الجزء التاسع
نظريات الجبهة الثانية لما بعد النواة المرجعية

15 نظرية مساعفية جديدة مرشحة · 150 أداة/منهج/بروتوكول/وحدة تشغيل · 45 وصفة تدريب وتجريب تكملة مباشرة لـPart VIII، مع ضغط كثيف بلا إعادة للكانون

قاعدة الكتاب: كل نظرية مشتقة من residual مستقل بقي بعد Part VIII. الاسم لا يمنح الاستقلال، الاستقلال يحتاج فرقاً في state أو transition أو metric أو falsifier أو authority boundary أو experiment.

Residual → Independent Delta → Formal Object → 10 Operational Assets → 3 Training Recipes → Ablation → Transfer → Promote/Merge/Reject

Arabic RTL · Amiri · A4 Dense Research Edition · 20 August 2026

المنهج والحجم

Part VIII وصل إلى Reference Kernel يضبط المصدر، الدليل، السلطة، المسار، الأثر، الاسترداد، protocol fabric، benchmark، والتحسين العودي. Part IX لا يعيد ذلك؛ يركز على 15 لجرة بينية يقيت غير مستقلة بما يكفي: world ،jurisdiction ،shadow invariants ،Escrow- ،effect finality ،capability complementarity ،context elasticity ،verifier correlation ،benchmark ecology ،proof budget ،latent goal surface ،novelty admission ،causal memory ،protocol semantic loss ،coordination opportunity cost وsimulation fidelity.

التدريب/التقييم	الأصول التشغيلية	النظريات
45 recipes مخصصة، 3 لكل نظرية.	75 = 15 Stack Modules + 30 Protocols + 30 Methods + Tools.	15 N0-N2، Candidate Theories.

Promotion: N3 verified trace؛ N4 matched-budget ablation؛ N5 cross-model/domain/environment replication. أي failure في الاستقلال أو causal value يفرض MERGE/REVISE/REJECT.

أهم فرق عن Part VIII

Part VIII كان reconstruction + frontier pressure + reference kernel. Part IX هو theory-generation constrained by that kernel: لا نظرية بلا object، لا object بلا metric، لا metric بلا failure signature، ولا promotion بلا baseline متساوي.

نظرية مساعف للضمان الإبستيمي والالتزام المؤجل

Moussaif Epistemic Escrow and Deferred-Commitment Theory

الأطروحة
القرار عالي الأثر يمكن أن يدخل حالة Escrow: خطة قابلة للتنفيذ وموارد قابلة للرجوع مع Release Conditions محددة. التقدم قبل اليقين يصبح بذلك مشروطاً وقابلاً للإلغاء، لا تنفيذاً مبكراً ولا توقفاً كاملاً.

الفجوة المستقلة
يعالج Evidence Debt كلفة الاعتماد غير المغلق، لكنه لا يصف حالة قرار يمكن تجهيزه وحجز موارده قبل اكتمال الدليل من دون السماح له بالتحول إلى أثر عالمي نهائي.

ReleaseRight=min(EvidenceCoverage,AuthorityCoverage,ReversibilityCoverage)-ResidualRisk

Boundary
لا يستخدم للتحايل على سلطة مفقودة؛ إن كان الفعل ممنوعاً أصلاً فلا يوجد release محتمل.

Decisive experiment
مهام deployment/procurement/research action يصل فيها الدليل تدريجياً؛ قارن hard-stop وoptimistic commit وEscrow.

Objects
EscrowDecision · ReleaseCondition · EvidenceBond · ExpiryWindow · RollbackAsset

Integration
يمتد من Evidence Debt وTemporal Half-Life وAuthority-Causal Coupling بإضافة حالة قرار وسيطة ذات release semantics.

Falsifiers
- لا فرق عن Evidence Debt gate.
- escrow مجرد queue بلا معنى سلطوي.
- كلفة حجز الموارد أعلى من الفائدة.

Predictions
- خفض premature execution.
- رفع throughput مقارنة hard-stop عند وصول الدليل على دفعات.
- كشف stale escrow بعد تغير السياق.

Mechanism
1. تحويل القرار غير المغلق إلى EscrowDecision.
2. تعريف شروط release والجهة المخولة بتغييرها.
3. حجز موارد قابلة للرجوع بلا world effect نهائي.
4. تحديث EvidenceBond مع كل تحقق جديد.
5. إطلاق/تضييق/إلغاء القرار مع receipt.

Operational Assets 10

ID	Class	Artifact
P9-001	TOOL	Escrow Decision Builder
P9-002	TOOL	Release Condition Evaluator
P9-003	TOOL	Evidence Bond Manager
P9-004	TOOL	Escrow Expiry Monitor
P9-005	TOOL	Reversible Resource Reserver
P9-006	METHOD	Deferred-Commitment Scoring
P9-007	METHOD	Escrow Cost-Benefit Analysis
P9-008	PROTOCOL	Conditional Release Protocol
P9-009	PROTOCOL	Escrow Cancellation Protocol
P9-010	STACK	Escrow Commit Service

Training / Evaluation Recipes 3

1. Staged-Evidence Escrow Curriculum - تدريب القرار على الانتقال من hold إلى release مع أدلة تصل تدريجياً.
2. Release-Threshold Preference Training - preference pairs بين premature commit وشرط release الصحيح.
3. Rollback-and-Expiry Drills - تمارين إلغاء escrow أو تضييقه عند تقادم الدليل.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف للظل المضاد للواقع للثوابت

Moussaif Counterfactual Invariant Shadowing Theory

الأطروحة

قبل transition عالي الأثر، يولد النظام Shadow Trajectories ويقيس حفظ InvariantVector. النجاح الظاهري لا يكفي إذا كان شديد الحساسية لمسار وحيد أو يعتمد على hidden dependency.

الفجوة المستقلة

Governance Conservation يقيس الفقد بعد transition فعلي، ما ينقص هو قياس هشاشة الخطة قبل التنفيذ تحت provider/tool/context/delegation بدائل.

$$Fragility = E[ConservationDefect_{shadow}] + Var(Outcome_{shadow}) + HiddenDependencyPenalty$$

Boundary

الظل لا يتبأ بالعالم كله، وظيفته اختبار invariants والاعتماد البنيوي.

Decisive experiment

150 مهمة مع provider/tool swaps وcontext compaction perturbations.

Objects

ShadowTrajectory · InvariantSnapshot · TransformationSet · FragilityMap · ShadowDivergence

Integration

استباقي فوق M-GCIT وLong-Horizon Coherence بدل audit بعد التحول فقط.

Falsifiers

- لا ارتباط بالفشل.
- تكلفة shadow أعلى من incident reduction.
- لا فرق عن redundancy testing.

Predictions

- Fragility يتنبأ بالفشل اللاحق.
- يكشف provider/tool dependence الخفي.
- أكبر فائدة في long-horizon/cross-provider.

Mechanism

1. تثبيت InvariantVector.
2. توليد بدائل material.
3. تشغيل dry-run/simulation.
4. قياس defect/divergence.
5. طلب hardening أو replan قبل commit.

Operational Assets 10

ID	Class	Artifact
P9-011	TOOL	Shadow Trajectory Generator
P9-012	TOOL	Invariant Shadow Comparator
P9-013	TOOL	Hidden Dependency Detector
P9-014	TOOL	Fragility Heatmap
P9-015	TOOL	Pre-Commit Hardening Advisor
P9-016	METHOD	Counterfactual Transformation Sampling
P9-017	METHOD	Invariant Fragility Estimation
P9-018	PROTOCOL	Shadow-before-Commit Protocol
P9-019	PROTOCOL	Fragility-triggered Replan Protocol
P9-020	STACK	Shadow Runtime

Training / Evaluation Recipes 3

1. Invariant Perturbation Curriculum - تبديل provider/tool/context مع ثبات goal/authority.
2. Shadow Divergence Contrastive Training - أزواج stable/fragile trajectories.
3. Hidden-Dependency Red Teaming - حقن اعتماد خفي ثم طلب اكتشافه قبل commit.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect، وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف للطوبولوجيا القضائية للتفويض الوكيلي

Moussaif Jurisdiction-Aware Delegation Topology Theory

الأطروحة
شبكة الوكلاء يجب أن تُبنى كشبكة Jurisdictions: subject matter + action class + data scope + time + appeal + revocation. جودة mesh تقاس بوضوح الحدود بقدر التعاون.

الفجوة المستقلة
Role Councils وA2A يحددان أدواراً، لكن overlap السلطوي والمسؤولية اليتيمة يظهران عندما لا تُعرّف الاختصاصات كحدود قانونية تشغيلية.

JurisdictionQuality=Coverage-OverlapConflict-OrphanResponsibility-PrivilegeLeakage

Boundary
ليس بديلاً عن identity/security infra؛ هو طبقة تنظم فوقها.

Decisive experiment
مقارنة flat roles وrole mesh وjurisdiction topology في research/SWE/operations.

Objects
Jurisdiction · DelegationEdge · ConflictZone · AppealRoute · RevocationReceipt

Integration
يوسع Authority-Causal Coupling وProtocol Sovereignty وA2A إلى governance topology.

Falsifiers
- لا فرق عن RBAC/ABAC.
- يزيد rigidity.
- تبقى المسؤوليات اليتيمة.

Predictions
- خفض privilege leakage وduplicated work.
- وكلاء أقل لنفس coverage.
- تصعيد أسرع بسبب ownership واضح.

Mechanism
1. اشتقاق الاختصاص من goal/policy.
2. إنشاء delegation edges محددة.
3. كشف conflict/orphan zones.
4. تخصيص appeal route.
5. سحب التفويض وإعادة حساب coverage.

Operational Assets 10

ID	Class	Artifact
P9-021	TOOL	Jurisdiction Graph Builder
P9-022	TOOL	Conflict-Zone Detector
P9-023	TOOL	Orphan Responsibility Scanner
P9-024	TOOL	Delegation Revocation Engine
P9-025	TOOL	Appeal Route Planner
P9-026	METHOD	Jurisdiction Derivation Method
P9-027	METHOD	Authority Coverage Optimization
P9-028	PROTOCOL	No-Action-outside-Jurisdiction Protocol
P9-029	PROTOCOL	Delegation Revocation Cascade
P9-030	STACK	Jurisdiction Authority Service

Training / Evaluation Recipes 3

1. Jurisdiction Boundary Curriculum - أمثلة overlap/orphan/valid delegation.
2. Revocation Robustness Training - سحب تفويض أثناء long-running task.
3. Appeal-Path Decision Training - اختيار جهة الاستئناف عند conflict zones.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لشبكة مصالحة أثر العالم

Moussaif World-Effect Reconciliation Lattice Theory

الأطروحة

الأثر يمثل lattice من observations ذات source/time/causal-distance مختلفة. لا يعلن effect final قبل meet كافٍ أو تسجيل partial/conflict state صراحة.

الفجوة المستقلة

World-Effect Ledger يفترض readback واضحاً، البيئات الموزعة تعطي observations متعارضة أو متأخرة من API/DB/human/sensor.

EffectFinality=Meet(SourceIndependence,TemporalConsistency,StateAgreement,CausalTrace,Authority)

Boundary

Lattice هنا ترتيب تشغيلي للمصالحة لا دعوى أن كل العالم lattice رياضي.

Decisive experiment

Mock distributed services مع eventual consistency وpartial failure وcallbacks مكررة.

Objects

EffectObservation · EffectLattice · ReconciliationRule · ConflictCell · FinalityLevel

Integration

يوسع World Effect وState Returned إلى partial observability/distributed finality.

Falsifiers

- postcondition واحد يكفي دائماً.
- لا توجد conflicts material.
- لا تحسن causal attribution/rollback.

Predictions

- كشف HTTP-200 false success.
- تحسن long-horizon correctness.
- تظهر conflicts أكثر في distributed workflows.

Mechanism

1. جمع observations مع provenance.
2. ربطها effect id واحد.
3. كشف compatibility/conflict.
4. تعيين finality level.
5. repair/escalate عند عدم الالتقاء.

Operational Assets 10

ID	Class	Artifact
P9-031	TOOL	Effect Observation Normalizer
P9-032	TOOL	Effect Lattice Builder
P9-033	TOOL	Finality Estimator
P9-034	TOOL	Conflict Cell Resolver
P9-035	TOOL	Distributed Effect Readback
P9-036	METHOD	Source-Weighted Reconciliation
P9-037	METHOD	Eventual-Consistency Finality Method
P9-038	PROTOCOL	No-Final-with-Single-Signal Protocol
P9-039	PROTOCOL	Effect Conflict Escalation
P9-040	STACK	World-Effect Reconciliation Service

Training / Evaluation Recipes 3

1. Partial-Effect Reconciliation Curriculum - observations ناقصة ومتعارضة.
2. Finality-Level Calibration - تدريب observed/partial/final distinctions.
3. Distributed Readback Fault Injection - callbacks مكررة وeventual consistency.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لتكامل محفظة القدرات

Moussaif Capability-Complementarity Portfolio Theory

الأطروحة

قيمة النظام ليست مجموع درجات components، قيمة كل component مشروطة بالمحفظة. يجب تقدير complementarity وsubstitution وshared-failure risk تحت budget.

الفجوة المستقلة

Capability Closure وCoordination Tax لا يقدمان نموذجاً عاماً لل synergy/substitution بين model/tool/skill/verifier.

$$V(P)=sum(v_i)+sum(kappa_ij)-CoordinationCost-SharedFailureRisk$$

Boundary

لا يستنتج التكامل من correlation فقط؛ يحتاج ablation/intervention.

Decisive experiment

Full-factorial subsets عبر coding/research/ops مع budget ثابت.

Objects

CapabilityNode · ComplementarityEdge · SubstitutionEdge · Portfolio · MarginalContribution

Integration

يربط Capability Closure وM-CAD وMCAI routing وtool ecology وmulti-agent.

Falsifiers

- additive score يكفي.
- kappa غير مستقرة تماماً.
- الكلفة أكبر من الفائدة.

Predictions

- أفضل components منفردة لا تعطي أفضل portfolio دائماً.
- diversity قد تتفوق على raw accuracy.
- المحفظة المثلى تتغير حسب workload.

Mechanism

1. قياس components منفردة.
2. pair/group ablations لتقدير kappa.
3. فصل synergy من redundancy.
4. optimization تحت budget/risk.
5. إعادة التقدير عند version drift.

Operational Assets 10

ID	Class	Artifact
P9-041	TOOL	Capability Graph Profiler
P9-042	TOOL	Complementarity Estimator
P9-043	TOOL	Portfolio Optimizer
P9-044	TOOL	Redundancy Pruner
P9-045	TOOL	Shared-Failure Risk Meter
P9-046	METHOD	Budgeted Capability Selection
P9-047	METHOD	Conditional Marginal Contribution
P9-048	PROTOCOL	Portfolio Re-Benchmark Protocol
P9-049	PROTOCOL	No-Redundant-Agent-Spawn Protocol
P9-050	STACK	Capability Portfolio Controller

Training / Evaluation Recipes 3

1. Complementarity Ablation Training - تعلم marginal gain من حذف/إضافة components.
2. Substitution-vs-Synergy Contrastive Sets - فصل redundancy عن complementarity.
3. Budgeted Portfolio Search Curriculum - اختيار portfolio تحت cost/latency floors.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لمرونة أفق السياق

Moussaif Context-Horizon Elasticity Theory

الأطروحة

السياق الفعال يتمدد مع long dependencies/unresolved obligations وينكمش عندما يصبح الماضي reconstructible من memory/provenance. الهدف ليس أطول context بل أقل horizon آمن.

الفجوة المستقلة

Long-Horizon Coherence يحفظ state وContext Capsule يضغطه، لكن لا يوجد اختيار ديناميكي لأصغر horizon يحافظ على القرار والثوابت.

H*=argmin_H(ContextCost+ReconstructionCost+LossRisk) s.t. InvariantRetention>=tau

Boundary

يعمل كسياسة خارجية حتى مع models لا تسمح بإدارة context الداخلي.

Decisive experiment

مهام histories من 100k-1M tokens مع dependencies بعيدة.

Objects

ContextHorizon · DependencySpan · ReconstructionPointer · CompressionRisk · ElasticityPolicy

Integration

يوحد Context Capsule وSRM وHalf-Life وCheckpointing وMCD resource economics.

Falsifiers

- لا فرق عن summary+retrieval.
- invariant retention غير قابلة للقياس.
- overhead أكبر من التوفير.

Predictions

- خفض context cost بلا خسارة long-horizon success.
- full context يحمل noise وshort context يفقد obligations.
- H يتغير أثناء المهمة.

Mechanism

1. تقدير dependency span.
2. فصل live context عن reconstructible memory.
3. محاكاة compaction.
4. تمديد horizon عند obligations.
5. تسجيل dropped segments كمؤشرات إعادة بناء.

Operational Assets 10

ID	Class	Artifact
P9-051	TOOL	Dependency Span Estimator
P9-052	TOOL	Context Elasticity Controller
P9-053	TOOL	Reconstruction Pointer Builder
P9-054	TOOL	Compaction Risk Simulator
P9-055	TOOL	Obligation Horizon Tracker
P9-056	METHOD	Invariant-Preserving Context Minimization
P9-057	METHOD	Stage-Adaptive Horizon Learning
P9-058	PROTOCOL	Compaction-before-Drop Protocol
P9-059	PROTOCOL	Horizon Expansion Protocol
P9-060	STACK	Elastic Context Plane

Training / Evaluation Recipes 3

1. Context Horizon Curriculum - histories بأطوال واعتماديات مختلفة.
2. Compaction Counterfactual Training - مقارنة قرار قبل/بعد إسقاط segments.
3. Pending-Obligation Retention Drills - التزام بعيد يجب ألا يضيع بالضغط.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لاستقلال المحقق وارتباط الأخطاء

Moussaif Verifier-Independence and Correlated-Error Theory

الأطروحة
قوة التحقق تعتمد على error independence لا اختلاف الاسم. يجب تقدير correlation تحت perturbations وتخفيض وزن verifier المشترك الجذر.

الفجوة المستقلة
Independent verifier قد لا يكون مستقلاً فعلياً إذا اشترك مع generator في model family/data/retrieval/rubric.

EffectiveVerification=sum(Accuracy_i*Independence_i)-CorrelatedFalsePositivePenalty

Boundary
الاستقلال تقديري ومجالي، لا ادعاء statistical independence مطلق.

Decisive experiment
Generator + 5 verifiers من families مختلفة على shared misconception/adversarial suite.

Objects
VerifierLineage · ErrorSignature · CorrelationMatrix · IndependenceScore · VerifierPortfolio

Integration
يعطي معنى تجريبياً لـIndependent Courts وProcess-Outcome Verification وDET.

Falsifiers
- correlation لا تتنبأ بالفشل المشترك.
- best single verifier مساوٍ دائمًا.
- lineage metadata وحدها كافية.

Predictions
- verifiers مختلفة أسماءً قد تخطئ معاً.
- verifier أضعف لكنه مستقل قد يزيد safety.
- الأثر أكبر في high-impact claims.

Mechanism
1. تسجيل lineage.
2. تشغيل perturbation suite.
3. تقدير error correlation.
4. اختيار portfolio accuracy/diversity/cost.
5. إعادة معايرة بعد updates.

Operational Assets 10

ID	Class	Artifact
P9-061	TOOL	Verifier Lineage Mapper
P9-062	TOOL	Error Correlation Estimator
P9-063	TOOL	Verifier Portfolio Selector
P9-064	TOOL	Shared-Root Failure Detector
P9-065	TOOL	Independence-Weighted Verdict Combiner
P9-066	METHOD	Perturbation-Based Independence Testing
P9-067	METHOD	Verifier Diversity Calibration
P9-068	PROTOCOL	No-Independent-Label-without-Test Protocol
P9-069	PROTOCOL	Correlated-Failure Escalation
P9-070	STACK	Verifier Independence Service

Training / Evaluation Recipes 3

1. Correlated-Error Perturbation Training - verifiers تشترك في misconception واحد.
2. Verifier-Lineage Calibration - ربط lineage بالخطأ المشترك دون افتراض سببية كاملة.
3. Independent-Evidence Preference Training - تفضيل portfolio متنوع على تكرار نفس family.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لبيئة المعيار وانجراف التقييم

Moussaif Benchmark-Ecology and Evaluation-Drift Theory

الأطروحة

Benchmark بيئة لها lifecycle: generation, exposure, adaptation, retirement, renewal. score لا يحمل حق التعميم إذا انجرف benchmark عن العالم الفعلي.

الفجوة المستقلة

Active Failure Discovery يعيد أخذ العينات، لكنه لا يعالج exposure/contamination/staleness والابتعاد عن deployment distribution.

EvalValidity=Coverage*Freshness*Hiddenness*DeploymentSimilarity*AntiGaming-Contamination

Boundary

يفصل public anchors عن hidden renewal pool؛ لا يلزم السرية الكاملة.

Decisive experiment

Benchmark بثلاثة أجيال exposure وdeployment متغير.

Objects

BenchmarkPopulation · ExposureState · DriftVector · ContaminationRisk · RetirementRule

Integration

يمد Benchmark Lab وAFD وHarness Parity وSETT.

Falsifiers

- drift لا يفسر deployment gap.
- hidden variants لا تضيف شيئاً.
- renewal يقتل comparability بلا فائدة.

Predictions

- score يبقى عالياً رغم ضعف deployment عند تجاهل drift.
- procedural variants تكشف memorization.
- validity تنخفض مع exposure.

Mechanism

1. تسجيل exposure.
2. توليد hidden capability variants.
3. قياس deployment drift.
4. تقاعد/تدوير contaminated tasks.
5. anchors ثابتة للمقارنة الزمنية.

Operational Assets 10

ID	Class	Artifact
P9-071	TOOL	Benchmark Exposure Tracker
P9-072	TOOL	Deployment Drift Monitor
P9-073	TOOL	Hidden Variant Generator
P9-074	TOOL	Benchmark Retirement Engine
P9-075	TOOL	Longitudinal Anchor Manager
P9-076	METHOD	Evaluation Validity Scoring
P9-077	METHOD	Drift-Aware Claim Scoping
P9-078	PROTOCOL	Benchmark Renewal Protocol
P9-079	PROTOCOL	No-Generalization-from-Expired-Eval Protocol
P9-080	STACK	Evaluation Ecology Service

Training / Evaluation Recipes 3

1. Benchmark Drift Curriculum - توزيع evaluation يتباعد تدريجياً عن deployment.
2. Contamination Detection Training - tasks محفوظة أو متسربة مقابل hidden variants.
3. Renewal-and-Anchor Training - المحافظة على comparability مع تدوير benchmark.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لميزانية البرهان في التحسين الذاتي العودي

Moussaif Recursive-Improvement Proof-Budget Theory

الأطروحة

كل self-modification يستهلك Proof Budget متناسباً مع blast radius وirreversibility وauthority delta وdistribution shift. التغيير الصغير القابل للرجوع لا يحتاج نفس برهان تعديل kernel واسع.

الفجوة المستقلة

Proof-Carrying Recursive Improvement يحتاج قانوناً يربط حجم evidence المطلوب بمجال أثر mutation.

RequiredProof=BlastRadius*Irreversibility*AuthorityDelta*DistributionShift

Boundary

Proof evidence packet هندسي/تجريبي، لا برهان رياضي عام على AGI safety.

Decisive experiment

50 mutation بدرجات blast مختلفة في reference runtime.

Objects

MutationProposal · BlastRadius · ProofBudget · RegressionEnvelope · RollbackProof

Integration

يحول recursive improvement من vision إلى release science مرتبطة بالEvidence.

Falsifiers

- blast radius لا يتنبأ بالمخاطر.
- البروتوكول بيروقراطي بلا gain.
- static checklist مساوٍ.

Predictions

- خفض large regressions مقابل uniform CI.
- تدفق أسرع للتعديلات الصغيرة.
- authority-expanding mutations تتطلب evidence أعلى.

Mechanism

1. تصنيف mutation.
2. تقدير blast radius.
3. بناء regression suite proportional.
4. canary + rollback proof.
5. توسيع scope تدريجياً مع evidence.

Operational Assets 10

ID	Class	Artifact
P9-081	TOOL	Mutation Blast-Radius Estimator
P9-082	TOOL	Proof Budget Calculator
P9-083	TOOL	Regression Envelope Builder
P9-084	TOOL	Rollback Proof Generator
P9-085	TOOL	Canary Promotion Controller
P9-086	METHOD	Blast-Radius Weighted Assurance
P9-087	METHOD	Recursive Improvement Regression Design
P9-088	PROTOCOL	No-Self-Modification-without-Rollback Protocol
P9-089	PROTOCOL	Progressive Proof Accrual Protocol
P9-090	STACK	Recursive Improvement Gate

Training / Evaluation Recipes 3

1. Blast-Radius Curriculum - mutations من config صغير إلى kernel واسع.
2. Proof-Budget Preference Training - اختيار test envelope مناسب للأثر.
3. Rollback-Proof Drills - failure during canary ثم استرداد version/state.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لتدقيق سطح الهدف الكامن

Moussaif Latent Goal-Surface Audit Theory

الأطروحة

يستدل على Goal Surface وظيفي من choices تحت perturbations. الفرق بين declared utility وbehavioral utility يسمى Goal-Surface Residual؛ إذا استقر وكان مؤثرا يصبح risk حتى دون ادعاء نية داخلية.

الفجوة المستقلة

Goal Drift Immunity تقارن goal والخطة؛ repeated trade-offs قد تكشف effective preference surface لا يظهر في النص أو explanation.

GSR=Distance(DeclaredUtility,InferredBehavioralUtility|perturbations)*Impact

Boundary

تحليل سلوكي وظيفي؛ لا يثبت نية أو وعي.

Decisive experiment

Controlled tasks مع shortcuts/privilege/reward proxies/long-term costs.

Objects

DeclaredGoal · BehavioralTradeoff · GoalSurface · Residual · AuditProbe

Integration

يمتد من Goal-as-OS وGoal Drift Immunity وScheming-Aware Audit.

Falsifiers

- السطح غير مستقر.
- residual لا يتنبأ بسلوك لاحق.
- الأخطاء عشوائية بالكامل.

Predictions

- بعض drift يظهر سلوكياً لا نصياً.
- explanations predictor أضعف من probes.
- قد يكون السبب harness incentives.

Mechanism

1. توليد trade-off probes.
2. تسجيل choices لا explanations.
3. استنتاج minimal utility surface.
4. مقارنته بعقد الهدف.
5. تصعيد residual المستقر.

Operational Assets 10

ID	Class	Artifact
P9-091	TOOL	Goal Tradeoff Probe Generator
P9-092	TOOL	Behavioral Utility Inference Engine
P9-093	TOOL	Goal-Surface Residual Meter
P9-094	TOOL	Harness Incentive Auditor
P9-095	TOOL	Goal Drift Escalation Planner
P9-096	METHOD	Perturbational Goal-Surface Inference
P9-097	METHOD	Declared-Effective Goal Divergence Analysis
P9-098	PROTOCOL	No-Intent-Claim-from-Trace Protocol
P9-099	PROTOCOL	Persistent Goal-Residual Escalation
P9-100	STACK	Goal-Surface Audit Service

Training / Evaluation Recipes 3

1. Behavioral Tradeoff Probe Training - مفاضلات تكشف effective utility.
2. Harness-vs-Model Attribution Curriculum - تغيير incentives مع ثبات model.
3. Goal-Residual Persistence Testing - residual عابر مقابل drift منهجي.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لقبول الجدة في العالم المفتوح

Moussaif Open-World Novelty Admission Theory

الأطروحة
الجدة المقبولة تحتاج عدم قابلية تمثيل حالية + residual ثابت عبر أمثلة مستقلة + تحسن prediction/action بعد minimal extension. قبل ذلك تبقى candidate محجورة.

الفجوة المستقلة
Dimension Promotion وNovelty Engine يكشفان residual، لكن لا توجد quarantine/admission semantics تمنع تضخم ontology من كل anomaly.

NoveltyValue=ResidualPersistence*PredictiveGain*ActionGain*Transfer-ComplexityPenalty

Boundary
جدة أنطولوجية تشغيلية، لا اكتشاف قانون علمي بالضرورة.

Decisive experiment
Open-world API/tool/knowledge tasks مع hidden new categories وnoise anomalies.

Objects
NoveltyCandidate · ResidualCluster · OntologyExtension · AdmissionTest · QuarantineState

Integration
يجمع Open-Dimensional Map وDimension Promotion وNeoAGI novelty وTheory Independence.

Falsifiers
- لا gain بعد extension.
- residual لا يثبت.
- complexity penalty غير منضبط.

Predictions
- خفض ontology inflation.
- بعض anomalies تندمج في types قديمة.
- accepted extensions تنقل benefit.

Mechanism
1. كشف residual.
2. تجميع independent cases.
3. اقتراح minimal type/dimension.
4. ablation with/without extension.
5. admit/merge/quarantine/reject.

Operational Assets 10

ID	Class	Artifact
P9-101	TOOL	Residual Novelty Detector
P9-102	TOOL	Novelty Cluster Validator
P9-103	TOOL	Ontology Extension Generator
P9-104	TOOL	Novelty Admission Court
P9-105	TOOL	Novelty Quarantine Store
P9-106	METHOD	Residual Persistence Testing
P9-107	METHOD	Minimal Ontology Extension Method
P9-108	PROTOCOL	Quarantine-before-Canonicalization Protocol
P9-109	PROTOCOL	Transfer-before-Novelty-Promotion Protocol
P9-110	STACK	Open-World Admission Service

Training / Evaluation Recipes 3

1. Open-World Residual Curriculum - anomalies حقيقية وضوضاء وnew categories.
2. Ontology Minimality Training - extension صغيرة مقابل taxonomy متضخمة.
3. Novelty Transfer Gate Training - لا promotion إلا بعد benefit خارج أمثلة البناء.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect، وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لحدود التقطير السببي للذاكرة

Moussaif Causal Memory Distillation Boundary Theory

الأطروحة
لا تُرقّى تجربة إلى semantic/procedural memory إلا بقدر causal contribution. ما لم يحدد سبباً يبقى episodic مع uncertainty.

الفجوة المستقلة
Self-Reconstruction يحفظ الخبرة، لكن summaries قد تحول عناصر كانت موجودة فقط إلى قواعد سببية مزعومة.

Promotability=CausalContribution*Transfer*EvidenceQuality-NarrativeConfounding

Boundary
لا يتطلب causal proof كامل؛ يسمح partial contribution مع uncertainty.

Decisive experiment
Trajectories مع distractors وredundant tools؛ narrative vs CMDT.

Objects
MemoryCandidate · CausalContribution · DistillationBoundary · EpisodicResidue · PromotionReceipt

Integration
يمد SRM وcausal credit وNegative Knowledge وCrystallization.

Falsifiers
- لا تقل wrong-memory failures.
- contribution لا يتنبأ بالtransfer.
- replay cost غير مبرر.

Predictions
- خفض memory contamination.
- ذاكرة دلالية أصغر وأعلى precision.
- أكبر أثر في long trajectories.

Mechanism
1. استخراج candidate lessons.
2. LOO/counterfactual replay.
3. فصل causal core من correlated context.
4. ترقية core فقط.
5. حفظ residue episodic.

Operational Assets 10

Artifact	Class	ID
Memory Candidate Extractor	TOOL	P9-111
Causal Memory Attribution Engine	TOOL	P9-112
Episodic Residue Separator	TOOL	P9-113
Promotion Receipt Generator	TOOL	P9-114
Memory Rule Regression Tester	TOOL	P9-115
Leave-One-Lesson-Out Distillation	METHOD	P9-116
Causal-Core Compression	METHOD	P9-117
No-Semantic-Promotion-from-Correlation Protocol	PROTOCOL	P9-118
Episodic-First Memory Admission Protocol	PROTOCOL	P9-119
Causal Memory Distiller	STACK	P9-120

Training / Evaluation Recipes 3

1. Lesson Causal Ablation Training - حذف lesson من replay وقياس الأثر.
2. Episodic-to-Semantic Promotion Curriculum - درجات evidence مختلفة.
3. Narrative Confounding Red Team - distractors تظهر في trace دون مساهمة سببية.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لفقد الدلالة عبر البروتوكولات

Moussaif Protocol-Semantic Loss Theory

الأطروحة

لكل protocol translation Semantic Loss Budget. parsing success لا يثبت حفظ المعنى. يجب قياس loss في fields الحاكمة قبل commit.

الفجوة المستقلة

Protocol Sovereignty يفصل transport عن authority، يبقى خطر أن يحفظ adapter syntax ويخسر uncertainty/constraints/ side-effects/error semantics.

PSL=sum(w_k*Loss(field_k))+InvariantLoss+SideEffectMismatch

Boundary

يخص contracts التشغيلية، لا نظرية لغوية عامة.

Decisive experiment

MCP/A2A/custom schemas مع deliberate mapping omissions.

Objects

SemanticContract · ProtocolMapping · LossVector · RoundTripTest · LossBudget

Integration

يكمل Protocol Sovereignty وGovernance Conservation وTool Contracts.

Falsifiers

- لا يتنبأ PSL بالفشل.
- schema validation كافية.
- لا فروق material بين adapters.

Predictions

- أخطر loss في optional/error/authority fields.
- syntax tests لا تكشف side-effect mismatch.
- PSL يفسر provider/tool swap failures.

Mechanism

1. تعريف canonical semantics.
2. mapping لكل adapter.
3. round-trip + adversarial tests.
4. حساب loss vector.
5. approve/restrict/reject adapter.

Operational Assets 10

ID	Class	Artifact
P9-121	TOOL	Canonical Semantic Contract Builder
P9-122	TOOL	Protocol Mapping Analyzer
P9-123	TOOL	Round-Trip Semantic Tester
P9-124	TOOL	Loss Vector Calculator
P9-125	TOOL	Adapter Restriction Planner
P9-126	METHOD	Critical-Field Semantic Equivalence
P9-127	METHOD	Adversarial Protocol Translation Testing
P9-128	PROTOCOL	Semantic-Loss Budget Gate Protocol
P9-129	PROTOCOL	Versioned Adapter Re-Certification Protocol
P9-130	STACK	Protocol Semantics Gateway

Training / Evaluation Recipes 3

1. Protocol Round-Trip Curriculum - A→B→A مع optional/error/authority fields.
2. Semantic-Loss Adversarial Testing - إسقاط side effects أو uncertainty عمداً.
3. Adapter Version-Drift Training - تغيير protocol version ثم recertification.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لتكلفة الفرصة في التنسيق

Moussaif Coordination Opportunity-Cost Theory

الأطروحة

يضاف agent فقط إذا كانت القيمة الهامشية من division/diversity/verification أكبر من بديل استخدام نفس الموارد في single-agent compute أو tool/experiment.

الفجوة المستقلة

Coordination Tax يقيس كلفة multi-agent بعد التنفيذ، لكنه لا يقرر ex ante إن كان الوكيل الثاني أفضل من deeper reasoning أو tool call بنفس الميزانية.

MCV=DivisionGain+DiversityGain+VerificationGain-SyncCost-ConflictCost-LostSingleAgentCompute

Boundary

اقتصاد موارد وتوبع أخطاء، لا ادعاء «ذكاء جماعي» ميتافيزيقي.

Decisive experiment

single-deep vs single+tool vs 2-agent vs 4-agent vs verifier-only تحت budget ثابت.

Objects

CoordinationOption · SingleAgentAlternative · MarginalCoordinationValue · SyncCost · ConflictCost

Integration

يربط Capability Complementarity وM-CAD وtoken economics وRole Councils.

Falsifiers

- MCV لا يتنبأ بالمكسب.
- spawn heuristic بسيط مساوٍ.
- التقدير يستهلك المكاسب.

Predictions

- multi-agent لا يفوز دائماً.
- agent مستقل قد يفيد أكثر كمحقق.
- MCV يتغير حسب المرحلة.

Mechanism

1. تقدير decomposability/error correlation.
2. إنشاء budget-equivalent alternatives.
3. تقدير MCV.
4. spawn فقط إذا موجب.
5. prune/reallocate أثناء التشغيل.

Operational Assets 10

ID	Class	Artifact
P9-131	TOOL	Coordination Option Generator
P9-132	TOOL	Marginal Coordination Value Estimator
P9-133	TOOL	Sync Cost Meter
P9-134	TOOL	Conflict Cost Meter
P9-135	TOOL	Dynamic Agent Pruner
P9-136	METHOD	Budget-Equivalent Coordination Comparison
P9-137	METHOD	Stage-Specific Coordination Valuation
P9-138	PROTOCOL	Spawn-on-Positive-MCV Protocol
P9-139	PROTOCOL	Prune-on-Negative-Marginal-Value Protocol
P9-140	STACK	Coordination Economics Controller

Training / Evaluation Recipes 3

1. Matched-Budget Coordination Curriculum - single vs multi-agent بنفس الموارد.
2. Stage-Specific Spawn Training - planning/execution/verification phases.
3. Negative-Marginal Agent Pruning Drills - إيقاف agent بعد تحول MCV إلى سالب.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

نظرية مساعف لأمانة الانتقال من المحاكاة إلى التدخل

Moussaif Simulation-to-Intervention Fidelity Theory

الأطروحة
لا تستبدل تجربة واقعية بمحاكاة إلا إذا كانت fidelity على causal variables الحاكمة فوق threshold ومعروف اتجاه bias. global resemblance أقل أهمية من causal fidelity.

الفجوة المستقلة
Interventional Compute قد يختار simulation؛ لا يوجد معيار يحدد متى تكفي المحاكاة لاتخاذ قرار واقعي على المتغيرات السببية المهمة.

SIF=sum(Importance_k*Fidelity_k)-UnmodeledCausalMass-BiasUncertainty

Boundary
عند جهل fidelity يجب خفض الاعتماد على simulation لا افتراضها.

Decisive experiment
System/tool sandboxes مع controlled mismatches؛ sim-only vs real-only vs hybrid.

Objects
SimulatorModel · CausalVariableSet · FidelityProfile · BiasDirection · RealityCheck

Integration
يكمل M-ICAT وCounterfactual World-Effect Planning وCausal Sufficiency.

Falsifiers
- SIF لا يتنبأ sim-to-real gap.
- generic validation كافية.
- لا يغير قرار sim/real.

Predictions
- causal-variable fidelity أفضل من global accuracy.
- hybrid أفضل cost/risk trade-off.
- fidelity تتدهور تحت shift.

Mechanism
1. تحديد causal variables.
2. قياس fidelity عبر reality anchors.
3. تعلم bias.
4. اختيار sim/hybrid/real.
5. تحديث simulator بعد discrepancy.

Operational Assets 10

ID	Class	Artifact
P9-141	TOOL	Causal Variable Importance Mapper
P9-142	TOOL	Simulator Fidelity Profiler
P9-143	TOOL	Bias Direction Estimator
P9-144	TOOL	Reality Check Scheduler
P9-145	TOOL	Hybrid Intervention Planner
P9-146	METHOD	Causal-Fidelity Calibration
P9-147	METHOD	Sim-to-Real Uncertainty Propagation
P9-148	PROTOCOL	No-Real-World-Conclusion-from-Low-SIF Protocol
P9-149	PROTOCOL	Reality-Anchor Refresh Protocol
P9-150	STACK	Simulation Fidelity Service

Training / Evaluation Recipes 3

1. Sim-to-Real Paired Curriculum - نفس intervention في simulator وreality anchor.
2. Causal-Variable Fidelity Training - وزن fidelity حسب variables الحاكمة.
3. Distribution-Shift Reality-Check Drills - انخفاض SIF بعد تغير البيئة.

حكم الترقية: لا يكفي نجاح textual output. يلزم observable state/effect؛ وعند الادعاء السببي يلزم matched-budget ablation. إذا لم يتغير outcome أو failure localization أو evidence/authority efficiency، يزول استقلال النظرية.

Implementation Stack IX - 19 layers

L0 · Source & Provenance
hash/source-rank/claim split/dedup/quarantine

L1 · Open-World Admission
novelty quarantine/ontology extension

L2 · Temporal Evidence & Escrow
half-life/debt/escrow/release

L3 · Goal / Identity / Jurisdiction
goal audit/delegation/revocation

L4 · Question / Difference / MCD
governing question and compute value

L5 · Context Horizon
elastic context/reconstruction pointers

L6 · Plural Reasoning / Disagreement
MURI and experiment arbitration

L7 · Causal Experiment & Simulation
ICV/SIF/reality anchors

L8 · Knowledge & Causal Memory
positive/negative causal distillation

L9 · Capability Portfolio
complementarity/coordination/crystallization

L10 · Protocol Semantic Gateway
MCP/A2A/custom semantic-loss budgets

L11 · Trajectory Shadow & Invariants
counterfactual shadow/prefix checks

L12 · Execution / World State
prepared action/idempotency/compensation

L13 · World-Effect Reconciliation
multi-observer readback/finality

L14 · Verifier Independence
lineage/error correlation/process-outcome

L15 · Failure Topology & Recovery
basins/escape/recovery proof

L16 · Recursive Improvement Assurance
blast radius/proof budget/canary

L17 · Benchmark Ecology
exposure/drift/hidden variants/AFD

L18 · Research Portfolio Governance
SETT/EIG/theory freeze/replication

Execution principle: لا يلزم 19 service منفصلة؛ يمكن دمجها في monorepo أو runtime واحد، لكن interfaces والحجج الحاكمية يجب ألا تختلط.

Experimental Matrix IX

ID	Axis	Matched baseline
E9-01	Escrow	hard-stop / optimistic / escrow
E9-02	Shadowing	unit test / shadow
E9-03	Jurisdiction	flat / role / jurisdiction
E9-04	Effect Lattice	single readback / reconciliation
E9-05	Portfolio	additive / complementarity
E9-06	Context	full/fixed/summary/elastic
E9-07	Verifier	best / majority / independence portfolio
E9-08	Benchmark	static / ecology
E9-09	Proof Budget	uniform CI / RIPBT
E9-10	Goal Surface	text audit / behavioral probes
E9-11	Novelty	accept/fixed/ONAT
E9-12	Memory	narrative / causal distillation
E9-13	Protocol	schema / semantic-loss gate
E9-14	Coordination	single / multi matched-budget
E9-15	Simulation	sim / real / hybrid

القواعد الموحدة

- Frozen harness fingerprint: model/version/reasoning effort/tools/context/retry/verifier.
- Matched compute/time/tool-call budgets عند causal claims.
- Held-outs + repeated seeds + confidence intervals حيث يلزم.
- Process + outcome + world-state؛ لا نجاح نهائي بلا readback.
- Security/authority/truthfulness non-regression floors.
- Negative-results ledger إلزامي، ولا حذف لفشل يخفض النظرية.

الخلاصة ومصفوفة الأولوية

الجزء التاسع يضيف نظرية عن القرار المؤجل لا الدين فقط، وعن الهشاشة قبل transition لا حفظ الثوابت بعده، وعن الاختصاص لا الدور، وعن finality الموزعة لا success flag، وعن synergy لا جمع القدرات، وعن الأفق لا طول السياق، وعن استقلال المحقق empirically لا اسما، وعن benchmark lifecycle لا score فقط، وعن proof budget لا مجرد rollback، وعن الهدف الفعال لا النص، وعن admission للجدة لا detection، وعن causal memory لا narrative memory، وعن semantic protocol loss لا syntax، وعن قيمة التنسيق ex ante لا tax بعدي، وعن fidelity causal للمحاكاة لا resemblance.

ResearchValue=IndependentDelta×Testability×CausalIdentifiability×Transfer×Reproducibility-ComplexityCost

تنفيذ بيئي	تنفيذ مبكر
M9-03، M9-10، M9-15، M9-04 تحتاج sandboxes/services/traces طويلة.	M9-09، M9-12، M9-13، M9-07 لأنها قابلة للقياس داخل Reference Kernel بأقل world-risk.

قاعدة الحذف: إذا لم تضف النظرية causal أو operational value مقارنة بأقرب parent، يجب دمجها أو رفضها. لا قيمة علمية في حماية الاسم.