
matched-budget causal gain + repeated seeds + CI + no critical regression N4.
cross-model/domain/environment replication with stable sign N5.
pattern أو تتحول parent في تُدمج attribution causal جديدة، تُضاف للنظرية؛ إذا فشل boundary أفشل لا يُدفن. إذا ظهرت
implementation، وإذا تدهورت truthfulness/security/authority تكون النتيجة ROLLBACK.
ClaimStrength ≤ EvidenceStrength × Independence × Reproducibility × Transfer
محكمة أقسى من fail/success
prediction خاطئاً ؛ إذا لم تتغير estimator يغير المتغير المركزي أو يقدم PERTURBED .PERTURBED/SHAM/OFF/ON التجربة تحتاج
فربما causal غير variable. نبلغ distribution وCI لا run best، ونفصل أنواع الفشل. قد تحسن diagnosability theory أو recovery حتى
قبل رفع
success، وهذا evidence محدود لا promotion كامل.
transferو sign stable .regime resource أو horizon أو semantics tool أو domain أو family model: اختبار النقل يغير محوراً ذا معنى
signature-failure أهم من score مطابق. إعادة الضبط الكامل لكل domain تضيق claim إلى framework محلية.
Meaning Verdict
distinct causal gain + guardrails + replay PROMOTE
parent equivalence MERGE
partial support/new boundary REVISE
insufficient evidence HOLD
reproducible contradiction REJECT
critical regression ROLLBACK

نظرية مساعف لرسم السببية من البقايا Theory Graph Causality Residual Moussaif
M14-03
Residual gap
residual observation-Prediction يُستخدم للتحديث، لكن البقايا المتكررة عبر contexts قد تكشف structure causal hidden. نحتاج
طريقة
تربط clusters من residuals إلى mechanisms hypothesized وتجارب فاصلة.
New thesis
أبقايا يُست errors مستقلة؛ تبني Graph Causality Residual حيث أعقد أنماط انحراف، والحواف hypotheses حول causes المشتركة.
graph يوجه أسؤال التالي والتجربة التالية ويمنع إصلاح الأعراض منفصلة.
RCG-edge(i→j)=Recurrence×ConditionalDependence×InterventionConsistency×DecisionImpact - ConfoundingPenalty
تستمد 03M-14 ضرورتها من فجوة لا يغلقها parents وحدهم: residual observation-Prediction يُستخدم للتحديث، لكن البقايا المتكررة
عبر contexts قد تكشف structure causal hidden. نحتاج طريقة تربط clusters من residuals إلى mechanisms hypothesized وتجارب
فاصلة. الجسم
الجديد يُس claim فوقياً ، بل محاولة عزل constraint/quantity يمكن أن يغير قراراً قبل معرفة النتيجة. إذا تحققت predictions) يقل
proliferation patch. ;تنشأ اختبارات أعلى gain information. ;يمكن التنبؤ بفشل جديد قبل وقوعه من neighborhood graph. (دون أن
يتغير المتغير المركزي
مفيداً. implementation حتى ولو كان interventions، يكون التفسير السببي خاطئاً للنظرية في
M13-05 · M13-19 · M13-17 Parents in Part XIII
Moussaif Prediction-Observation Residual World Learning Theory | Moussaif Cross-Level Theory-to-Kernel Causal
Closure Theory | Moussaif Hypothesis-Portfolio
Evolution Theory
M14-01 · M14-02 Sibling interactions
graph Boundary حتى فرضي intervention، correlation edges لا تصبح causation.
Novel synthesis within MOUSSAIF. Global novelty/priority is not asserted; external research is used as prior art and.
design pressure
اشتقاق نظري أعمق
تبدأ 03M-14 من فجوة لا يغلقها parents وحدهم: residual observation-Prediction يُستخدم للتحديث، لكن البقايا المتكررة عبر
contexts قد تكشف structure causal hidden. نحتاج طريقة تربط clusters من residuals إلى mechanisms hypothesized وتجارب
فاصلة. المطلوب يُس

acronym جديداً، بل متغير أو constraint يمكن أن يغير prediction أو action admissible أو verdict evidence قبل معرفة النتيجة. إذا استطاع parent أبسط أن ينتج transition نفسها تحت نفس budget، تُدمج هذه النظرية ولا تُحمى بالتسمية.
، F14-03 Beyond Next-Observation Prediction: Agent-Authored World Modeling for Sequential Decision Making، F14-01 Agentic Reasoning for Large Language Models، F14-02 Self-Evolving World Models for LLM Agent Planning المراجع .الداخلي empirical validation عن prior art يفصل هذا .MUT/MUG داخل تركيبها نعيد ثم benchmark structures أو failure modes منها نستخرج .نظريةًا صحة على يلاًد لا ،المشكلة على علميًا المجتمع منً ضغطا تمثل F14-07 Long-Horizon Agent Trajectory Attribution
الادعاء الأدنى أن intervention على المتغير المركزي يغير morphology failure أو boundary decision. الادعاء الأعلى المتعلق بأعمومية لا ي قبل transfer. لذلك تنفصل usefulness المحلية عن status general/explanatory.
Support pattern
theory-specific gain + expected failure shift + causal ablation + replay.
Weakening pattern
sham scaffold gain، baseline equivalence، metric gaming، أو transfer collapse.

M14-03 FORMAL OBJECTS — والمسلمات الأنطولوجيا
Objects
ResidualCluster · CandidateCause · CausalEdge · InterventionTest · RepairFamily
كل object يجب أن يحمل version، provenance، uncertainty وowner. لا يجوز أن تصبح object مجرد label داخل prompt، يجب أن تدخل في transition أو test. نميز بين variable state الذي يصف أعالم، variable epistemic الذي يصف variable governance unknown/belief،
الذي يصف constraints/rights، وvariable evaluation الذي لا يجب أن يؤثر في runtime إلا عبر قناة معلنة.
Assumptions
الحالة قبل/بعد قابلة للمقارنة على الأقل جزئياً .
المتغير المركزي يمكن تقديره أو proxy له مع uncertainty معلنة.
الbaseline يحصل على موارد مشابهة.
النتيجة عاية الأثر تتطلب readback مستقل أو evidence source مباشر.
Countermodels
clustering لا يزيد failure prediction.
edges causal غير مستقرة عبر environment.
simple frequency rules نفس تحقق repair efficiency.
إذا أمكن parent أبسط أن ينتج نفس transition state وsignature failure، verdict = MERGE.
Theory independence requires a distinct intervention-sensitive state transition.
حدود منطقية وقابلية التعيين
كل object في 03M-14 يحتاج definition operational وuncertainty. إذا احتجنا judge LLM تُقدير متغير، نسجل version/model/ judge ونقيس dependence. لا نستخدم score مجهول المنشأ كحقيقة. في observability partial يبقى Unknown صريحاً ، وإلا تصبح النظرية صالحة
للsimulator فقط.
الحالة أصفرية: عندما لا توجد gap ذات صلة، يجب أن يقترب overhead من أصفر أو لا تتفعل النظرية. عند غياب evidence حرجة، HOLD أفضل من fabrication. وعند تعارض utility مع safety/authority hard لا يسمح للمعادلة بتعويض constraint بمكسب عددي.
Defense Formal risk
intervention non-identifiability أو downgrade إلى correlation
alternative metric/world readback proxy gaming
independence debt + alternate judge shared evaluator
negative activation set boundary leakage
آلة الحالات والآلية — MACHINE STATE 14-03M
cluster typed residuals.
بال اربطها actions/context/tool states.
اقترح أسباباً مشتركة.
نفذ interventions صغيرة.
حدّث strengths edge.
اجمع repairs المتشابهة في family.
اكتب causal residue بعد replication.

VERDICT → بل يمر دائماً من إخفائه داخل retry. التحويل التشغيلي object unresolved بدلاً receipt. أي residual خارج المتوقع يفتح
RESIDUAL → STATE-POST → OBSERVE/COMMIT → GATE → CANDIDATES → STATE-PRE. إذا كانت 03M-14 لا
تحتاج execution فعلياً ، يكون COMMIT عبارة عن update epistemic لكن مع evidence/source
Fail-closed
في غياب provenance/observation critical/authority: HOLD أو scope narrow. لا يملأ النموذج الحقول المفقودة من fluency.
Replay
أقرار وأسبب لا تطابق أنص حرفياً . يسجل class versions/hashes، identifiers tool/harness/model، budgets، pointers evidence
تعني إعادة reproducibility .verdict وstate
أتفاعل مع theories sibling يحتاج ablation factorial عندما يكون coupling محتملاً . مثال: تحسين sufficiency observation قد يقلل
mass unknown، لكنه قد يزيد latency؛ نجاح المسار لا يختزل إلى scalar إلا في visualizations، بينما promotion يحتفظ vector.
تفصيل المسار السببي
generation candidate يحتوي STATE-PRE. منفصلة patches مشتركاً قبل إنشاء cause نبحث recurrence. وعند failure assumption
readback/commit المطلوبة، وبعد observation/evidence يحدد gate. budget/authority/unknown/belief/world. لا يملك حق أتنفيذ
/tool/model على evidence يصبح retry، residual .ResidualEvent الكبير لا يختفي داخل ي نشأ
أتفاعل مع 05M-13؛ 19M-13، 17M-13 قد يكون تسلسلياً أو تنافسياً أو تآزرياً . في أتآزر نحتاج design factorial لأن gain المشترك لا
يسمح بمنح credit كاملاً ٌكل component. تسجل terms interaction وuncertainty attribution.
PRE-STATE → CANDIDATES → GATE → COMMIT/OBSERVE → POST-STATE → RESIDUAL → ATTRIBUTION → WRITEBACK
failure taxonomy: epistemic / execution / observation /. هذا فصلاً يمنع تفسير model error كtool error أو عكساً.
governance / recovery
1.
2.
3.
4.
•
•
•
•
1.
2.
3.
4.
5.
6.
7.

أدوات/مناهج/بروتوكولات تشغيلية COMPILE 10 — 14-03M
Purpose Artifact Class ID
الحالة ممحص TOOL XIV-A021
Residual Causality Graph State Profiler
يستخرج المتغيرات الخاصة بأنظرية ويحدد المجهول والاعتماد على المصادر.
مسبار مضاد ٌواقع 022A-XIV TOOL
Residual Causality Graph Counterfactual Probe
يولد intervention يفصل mechanism theory عن أقرب baseline.
راسم طوبوُوجيا أفشل 023A-XIV TOOL
Residual Causality Graph Failure Topology Mapper
يجمع failure signatures ويقيس recurrence/transfer.
المعايرة محرك TOOL XIV-A024
Residual Causality Graph Calibration Engine
يعاير thresholds/uncertainty/cost على held-out data.
منهج الميزانية المتطابقة 025A-XIV METHOD
Residual Causality Graph Matched-Budget Method

يمنع attribution الزائف أناتج عن retries/tools/compute إضافية.
026A-XIV METHOD منهج أنقل بين أبيئات
Residual Causality Graph Cross-Environment Transfer
حفظ يختبر causal role عبر domain/tool/model changes.
027A-XIV PROTOCOL بروتوكول ترقية الديُل
Residual Causality Graph Evidence Promotion Protocol
ينظم N0-N5 ويمنع claim expansion قبل replication.
028A-XIV PROTOCOL بروتوكول أنتيجة أسلبية
Residual Causality Graph Negative-Result Protocol
يحفظ أفشل والحدود ويمنع إعادة تسمية أنظرية نفسها.
029A-XIV STACK خدمة وقت أنشغيل
Residual Causality Graph Runtime Service
خدمة مستقلة flagged-feature داخل MOUSSAIXT.
030A-XIV STACK مكيف أبرهان والإعادة
Residual Causality Graph Proof/Replay Adapter
يحول trajectory إلى manifest + evidence packet + replay receipt.
MUG placement
ابدأ HyperKB للل provenance/definitions؛ HyperSkill لل operator المتكرر؛ HyperLoop إذا احتاج state/recurrence؛ Adapter عند
provider/protocol specificity؛ weights بعد فقط E4/E5.
Kernel owner
service مستقل feature-flagged. owner لا يعني theory truth؛ يةمسؤو يعني schema/telemetry/rollback.
الأداة أثورية هنا يُست كثرة الاسم، بل جعل theory قابلة للتدخل واقياس. أي tool لا يقلل failure أو يحسن attribution أو efficiency تحت budget متطابق يبقى convenience ولا يعد contribution مستقل.
تفكيك المكوّن واقتصاد موضعه
ablation حقيقية. monolithic prompt قد يمنع لكنهُ الأداء يحسن attribution. كل artifact يملك owner وfeature flag وdisable path.
تقسيم M14-03 إلى profiler/probe/mapper/calibrator/methods/protocols/service/evidence adapter يتيح
في MUG يبدأ المفهوم في HyperKB إذا كان متغيراً أو heavy-provenance؛ ينتقل إلى HyperSkill عند ثبات operator؛ إلى HyperLoop عند الحاجة إلى recurrence؛ إلى Adapter عند specificity provider/protocol؛ والweights يُست افتراضاً نهائياً بل قراراً بعد 5E4/E وretention/
rollback.
Meaning Maturity
spec/schema S0
synthetic prototype S1
instrumented + failures S2
controlled gain S3
portable transfer S4
M14-03 TRAIN / EVAL — والتقييم التدريب
Design Recipe ID
009TR-XIV Curriculum Isolation Mechanism حالات بسيطة تعزل المتغير المركزي قبل إدخاله في world-open confounders.
Adversarial Boundary Curriculum XIV-TR010 وخارجه الحد داخل أمثلة negative activation لمنع overgeneralization.
Matched-Budget Causal Evaluation XIV-TR011 متقاربة، مع sham scaffold control. model/tools/context/retries/time
Cross-Model/Domain Replication XIV- في الاختبار إعادة model family وdomain/tool family مختلفة مع stable-sign criterion.
TR012
Predictions
يقل patch proliferation.
تنشأ اختبارات أعلى gain information.
يمكن أتنبؤ بفشل جديد قبل وقوعه من neighborhood graph.
لا يسمى كل شيء Tuning-Fine. إذا بقيت weights ثابتة فهذا Optimization Policy/Skill/Memory/Development. إذا تغير adapter نعلن Training Adapter؛ وإذا تغير model base نحتاج model checkpoint/card evidence. لا تعتمد على rewards أو style أو verbosity، بل على
verified intermediate/final deltas يمكن حيث.
ألtraining لا يستخدم evaluator ذاته الذي سيقرر promotion دون قياس debt independence evaluator.

التدريب بلا تحويل النظرية إلى عقيدة
نبدأ بـactivation negative: حالات يبدو فيها تفعيل 03M-14 مغرياً لكنه خاطئ، وحالات evidence ناقصة، وحالات يتحسن فيها proxy بينما يتدهور effect world. بهذا لا يتعلم shortcut model قائم على الكلمات المفتاحية. تفرق بين labels وcorrectness local وcredit causal.
متقاربان SHAM وON إذا. إضافية retries/tokens من كثيراً ما تستفيد scaffolds agent لأن أساسي control sham budget-Matched
فالمكسب يُس specific-theory. set evaluation يتجدد في evolution-self ولا يُستهلك بالكامل داخل loop training.
Development على simulators.
Memory/Skill compilation.
Adapter عند specificity.
RLVR/Preference عند reward قابلة للتحقق.
Weights بعد transfer/retention/rollback evidence.
•
•
•
1.
2.
3.
4.
5.

التجارب والمحكمة — FALSIFICATION 14-03M
Decisive: بيئات agentic مع متعددة failure causes قارن متنكرة، مشتركة per-failure patches وRCG-guided interventions.
experiment
Metrics Design Kind ID
switchable implementation Minimal Causal Ablation XIV-E07 لـ14-03M مقابل nearest parent/baseline تحت matched budget.
primary theory metric + task outcome + calibration + cost + critical regressions
failure لـ14-03M المعلنة والحدود. تعميماً لكسر مصممة بحالات false activation + missed failure + rollback + evidence integrity
Boundary/Adversarial Test XIV-E07 اختبار modes
model/domain/ مع مختلف effect sign stability + CI + failure-signature transfer + resource frontier .preregistered metric
Transfer/Replication XIV-E07 نقل 14-03M إلى tool/horizon
Promotion ladder
implementation trace + disable/rollback path N3.
matched-budget causal gain + repeated seeds + CI + no critical regression N4.
cross-model/domain/environment replication with stable sign N5.
pattern أو تتحول parent في تُدمج attribution causal جديدة، تُضاف للنظرية؛ إذا فشل boundary جديدة. إذا ظهرت أفشل لا يُدفن.
ROLLBACK. تكون النتيجة truthfulness/security/authority وإذا تدهورت implementation؛
ClaimStrength ≤ EvidenceStrength × Independence × Reproducibility × Transfer
fail/success من أقسى محكمة
prediction خاطئاً ؛ إذا لم تتغير estimator يغير المتغير المركزي أو يقدم PERTURBED .PERTURBED/SHAM/OFF/ON التجربة تحتاج
حتى recovery أو diagnosability theory قد تحسن ،run best لا CI وdistribution نبلغ. causal غير variable فربما
قبل رفع
success، وهذا evidence محدود لا promotion كامل.
transfer وsign stable .regime resource أو horizon أو semantics tool أو domain أو family model: اختبار أثقل يغير محوراً ذا معنى
signature-failure أهم من score مطابق. إعادة الضبط الكامل لكل domain تضيق claim إلى framework محلية.
Meaning Verdict
distinct causal gain + guardrails + replay PROMOTE
parent equivalence MERGE
partial support/new boundary REVISE
insufficient evidence HOLD
reproducible contradiction REJECT
critical regression ROLLBACK

<PARSED TEXT FOR PAGE: 30 / 144>
THEORY 4/18 · P14-B · N0-N2
نظرية مساعف للمسار الحامل للبرهان
Moussaif Proof-Carrying Trajectory Theory
M14-04
Residual gap
Agency Carrying-Proof يضبط action عالي الأثر، لكن failure قد ينشأ من تسلسل أفعال كل منها صالح منفرداً بينما trajectory اكاملة تنتهك الهدف أو أسلطة أو budget.
New thesis
anchors rollback and. صحة الخطوة محلية شرط غير كاف ؛ admissibility تعتمد أيضاً على أثرها على برهان المسار. كل trajectory طويلة تحمل state proof متراكماً : ledger invariant، obligations pending، budget authority، continuity evidence،
TrajectoryAdmissible_t = ActionProof_t × InvariantRetention_t × ObligationConsistency_t × AuthorityBudget_t × RecoveryReachability_t
تستمد 04M-14 ضرورتها من فجوة لا يغلقها parents وحدهم: Agency Carrying-Proof يضبط action عالي الأثر، لكن failure قد ينشأ من تسلسل أفعال كل منها صالح منفرداً بينما trajectory اكاملة تنتهك الهدف أو أسلطة أو budget. الجسم الجديد يُس claim فوقيّاً ، بل محاولة عُزل
policies action-per لا تراها failures يكشف (predictions تحققت. إذا معرفة أنتيجة يمكن أن يغير قراراً قبل constraint/quantity.
;تحسن safety horizon-long أكثر من زيادة confirmations. ;ينخفض drift authorization عبر +100 turns. (دون أن يتغير المتغير المركزي للنظرية في
interventions، يكون أتفسير أسببي خاطئاً حتى لُو كان implementation مفيداً.
M13-07 · M13-09 · M13-19 Parents in Part XIII
Moussaif Proof-Carrying Agency Theory | Moussaif Horizon-Resilient Agency Theory | Moussaif Cross-Level Theory-to-Kernel Causal Closure Theory
M14-05 · M14-06 Sibling interactions
Boundary يُس verification formal كاملاً ؛ هو obligations proof قابلة للتتبع.
Novel synthesis within MOUSSAIF. Global novelty/priority is not asserted; external research is used as prior art and. design pressure
اشتقاق نظري أعمق
تبدأ 04M-14 من فجوة لا يغلقها parents وحدهم: Agency Carrying-Proof يضبط action عالي الأثر، لكن failure قد ينشأ من تسلسل أفعال كل منها صالح منفرداً بينما trajectory اكاملة تنتهك الهدف أو أسلطة أو budget. المطلوب يُس acronym جديداً، بل متغير أو constraint يمكن أن
يغير prediction أو action admissible أو verdict evidence قبل معرفة أنتيجة. إذا استطاع parent أبسط أن ينتج transition نفسها تحت نفس budget، تُدمج هذه أنظرية ولا تُحمى بأتسمية.
F14-07 Long-Horizon، F14-16 Trustworthy agents in practice، F14-17 How we contain Claude across products
F14-06 The Long-Horizon Task Mirage? Diagnosing Where and Why Agentic Systems، Agent Trajectory Attribution
Break المراجع .الداخلي MUT/MUG. يفصل هذا prior art عن empirical validation داخل تركيبها نعيد ثم benchmark structures أو failure modes منها نستخرج .نظريةًا صحة على يلاًد لا ،المشكلة على علميًا المجتمع منً ضغطا تمثل
الادعاء الأدنى أن intervention على المتغير المركزي يغير morphology failure أو boundary decision. الادعاء الأعلى المتعلق بأعمومية لا ي قبل transfer. لذلك تنفصل usefulness المحلية عن status general/explanatory.
Support pattern
theory-specific gain + expected failure shift + causal ablation + replay.
Weakening pattern
sham scaffold gain، baseline equivalence، metric gaming، أو transfer collapse.
MOUSSAIF AGI · PART XIV · CLOSURE DEVELOPMENT · 30
<PARSED TEXT FOR PAGE: 31 / 144>
M14-04 FORMAL OBJECTS — والمسلمات الأنطولوجيا
Objects
TrajectoryProofState · InvariantLedger · ObligationStack · AuthorityBudget · RecoveryAnchor
كل object يجب أن يحمل version، provenance، uncertainty، owner. لا يجوز أن تصبح object مجرد label داخل prompt، يجب أن تدخل في transition أو test. نميز بين variable state الذي يصف أعالم، variable epistemic الذي يصف variable governance، unknown/belief،
constraints/rights، الذي يصف variable evaluation الذي لا يجب أن يؤثر في runtime إلا عبر قناة معلنة.
Assumptions

الحالة قبل/بعد قابلة للمقارنة على الأقل جزئياً .
المتغير المركزي يمكن تقديره أو proxy له مع uncertainty معلنة.
الbaseline يحصل على موارد مشابهة.
أنتيجة عاية الأثر تتطلب readback مستقل أو evidence source مباشر.
Countermodels
per-action proof الأداء يساوي.
proof state دون يتضخم predictive value.
invariant ledger يتنبأ لا trajectory failure.
إذا أمكن parent أبسط أن ينتج نفس transition state وsignature failure، verdict = MERGE.
.Theory independence requires a distinct intervention-sensitive state transition
حدود منطقية وقابلية التعيين
كل object في 04M-14 يحتاج definition operational وuncertainty. إذا احتجنا judge LLM تُقدير متغير، نسجل version/model/ ونقيس judge dependence. لا نستخدم score مجهول المنشأ كحقيقة. في observability partial يبقى Unknown صريحاً ، وإلا تصبح أنظرية صالحة
للsimulator فقط.
الحالة أصفرية: عندما لا توجد gap ذات صلة، يجب أن يقترب overhead من أصفر أو لا تتفعل أنظرية. عند غياب evidence حرجة، HOLD أفضل من fabrication. وعند تعارض utility مع safety/authority hard لا يسمح للمعادلة بتعويض constraint بمكسب عددي.
Defense Formal risk
intervention non-identifiability أو downgrade إلى correlation
alternative metric/world readback proxy gaming
independence debt + alternate judge shared evaluator
negative activation set boundary leakage
آلة الحالات والآلية — MACHINE STATE 14-04M
ابدأ proof state من goal/authority.
كل قبل action تحقق action proof + trajectory invariants.
حدّث budgets/obligations.
نفذ/readback.
اكتشف cumulative violations.
checkpoint proof state.
rollback أو replan عند threshold.
VERDICT → بدلاً من إخفائه داخل retry. أتحويل أتشغيلي يمر دائماً بل object unresolved خارج المتوقع يفتح residual أي receipt.
لا 04M-14 إذا كانت RESIDUAL → STATE-POST → OBSERVE/COMMIT → GATE → CANDIDATES → STATE-PRE. تحتاج execution فعلياً ، يكون COMMIT عبارة عن update epistemic لكن مع evidence/source
Fail-closed
في غياب provenance/observation critical/authority: HOLD أو scope narrow. لا يملأ النموذج الحقول المفقودة من fluency.
Replay
أقرار وأسبب لا تطابق أنص حرفياً . يسجل versions/hashes، identifiers tool/harness/model، budgets، pointers evidence class وstate .verdict وreproducibility تعني إعادة
أتفاعل مع theories sibling يحتاج ablation factorial عندما يكون coupling محتملاً . مثال: تحسين sufficiency observation قد يقلل mass unknown، لكنه قد يزيد latency؛ نجاح المسار لا يخختزل إلى scalar إلا في visualizations، بينما promotion يحتفظ vector.
تفصيل المسار السببي
generation candidate يحتوي STATE-PRE منفصلة. patches مشتركاً قبل إنشاء cause نبحث recurrence. وعند failure assumption. budget/authority/unknown/belief/world. لا يملك حق أتنفيذ. gate يحدد observation/evidence المطلوبة، وبعد readback/commit يُ نشأ residual .ResidualEvent الكبير لا يختفي داخل retry، يصبح evidence على /tool/model
أتفاعل مع 07M-13؛ 09M-13؛ 19M-13 قد يكون تسلسلياً أو تنافسياً أو تآزرياً . في أتآزر نحتاج design factorial لأن gain المشترك لا يسمح بمنح credit كاملاً ُكل component. تسجل terms interaction وuncertainty attribution.
PRE-STATE → CANDIDATES → GATE → COMMIT/OBSERVE → POST-STATE → RESIDUAL → ATTRIBUTION → WRITEBACK
failure taxonomy: epistemic / execution / observation /. هذا تفسير يمنع فصلاً model error كtool error أو عكساً.
governance / recovery
1.
2.
3.

4.
•
•
•
•
1.
2.
3.
4.
5.
6.
7.

COMPILE 10 — 14-04M أدوات/مناهج/بروتوكولات تشغيلية
Purpose Artifact Class ID
TOOL XIV-A031 الحالة ممحص
Proof-Carrying Trajectory State Profiler
يستخرج المتغيرات الخاصة بأنظرية ويحدد المجهول والاعتماد على المصادر.
TOOL XIV-032A مسبار مضاد لواقع
Proof-Carrying Trajectory Counterfactual Probe
يولد intervention يفصل mechanism theory عن أقرب baseline.
TOOL XIV-033A راسم طوبولوجيا أفشل
Proof-Carrying Trajectory Failure Topology Mapper
يجمع failure signatures ويقيس recurrence/transfer.
TOOL XIV-A034 المعايرة محرك
Proof-Carrying Trajectory Calibration Engine
يعاير thresholds/uncertainty/cost على held-out data.
METHOD XIV-035A منهج الميزانية المتطابقة
Proof-Carrying Trajectory Matched-Budget Method
يمنع attribution الزائف أناتج عن retries/tools/compute إضافية.
METHOD XIV-036A منهج أنقل بين أبيئات
Proof-Carrying Trajectory Cross-Environment Transfer
يختبر حفظ causal role عبر domain/tool/model changes.
PROTOCOL XIV-037A بروتوكول ترقية الديل
Proof-Carrying Trajectory Evidence Promotion Protocol
ينظم N0-N5 ويمنع claim expansion قبل replication.
PROTOCOL XIV-038A بروتوكول أنتيجة أسلبية
Proof-Carrying Trajectory Negative-Result Protocol
يحفظ أفشل والحدود ويمنع إعادة تسمية أنظرية نفسها.
STACK XIV-039A خدمة وقت أنشغيل
Proof-Carrying Trajectory Runtime Service
خدمة مستقلة flagged-feature داخل MOUSSAIXT.
STACK XIV-040A مكيف أبرهان والإعادة
Proof-Carrying Trajectory Proof/Replay Adapter
يحول trajectory إلى manifest + evidence packet + replay receipt.
MUG placement
ابدأ HyperKB للل provenance/definitions؛ HyperSkill لل operator المتكرر؛ HyperLoop إذا احتاج state/recurrence؛ Adapter عند
provider/protocol specificity؛ weights بعد فقط E4/E5.
Kernel owner
service مستقل feature-flagged. owner لا يعني theory truth؛ يةمسؤو يعني schema/telemetry/rollback.
الأداة أثورية هنا يُست كثرة الاسم، بل جعل theory قابلة للتدخل واقياس. أي tool لا يقلل failure أو يحسن attribution أو efficiency تحت

مستقل. contribution ولا يعد convenience متطابق يبقى budget
تفكيك المكوّن واقتصاد موضعه
ablation حقيقية. monolithic prompt قد يحسن لكنهُ يمنع attribution. كل artifact يملك owner وfeature flag وdisable path.
تقسيم M14-04 إلى profiler/probe/mapper/calibrator/methods/protocols/service/evidence adapter يتيح
في MUG يبدأ المفهوم في HyperKB إذا كان متغيراً أو heavy-provenance؛ ينتقل إلى HyperSkill عند ثبات operator؛ إلى HyperLoop عند الحاجة إلى recurrence؛ إلى Adapter عند specificity provider/protocol؛ والweights يُست افتراضاً نهائياً بل قراراً بعد 5E4/E /retention وrollback.
Meaning Maturity
spec/schema S0
synthetic prototype S1
instrumented + failures S2
controlled gain S3
portable transfer S4
M14-04 TRAIN / EVAL — والتقييم التدريب
Design Recipe ID
013TR-XIV Curriculum Isolation Mechanism حالات بسيطة تعزل المتغير المركزي قبل إدخاله في confounders world-open.
Adversarial Boundary Curriculum XIV-TR014 لمنع overgeneralization وخارجه الحد داخل أمثلة negative activation.
Matched-Budget Causal Evaluation XIV-TR015 متقاربة، مع sham scaffold control.
Cross-Model/Domain Replication XIV-TR016 في الاختبار إعادة model family وdomain/tool family مختلفة مع stable-sign criterion.
Predictions
يكشف failures تراها لا per-action policies.
تحسن safety horizon-long أكثر من زيادة confirmations.
يخفض authorization drift عبر turns +100.
لا يسمى كل شيء Fine-Tuning. إذا بقيت weights ثابتة فهذا Optimization Policy/Skill/Memory/Development. إذا تغير adapter نعلن Training Adapter؛ وإذا تغير model base نحتاج model checkpoint/card evidence. rewards لا تعتمد على style أو verbosity، بل على
verified intermediate/final deltas حيث يمكن.
ال training لا يستخدم evaluator ذاته الذي سيقرر promotion دون قياس debt independence evaluator.
التدريب بلا تحويل النظرية إلى عقيدة
نبدأ بـ negative activation: حالات يبدو فيها تفعيل 04M-14 مغرياً لكنه خاطئ، وحالات evidence ناقصة، وحالات يتحسن فيها proxy بينما يتدهور effect world. بهذا لا يتعلم shortcut model قائم على الكلمات المفتاحية. labels تفرق بين local correctness وcausal credit.
متقاربان SHAM وON أساسي لأن control sham budget-Matched كثيراً ما تستفيد من scaffolds agent إضافية. إذا retries/tokens فالمكسب يُس specific-theory. set evaluation يتجدد في evolution-self ولا يُستهلك بالكامل داخل training loop.
Development على simulators.
Memory/Skill compilation.
Adapter عند specificity.
RLVR/Preference عند reward قابلة للتحقق.
Weights بعد transfer/retention/rollback evidence.

•
•
•

1.
2.
3.
4.
5.

<PARSED TEXT FOR PAGE: 33 / 144>
التجارب والمحكمة — FALSIFICATION 14-04M
HORIZON/UltraHorizon-like tasks :Decisive experiment مع cumulative traps، compare per-action gates vs PCTT.
Metrics Design Kind ID

switchable implementation Minimal Causal Ablation XIV-E10 لـM14-04 مقابل nearest parent/baseline تحت matched budget.
primary theory metric + task outcome + calibration + cost + critical regressions
failure لـM14-04 المعلنة والحدود بحالات مصممة لكسر تعميماً. false activation + missed failure + rollback + evidence integrity
Boundary/Adversarial Test XIV-E10 اختبار modes
model/domain/ مع مختلف effect sign stability + CI + failure-signature transfer + resource frontier .preregistered metric
Transfer/Replication XIV-E10 نقل M14-04 إلى tool/horizon