RECONSTRUCTED_FROM_INDEXED_TEXT - نسخة أرشيفية معاد بناؤها من النص المفهرس

<PARSED TEXT FOR PAGE: 1 / 40>
نظرية مساعف لإغلاق القدرة الوكيلية والأثر العالمي 2026 MOUSSAIF UNIFIED AGI RESEARCH PROGRAM / September 05
Moussaif Agentic Capability Closure Theory (M-ACCT) - Dense formal research paper
Evidence-ranked Research Candidate Amiri Arabic RTL
زهير مساعف / Moussaif Zouhair
بنية افتراضية دقيقة قابلة للهدم والقياس. دمج داخلي لمخرجات MOUSSAIXT/HKOS-M/MUG/MUT مع قيود ونتائج المجتمع العلمي، من دون تحويل art prior إلى إثبات للمشروع.
1
<PARSED TEXT FOR PAGE: 2 / 40>
P14-B / M-ACCT / PAGE 002 / 05 September 2026
الادعاء وحدود المركزية المشكلة - Moussaif Agentic Capability Closure Theory
)M-ACCT(
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعلياً بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالمشكلة المركزية وحدود الادعاء.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد. الإغلاق معيار هو World Effect، Agent jurisdictional executor، HyperLoop bounded controller، HyperAbility، HyperSkill primitive، compositional policy
registry capability تخريطة reachability مثبتة. التركيز هنا هو المشكلة المركزية وحدود الادعاء، مع منع الخلط بين schema نظري ونتيجة تجريبية.
CapabilityClosure = Reachability × Reliability × Verifiability × Recoverability × Transferability
المعادلة هنا object research formal وليست رقماً تجريبياً جاهزاً. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
APB 2026
Agent Planning Benchmark: 4,209 multimodal cases across 22 domains; planning, broken tools, unsolvable tasks, refinement.
arXiv:2606.04874
AMA-Bench 2026
Long-horizon agent memory benchmark; highlights missing causality/objective information and lossy similarity retrieval.
arXiv:2602.22769
AgentLAB 2026
Long-horizon attack benchmark covering intent hijacking, tool chaining, task injection, objective drift and memory poisoning.
arXiv:2602.16901
RoadmapBench 2026
Long-horizon software-development benchmark over real version upgrades; exposes multi-file, multi-target planning limits.
arXiv:2605.15846
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفياً محتملاً.
3. الربط الداخلي
MUT: theory unification, lineage, relation graph, research questions and experiment candidates.
MUG: governance of placement, internalization, authority, evolution and evidence status.
MOUSSAIXT: cross-environment execution shell binding models, tools, memory, agents, verification and local/cloud runtimes.

MCD / QMCD / MCAI / HAKIM-MCD: difference-first, question-conditioned,
plural-path and falsification-oriented cognitive routing.
4. التجربة الحاسمة
مع التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد، skill versus ability
تجربة صمم matched-budget حول ablation
حذف المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عاماً.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر capability مستقلة.
6. ممر التطبيق
المركزية المشكلة محور في .pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope: قدرة
إنشاء Capability Closure Certificate لكل
مع ،الادعاء وحدود typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: Candidate theoretical formalization. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•
2
<PARSED TEXT FOR PAGE: 3 / 40>
P14-B / M-ACCT / PAGE 003 / 05 September 2026
MUT/MUG/MOUSSAIXT - Moussaif Agentic Capability Closure إلى النسب
)Theory)M-ACCT
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعلياً بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالنسب إلى MUT/
MUG/MOUSSAIXT.
1. الكائن النظري والدقة الافتراضية
HyperAbility؛ HyperLoop bounded controller؛ Agent jurisdictional executor؛ World Effect هو .الإغلاق معيار تعريف يعاد
HyperSkill primitive؛ compositional policy
schema مع منع الخلط بين ،MOUSSAIXT/MUG/MUT إلى النسب هنا هو التركيز .مثبتة reachability تكريطة registry capability
نظري ونتيجة تجريبية.
ΔState = State_after − State_before ; VerifiedSuccess ⇔ Postcondition(State_after)=true
المعادلة هنا object research formal وليست رقماً تجريبياً جاهزاً. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
RoadmapBench 2026
Long-horizon software-development benchmark over real version upgrades; exposes
multi-file, multi-target planning limits.
arXiv:2605.15846
MM-ToolSandBox 2026
Stateful visual tool-calling across 500+ tools and 16 domains; visual precision
becomes a major bottleneck.
arXiv:2607.11818

World Model Agentic Mid-Training 2026
Capability-first world-model planning; distinguishes predictive capability from merely imitating foresight format.
arXiv:2606.27483
WorldEvolver 2026
Deployment-time world-model revision via episodic memory, semantic memory and selective foresight.
arXiv:2606.30639
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملًا.
3. الربط الداخلي
MCD / QMCD / MCAI / HAKIM-MCD: difference-first, question-conditioned, plural-path and falsification-oriented cognitive routing.
M-HKOS / KB-as-OS: knowledge as governed operational substrate, not passive archive.
HyperSkill -> HyperAbility -> Bounded HyperLoop -> Agent/Tool -> World Effect -< Self-Reconstruction Memory.
M-EAPOS / M-RPROS / M-XEOS / M-CAROS / M-MCBOS: action-state separation, reproducibility, externalities, accountability and multi-currency budgets.
4. التجربة الحاسمة
التجربة أعد ثم ،outcome score عن process score افصل ،raw trajectory سجل ،model/tool/context جمد ؛long-horizon recovery
benchmark حول matched-budget تجربة صمم
مع حذف المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كغضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ.

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر capability مستقلة.
6. ممر التطبيق
pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope. في محور MUT/ إلى النسب
إنشاء Capability Closure Certificate قدرة لكل:
MUG/MOUSSAIXT، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: formalization theoretical Candidate. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•
3
<PARSED TEXT FOR PAGE: 4 / 40>
P14-B / M-ACCT / PAGE 004 / 05 September 2026
قاموس الحالات والكائنات - Moussaif Agentic Capability Closure Theory)M-ACCT(
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعليًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بقاموس الحالات والكائنات.
1. الكائن النظري والدقة الافتراضية

HyperAbility، HyperLoop bounded controller، Agent jurisdictional executor، World Effect تعريف يعاد .الإغلاق معيار هو
HyperSkill primitive، compositional policy
registry capability نخريطة reachability مثبتة. التركيز هنا هو قاموس الحالات والكائنات، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(EvidenceStrength_eff = Σ w_i E_i − CorrelationPenalty(E_1…E_n
المعادلة هنا object research formal وليست رقمً ا تجريبيً ا جاهزًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
WorldEvolver 2026
Deployment-time world-model revision via episodic memory, semantic memory and selective foresight.
arXiv:2606.30639
Causal Memory Intervention 2026
Selects memories by causal usefulness under intervention rather than semantic relevance alone.
arXiv:2605.17641
Continual Learning Moves to Memory 2026
Shows stability-plasticity reappears at memory representation/retrieval level.
arXiv:2604.27003
LWM-Planner 2026
Fact-augmented lookahead planning with validated atomic facts and limited-horizon simulation.
AISTATS 2026 / PMLR 300
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»، بل بوصفها مرجع آلية أو معيار اختبار أو نفيً ا محتملاً .
3. الربط الداخلي
M-EAPOS / M-RPROS / M-XEOS / M-CAROS / M-MCBOS: action-state separation, reproducibility, externalities, accountability and multi-currency budgets.
Evidence discipline: AGI-oriented architecture is not evidence of achieved AGI; E3/E4/E5 require execution, controlled baselines and replication.
MUT: theory unification, lineage, relation graph, research questions and experiment candidates.
MUG: governance of placement, internalization, authority, evolution and evidence status.
4. التجربة الحاسمة
حذف مع التجربة أعد ثم ،outcome score عن process score افصل ،raw trajectory سجل ،model/tool/context جمد ،tool-noise
تجربة صمم matched-budget حول robustness
المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر capability مستقلة.
6. ممر التطبيق
الحالات قاموس محور في .transfer envelope ،failure envelope ،proof bundle ،constraints ،post-state class ،pre-state class
إنشاء Capability Closure Certificate لكل قدرة:
مع ،typed schema وtrace_id وsource/evidence pointers وrollback receipt.والكائنات

حالة هذا البناء: Candidate theoretical formalization. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•

المسلمات الدنيا -)Moussaif Agentic Capability Closure Theory)M-ACCT
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعلي ًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالمسلمات الدنيا.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد. الإغلاق معيار هو World Effect، Agent jurisdictional executor، HyperLoop bounded controller، HyperAbility، HyperSkill primitive، compositional policy
registry capability نخريطة reachability مثبتة. التركيز هنا هو المسلمات الدنيا، مع منع الخلط بين schema نظري ونتيجة تجريبية.
RiskAdjustedUtility = ExpectedVerifiedUtility − λ_risk·Risk − λ_cost·Cost
المعادلة هنا formal object research وليست رقم ًا تجريبي ًا جاهز ًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
LWM-Planner 2026
Fact-augmented lookahead planning with validated atomic facts and limited-horizon simulation.
AISTATS 2026 / PMLR 300
PaperBench 2025
Research replication benchmark: 20 ICML papers and 8,316 gradable outcomes; stresses full research lifecycle competence.
OpenAI 2025
ToolSandbox 2024
Stateful conversational tool-use evaluation with dynamic milestone checks and state dependencies.
arXiv:2408.04682
BrowseComp-Plus 2025
Fixed-corpus deep-research benchmark for retriever/citation/context-engineering disentanglement.
arXiv:2508.06600
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفي ًا محتملاً.
3. الربط الداخلي
MUG: governance of placement, internalization, authority, evolution and evidence status.
MOUSSAIXT: cross-environment execution shell binding models, tools, memory, agents, verification and local/cloud runtimes.
MCD / QMCD / MCAI / HAKIM-MCD: difference-first, question-conditioned, plural-path and falsification-oriented cognitive routing.
M-HKOS / KB-as-OS: knowledge as governed operational substrate, not passive archive.
4. التجربة الحاسمة
حذف مع التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ unsolvable-task refusal حول matched-budget تجربة صمم.
المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى

Task/world-effect success تحت budget مجمد.
ا من القرارٍ.

حيث تكون الثقة جزء error confidence وCalibration
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
عندما يكون الادعاء عامًا. environment/domain/model عبر Transfer
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر capability مستقلة.
6. ممر التطبيق
مع الدنيا المسلمات محور في pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope.
إنشاء Capability Closure Certificate لكل قدرة:
typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: formalization theoretical Candidate. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•
5
<PARSED TEXT FOR PAGE: 6 / 40>
الأولى الرياضية - P14-B / M-ACCT / PAGE 006 / 05 September 2026 -Moussaif Agentic Capability Closure Theory)M
الصياغة
ACCT(
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعليًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالصياغة الرياضية الأولى.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد. الإغلاق معيار هو World Effect؛ Agent jurisdictional executor؛ HyperLoop bounded controller؛ HyperAbility؛ HyperSkill primitive؛ compositional policy
registry capability نكريطة reachability مثبتة. التركيز هنا هو الصياغة الرياضية الأولى، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(Generalization = Conservation(Invariants | model,domain,tool,context change
المعادلة هنا object research formal وليست رقمًا تجريبيًا جاهزًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
BrowseComp-Plus 2025
Fixed-corpus deep-research benchmark for retriever/citation/context-engineering
disentanglement.
arXiv:2508.06600
Agentic Misalignment 2025
Controlled simulations of harmful agentic actions under goal/replacement pressure;
motivates deployment-time governance.
Anthropic 2025
Automated Alignment Auditing 2025
Agents that discover hidden goals, build evaluations and surface concerning behavior.
Anthropic 2025
Bloom Auto-Evals 2025
Reproducible targeted behavioral evaluation generation and severity/frequency
measurement.
Anthropic 2025
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملاً .

3. الربط الداخلي
M-HKOS / KB-as-OS: knowledge as governed operational substrate, not passive archive.
HyperSkill -> HyperAbility -> Bounded HyperLoop -> Agent/Tool -> World Effect -< Self-Reconstruction Memory.
M-EAPOS / M-RPROS / M-XEOS / M-CAROS / M-MCBOS: action-state separation, reproducibility, externalities, accountability and multi-currency budgets.
Evidence discipline: AGI-oriented architecture is not evidence of achieved AGI; E3/E4/E5 require execution, controlled baselines and replication.
4. التجربة الحاسمة
أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد، world-effect postcondition
تجربة صمم matched-budget حول accuracy
التجربة مع حذف المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ.

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر capability مستقلة.
6. ممر التطبيق
الرياضية الصياغة محور في .transfer envelope، failure envelope، proof bundle، constraints، post-state class، pre-state class
إنشاء Capability Closure Certificate لكل قدرة:
الأولى، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: formalization theoretical Candidate. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•
6
<PARSED TEXT FOR PAGE: 7 / 40>
P14-B / M-ACCT / PAGE 007 / 05 September 2026
القياس قابليةو identifiability - Moussaif Agentic Capability Closure Theory
)M-ACCT(
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعليًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بقابلية القياس
.identifiability و
1. الكائن النظري والدقة الافتراضية
تعريف يعاد .الإغلاق معيار هو World Effect، Agent jurisdictional executor، HyperAbility، HyperLoop bounded controller، HyperSkill primitive، compositional policy
registry capability نخريطة reachability مثبتة. التركيز هنا هو قابلية القياس وidentifiability، مع منع الخلط بين schema نظري ونتيجة تجريبية.
CapabilityClosure = Reachability × Reliability × Verifiability × Recoverability × Transferability
المعادلة هنا object research formal وليست رقمًا تجريبيًا جاهزًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
Bloom Auto-Evals 2025

Reproducible targeted behavioral evaluation generation and severity/frequency measurement.
Anthropic 2025
AI Scientist / Nature 2026
End-to-end scientific automation from idea through code, experiment, manuscript and review; shows research workflow can be agentized.
Nature 2026
Belief-Based World Models 2026
Exposes uncertainty about current state alongside simulations under partial observability.
arXiv:2609.00455
TASPO 2026
Action-level credit allocation grounded in verified trajectory outcomes; addresses process-supervision versus outcome-credit mismatch.
arXiv:2608.31077
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملًا.
3. الربط الداخلي
Evidence discipline: AGI-oriented architecture is not evidence of achieved AGI; E3/E4/E5 require execution, controlled baselines and replication.
MUT: theory unification, lineage, relation graph, research questions and experiment candidates.
MUG: governance of placement, internalization, authority, evolution and evidence status.
MOUSSAIXT: cross-environment execution shell binding models, tools, memory, agents, verification and local/cloud runtimes.
4. التجربة الحاسمة
التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ cross-domain
تجربة صمم matched-budget حول capability transfer
كبيانات أساسية لا كضوضاء. استخدم cases failure مع حذف المكوّن المستهدف.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج ولا تعتبر loop/ability مستقلة capability.
6. ممر التطبيق
القياس قابلية محور في .قدرة: pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope
إنشاء Capability Closure Certificate لكل
وidentifiability، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: Candidate theoretical formalization. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•

-)P14-B / M-ACCT / PAGE 008 / 05 September 2026 Moussaif Agentic Capability Closure Theory)M-ACCT
الأنطولوجية الحدود
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعلي ًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالحدود الأنطولوجية.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد .الإغلاق معيار هو World Effect، Agent jurisdictional executor، HyperLoop bounded controller، HyperAbility
HyperSkill primitive، compositional policy
registry capability تخريطة reachability مثبتة. التركيز هنا هو الحدود الأنطولوجية، مع منع الخلط بين schema نظري ونتيجة تجريبية.
ΔState = State_after − State_before ; VerifiedSuccess ⇔ Postcondition(State_after)=true
المعادلة هنا object research formal وليست رقم ًا تجريبي ًا جاهز ًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
TASPO 2026
Action-level credit allocation grounded in verified trajectory outcomes; addresses process-supervision versus outcome-credit mismatch.
arXiv:2608.31077
APB 2026
Agent Planning Benchmark: 4,209 multimodal cases across 22 domains; planning, broken tools, unsolvable tasks, refinement.
arXiv:2606.04874
AMA-Bench 2026
Long-horizon agent memory benchmark; highlights missing causality/objective information and lossy similarity retrieval.
arXiv:2602.22769
AgentLAB 2026
Long-horizon attack benchmark covering intent hijacking, tool chaining, task injection, objective drift and memory poisoning.
arXiv:2602.16901
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفي ًا محتملا .
3. الربط الداخلي
MOUSSAIXT: cross-environment execution shell binding models, tools, memory, agents, verification and local/cloud runtimes.
MCD / QMCD / MCAI / HAKIM-MCD: difference-first, question-conditioned, plural-path and falsification-oriented cognitive routing.
M-HKOS / KB-as-OS: knowledge as governed operational substrate, not passive archive.
HyperSkill -> HyperAbility -> Bounded HyperLoop -> Agent/Tool -> World Effect -< Self-Reconstruction Memory.
4. التجربة الحاسمة
مع التجربة أعد ثم ،outcome score عن process score افصل ،raw trajectory سجل ،model/tool/context جمد ،skill versus ability
تجربة صمم matched-budget حول ablation
حذف المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم

إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تَ دمج loop/ability ولا تعتبر مستقلة. capability
6. ممر التطبيق
الحدود محور في .transfer envelope، failure envelope، proof bundle، constraints، post-state class، pre-state class: قدرة لكل إنشاء Capability Closure Certificate
مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.، الأنطولوجية
حالة هذا البناء: formalization theoretical Candidate. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•
8
<PARSED TEXT FOR PAGE: 9 / 40>
P14-B / M-ACCT / PAGE 009 / 05 September 2026
MCD - Moussaif Agentic Capability Closure Theory)M-ACCT(بـ العلاقة
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعلي ًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـMCD.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد .الإغلاق معيار هو World Effect، Agent jurisdictional executor، HyperLoop bounded controller، HyperAbility HyperSkill primitive، compositional policy
registry capability تخريطة reachability مثبتة. التركيز هنا هو العلاقة بـMCD، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(EvidenceStrength_eff = Σ w_i E_i − CorrelationPenalty(E_1…E_n
المعادلة هنا object research formal وليست رقم ًا تجريبي ًا جاهز ًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
AgentLAB 2026
Long-horizon attack benchmark covering intent hijacking, tool chaining, task injection, objective drift and memory poisoning.
arXiv:2602.16901
RoadmapBench 2026
Long-horizon software-development benchmark over real version upgrades; exposes multi-file, multi-target planning limits.
arXiv:2605.15846
MM-ToolSandBox 2026
Stateful visual tool-calling across 500+ tools and 16 domains; visual precision becomes a major bottleneck.
arXiv:2607.11818
World Model Agentic Mid-Training 2026
Capability-first world-model planning; distinguishes predictive capability from merely imitating foresight format.
arXiv:2606.27483
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُ دمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفي ًا محتملا .
3. الربط الداخلي
HyperSkill -> HyperAbility -> Bounded HyperLoop -> Agent/Tool -> World Effect -< Self-Reconstruction Memory.
M-EAPOS / M-RPROS / M-XEOS / M-CAROS / M-MCBOS: action-state separation, reproducibility, externalities, accountability and multi-currency budgets.
Evidence discipline: AGI-oriented architecture is not evidence of achieved AGI; E3/E4/E5 require execution, controlled baselines and replication.
MUT: theory unification, lineage, relation graph, research questions and experiment

candidates.
4. التجربة الحاسمة
التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ long-horizon recovery
benchmark حول matched-budget تجربة صمم
مع حذف المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج ولا تعتبر loop/ability مستقلة capability.
6. ممر التطبيق
MCD، بـ العلاقة محور في pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope.
إنشاء Capability Closure Certificate لكل قدرة:
مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: formalization theoretical Candidate. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•
9
<PARSED TEXT FOR PAGE: 10 / 40>
P14-B / M-ACCT / PAGE 010 / 05 September 2026
QMCD - Moussaif Agentic Capability Closure Theory)M-ACCT (بـ العلاقة
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعليًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـQMCD.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد. الإغلاق معيار هو World Effect؛ Agent jurisdictional executor؛ HyperLoop bounded controller؛ HyperAbility؛ HyperSkill primitive؛ compositional policy
registry capability تكريطة reachability مثبتة. التركيز هنا هو العلاقة بـQMCD، مع منع الخلط بين schema نظري ونتيجة تجريبية.
RiskAdjustedUtility = ExpectedVerifiedUtility − λ_risk·Risk − λ_cost·Cost
المعادلة هنا formal research object وليست رقمًا تجريبيًا جاهزًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
World Model Agentic Mid-Training 2026
Capability-first world-model planning; distinguishes predictive capability from merely imitating foresight format.
arXiv:2606.27483
WorldEvolver 2026
Deployment-time world-model revision via episodic memory, semantic memory and selective foresight.
arXiv:2606.30639
Causal Memory Intervention 2026
Selects memories by causal usefulness under intervention rather than semantic relevance alone.
arXiv:2605.17641

Continual Learning Moves to Memory 2026
Shows stability-plasticity reappears at memory representation/retrieval level.
arXiv:2604.27003
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملًا.
3. الربط الداخلي
MUT: theory unification, lineage, relation graph, research questions and experiment candidates.
MUG: governance of placement, internalization, authority, evolution and evidence status.
MOUSSAIXT: cross-environment execution shell binding models, tools, memory, agents, verification and local/cloud runtimes.
MCD / QMCD / MCAI / HAKIM-MCD: difference-first, question-conditioned, plural-path and falsification-oriented cognitive routing.
4. التجربة الحاسمة
حذف مع التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ tool-noise
تجربة صمم matched-budget حول robustness
المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر capability مستقلة.
6. ممر التطبيق
QMCD، بـ العلاقة محور في pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope.
إنشاء Capability Closure Certificate لكل قدرة:
مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: formalization theoretical Candidate. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•
10
<PARSED TEXT FOR PAGE: 11 / 40>
P14-B / M-ACCT / PAGE 011 / 05 September 2026
MCAI - Moussaif Agentic Capability Closure Theory)M-ACCT(بـ العلاقة
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعليًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـMCAL.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد .الإغلاق معيار هو World Effect ؛Agent jurisdictional executor ؛HyperLoop bounded controller ؛HyperAbility
HyperSkill primitive؛ compositional policy
registry capability تكريطة reachability مثبتة. التركيز هنا هو العلاقة بـMCAL، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(Generalization = Conservation(Invariants | model,domain,tool,context change
المعادلة هنا object research formal وليست رقمًا تجريبيًا جاهزًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.

2. ضغط المجتمع العلمي
Continual Learning Moves to Memory 2026
Shows stability-plasticity reappears at memory representation/retrieval level.
arXiv:2604.27003
LWM-Planner 2026
Fact-augmented lookahead planning with validated atomic facts and limited-horizon simulation.
AISTATS 2026 / PMLR 300
PaperBench 2025
Research replication benchmark: 20 ICML papers and 8,316 gradable outcomes; stresses full research lifecycle competence.
OpenAI 2025
ToolSandbox 2024
Stateful conversational tool-use evaluation with dynamic milestone checks and state dependencies.
arXiv:2408.04682
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفياً محتملاً.
3. الربط الداخلي
MCD / QMCD / MCAI / HAKIM-MCD: difference-first, question-conditioned, plural-path and falsification-oriented cognitive routing.
M-HKOS / KB-as-OS: knowledge as governed operational substrate, not passive archive.
HyperSkill -> HyperAbility -> Bounded HyperLoop -> Agent/Tool -> World Effect -< Self-Reconstruction Memory.
M-EAPOS / M-RPROS / M-XEOS / M-CAROS / M-MCBOS: action-state separation, reproducibility, externalities, accountability and multi-currency budgets.
4. التجربة الحاسمة
حذف مع التجربة أعد ثم ،outcome score عن process score افصل ،raw trajectory سجل ،model/tool/context جمد ،unsolvable-
تجربة صمم matched-budget حول task refusal
المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرار.

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عاماً.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر capability مستقلة.
6. ممر التطبيق
MCAI، بـ العلاقة محور في .transfer envelope ،failure envelope ،constraints ،post-state class ،pre-state class
إنشاء Capability Closure Certificate لكل قدرة:
مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: Candidate theoretical formalization. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•

<PARSED TEXT FOR PAGE: 12 / 40>
P14-B / M-ACCT / PAGE 012 / 05 September 2026
العلاقة بـHAKIM-MCD - Moussaif Agentic Capability Closure Theory)M-ACCT(
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعلياً بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـHAKIM-MCD.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد. الإغلاق معيار هو World Effect، Agent jurisdictional executor، HyperLoop bounded controller، HyperAbility، HyperSkill primitive، compositional policy
registry capability كخريطة reachability مثبتة. التركيز هنا هو العلاقة بـMCD-HAKIM، مع منع الخلط بين schema نظري ونتيجة تجريبية.
CapabilityClosure = Reachability × Reliability × Verifiability × Recoverability × Transferability
المعادلة هنا formal research object وليست رقماً تجريبياً جاهزاً. كل متغير يجب أن يحصل على operational definition، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
ToolSandbox 2024
Stateful conversational tool-use evaluation with dynamic milestone checks and state dependencies.
arXiv:2408.04682
BrowseComp-Plus 2025
Fixed-corpus deep-research benchmark for retriever/citation/context-engineering disentanglement.
arXiv:2508.06600
Agentic Misalignment 2025
Controlled simulations of harmful agentic actions under goal/replacement pressure; motivates deployment-time governance.
Anthropic 2025
Automated Alignment Auditing 2025
Agents that discover hidden goals, build evaluations and surface concerning behavior.
Anthropic 2025
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»، بل بوصفها مرجع آلية أو معيار اختبار أو نفياً محتملاً.
3. الربط الداخلي
M-EAPOS / M-RPROS / M-XEOS / M-CAROS / M-MCBOS: action-state separation, reproducibility, externalities, accountability and multi-currency budgets.
Evidence discipline: AGI-oriented architecture is not evidence of achieved AGI; E3/E4/E5 require execution, controlled baselines and replication.
MUT: theory unification, lineage, relation graph, research questions and experiment candidates.
MUG: governance of placement, internalization, authority, evolution and evidence status.
4. التجربة الحاسمة
أعد ثم outcome score عن process score، افصل raw trajectory، سجل model/tool/context، جمد world-effect postcondition؛ تجربة صمم matched-budget حول accuracy
التجربة مع حذف المكوّن المستهدف. استخدم failure cases كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.

عندما يكون الادعاء عامًا. environment/domain/model عبر Transfer
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر مستقلة. capability
6. ممر التطبيق
HAKIM-MCD، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope. في العلاقة محور بـ
إنشاء Capability Closure Certificate قدرة لكل:
حالة هذا البناء: formalization theoretical Candidate. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•

M-ACCT (بـ العلاقة M-HKOS - Moussaif Agentic Capability Closure Theory)
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعليًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـHKOS-M.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد. الإغلاق معيار هو World Effect، Agent jurisdictional executor، HyperLoop bounded controller، HyperAbility، HyperSkill primitive، compositional policy
registry capability نخريطة reachability مثبتة. التركيز هنا هو العلاقة بـHKOS-M، مع منع الخلط بين schema نظري ونتيجة تجريبية.
ΔState = State_after − State_before ; VerifiedSuccess ⇔ Postcondition(State_after)=true
المعادلة هنا formal object research وليست رقمًا تجريبيًا جاهزًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
Automated Alignment Auditing 2025
Agents that discover hidden goals, build evaluations and surface concerning behavior.
Anthropic 2025
Bloom Auto-Evals 2025
Reproducible targeted behavioral evaluation generation and severity/frequency measurement.
Anthropic 2025
AI Scientist / Nature 2026
End-to-end scientific automation from idea through code, experiment, manuscript and review; shows research workflow can be agentized.
Nature 2026
Belief-Based World Models 2026
Exposes uncertainty about current state alongside simulations under partial observability.
arXiv:2609.00455
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملاً.
3. الربط الداخلي
MUG: governance of placement, internalization, authority, evolution and evidence status.
MOUSSAIXT: cross-environment execution shell binding models, tools, memory, agents, verification and local/cloud runtimes.
MCD / QMCD / MCAI / HAKIM-MCD: difference-first, question-conditioned, plural-path and falsification-oriented cognitive routing.

M-HKOS / KB-as-OS: knowledge as governed operational substrate, not passive
archive.
4. التجربة الحاسمة
cross-domain؛ جمد model/tool/context، سجل raw trajectory، افصل process score عن outcome score، التجربة أعد ثم
تجربة صمم matched-budget حول capability transfer
مع حذف المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

حيث تكون الثقة جزء error confidenceو Calibration
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
عندما يكون الادعاء عاماً. environment/domain/model عبر Transfer
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر
مستقلة. capability
6. ممر التطبيق
M-بـ العلاقة محور في .قدرة: pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope
إنشاء Capability Closure Certificate لكل
HKOS، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
.replication دون 5E لا ؛baseline frozen دون 4E لا ؛claim AGI لا .formalization theoretical Candidate :حالة هذا البناء
•
•
•
•
•
13
<PARSED TEXT FOR PAGE: 14 / 40>
P14-B / M-ACCT / PAGE 014 / 05 September 2026
M-بـ العلاقة HyperSkills - Moussaif Agentic Capability Closure Theory)
ACCT(
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعلياً بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة
يختبر التركيز الخاص بالعلاقة بـHyperSkills.
1. الكائن النظري والدقة الافتراضية
HyperAbility، HyperLoop bounded controller، Agent jurisdictional executor، World Effect هو معيار الإغلاق. تعريف يعاد
HyperSkill primitive، compositional policy
نظري schema مع منع الخلط بين ،HyperSkills بـ العلاقة هنا هو التركيز .مثبتة reachability تخريطة registry capability ونتيجة تجريبية.
(EvidenceStrength_eff = Σ w_i E_i − CorrelationPenalty(E_1…E_n
المعادلة هنا object research formal وليست رقماً تجريبياً جاهزاً. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو
scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
Belief-Based World Models 2026
Exposes uncertainty about current state alongside simulations under partial
observability.
arXiv:2609.00455
TASPO 2026
Action-level credit allocation grounded in verified trajectory outcomes; addresses
process-supervision versus outcome-credit mismatch.
arXiv:2608.31077
APB 2026
Agent Planning Benchmark: 4,209 multimodal cases across 22 domains; planning,

broken tools, unsolvable tasks, refinement.
arXiv:2606.04874
AMA-Bench 2026
Long-horizon agent memory benchmark; highlights missing causality/objective
information and lossy similarity retrieval.
arXiv:2602.22769
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملًا.
3. الربط الداخلي
M-HKOS / KB-as-OS: knowledge as governed operational substrate, not passive
archive.
HyperSkill -> HyperAbility -> Bounded HyperLoop -> Agent/Tool -> World Effect -<
Self-Reconstruction Memory.
M-EAPOS / M-RPROS / M-XEOS / M-CAROS / M-MCBOS: action-state
separation, reproducibility, externalities, accountability and multi-currency budgets.
Evidence discipline: AGI-oriented architecture is not evidence of achieved AGI; E3/
E4/E5 require execution, controlled baselines and replication.
4. التجربة الحاسمة
مع التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ skill versus ability
تجربة صمم matched-budget حول ablation
حذف المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج ولا تعتبر loop/ability مستقلة capability.
6. ممر التطبيق
العلاقة محور في .قدرة لكل: pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope
إنشاء Capability Closure Certificate
بـHyperSkills، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: theoretical Candidate formalization. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•

العلاقة بـHyperAbilities - Moussaif Agentic Capability Closure Theory)M-
ACCT(
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعليًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة
يختبر التركيز الخاص بالعلاقة
.HyperAbilities بـ
1. الكائن النظري والدقة الافتراضية
تعريف يعاد .الإغلاق معيار هو World Effect؛ Agent jurisdictional executor؛ HyperLoop bounded controller؛ HyperAbility
HyperSkill primitive؛ compositional policy

نظري ونتيجة تجريبية. schema مع منع الخلط بين HyperAbilities، مثبتة. التركيز هنا هو العلاقة بـ reachability تكريطة registry capability
RiskAdjustedUtility = ExpectedVerifiedUtility − λ_risk·Risk − λ_cost·Cost
المعادلة هنا formal research object وليست رقمًا تجريبيًا جاهزًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
AMA-Bench 2026
Long-horizon agent memory benchmark; highlights missing causality/objective information and lossy similarity retrieval.
arXiv:2602.22769
AgentLAB 2026
Long-horizon attack benchmark covering intent hijacking, tool chaining, task injection, objective drift and memory poisoning.
arXiv:2602.16901
RoadmapBench 2026
Long-horizon software-development benchmark over real version upgrades; exposes multi-file, multi-target planning limits.
arXiv:2605.15846
MM-ToolSandBox 2026
Stateful visual tool-calling across 500+ tools and 16 domains; visual precision becomes a major bottleneck.
arXiv:2607.11818
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملًا.
3. الربط الداخلي
Evidence discipline: AGI-oriented architecture is not evidence of achieved AGI; E3/E4/E5 require execution, controlled baselines and replication.
MUT: theory unification, lineage, relation graph, research questions and experiment candidates.
MUG: governance of placement, internalization, authority, evolution and evidence status.
MOUSSAIXT: cross-environment execution shell binding models, tools, memory, agents, verification and local/cloud runtimes.
4. التجربة الحاسمة
التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ long-horizon recovery benchmark حول matched-budget تجربة صمم
مع حذف المكوّن المستهدف. استخدم failure cases كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج ولا تعتبر loop/ability مستقلة capability.
6. ممر التطبيق
العلاقة محور في. pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope: قدرة لكل
إنشاء Capability Closure Certificate
بـHyperAbilities، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: Candidate theoretical formalization. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.

.
.
.
.
.
15
<PARSED TEXT FOR PAGE: 16 / 40>
P14-B / M-ACCT / PAGE 016 / 05 September 2026
Bounded HyperLoops - Moussaif Agentic Capability Closure بـالعلاقة
)Theory)M-ACCT
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعلي ًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بـالعلاقة بـBounded
HyperLoops.
1. الكائن النظري والدقة الافتراضية
HyperAbility، HyperLoop bounded controller، Agent jurisdictional executor، World Effect هو معيار الإغلاق. تعريف يعاد
HyperSkill primitive، compositional policy
registry capability نكريطة reachability مثبتة. التركيز هنا هو العلاقة بـHyperLoops Bounded، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(Generalization = Conservation(Invariants | model,domain,tool,context change
المعادلة هنا object research formal وليست رقم ًا تجريبي ًا جاهز ًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
MM-ToolSandBox 2026
Stateful visual tool-calling across 500+ tools and 16 domains; visual precision becomes a major bottleneck.
arXiv:2607.11818
World Model Agentic Mid-Training 2026
Capability-first world-model planning; distinguishes predictive capability from merely imitating foresight format.
arXiv:2606.27483
WorldEvolver 2026
Deployment-time world-model revision via episodic memory, semantic memory and selective foresight.
arXiv:2606.30639
Causal Memory Intervention 2026
Selects memories by causal usefulness under intervention rather than semantic relevance alone.
arXiv:2605.17641
الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفي ًا محتملا .
3. الربط الداخلي
MOUSSAIXT: cross-environment execution shell binding models, tools, memory, agents, verification and local/cloud runtimes.
MCD / QMCD / MCAI / HAKIM-MCD: difference-first, question-conditioned, plural-path and falsification-oriented cognitive routing.
M-HKOS / KB-as-OS: knowledge as governed operational substrate, not passive archive.
HyperSkill -> HyperAbility -> Bounded HyperLoop -> Agent/Tool -> World Effect -< Self-Reconstruction Memory.
4. التجربة الحاسمة
حذف مع التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد، tool-noise
تجربة صمم matched-budget حول robustness

المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرارٍ.

Calibration وerror confidence حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج ولا تعتبر loop/ability مستقلة capability.
6. ممر التطبيق
Bounded بـ العلاقة محور في pre-state class، post-state class، constraints، proof bundle، failure envelope، transfer envelope.
إنشاء Capability Closure Certificate قدرة لكل:
HyperLoops، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
حالة هذا البناء: formalization theoretical Candidate. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•
16
<PARSED TEXT FOR PAGE: 17 / 40>
P14-B / M-ACCT / PAGE 017 / 05 September 2026
Agent Mesh - Moussaif Agentic Capability Closure Theory)M-ACCT(بـ العلاقة
الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعليًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بـالعلاقة بـMesh Agent.
1. الكائن النظري والدقة الافتراضية
تعريف يعاد. الإغلاق معيار هو World Effect، HyperLoop bounded controller، Agent jurisdictional executor، HyperAbility، HyperSkill primitive، compositional policy
registry capability نخريطة reachability مثبتة. التركيز هنا هو العلاقة بـMesh Agent، مع منع الخلط بين schema نظري ونتيجة تجريبية.
CapabilityClosure = Reachability × Reliability × Verifiability × Recoverability × Transferability
المعادلة هنا formal object research وليست رقمًا تجريبيًا جاهزًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.
الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.
2. ضغط المجتمع العلمي
Causal Memory Intervention 2026
Selects memories by causal usefulness under intervention rather than semantic relevance alone.
arXiv:2605.17641
Continual Learning Moves to Memory 2026
Shows stability-plasticity reappears at memory representation/retrieval level.
arXiv:2604.27003
LWM-Planner 2026
Fact-augmented lookahead planning with validated atomic facts and limited-horizon simulation.
AISTATS 2026 / PMLR 300
PaperBench 2025
Research replication benchmark: 20 ICML papers and 8,316 gradable outcomes; stresses full research lifecycle competence.
OpenAI 2025

الاستفادة العلمية تتم عبر تحويل نتيجة المجتمع إلى constraint أو baseline أو mode failure. لا تَدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملًا.

3. الربط الداخلي

HyperSkill -> HyperAbility -> Bounded HyperLoop -> Agent/Tool -> World Effect -< Self-Reconstruction Memory.

M-EAPOS / M-RPROS / M-XEOS / M-CAROS / M-MCBOS: action-state separation, reproducibility, externalities, accountability and multi-currency budgets.

Evidence discipline: AGI-oriented architecture is not evidence of achieved AGI; E3/E4/E5 require execution, controlled baselines and replication.

MUT: theory unification, lineage, relation graph, research questions and experiment candidates.

4. التجربة الحاسمة

حذف مع التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ unsolvable-task refusal حول matched-budget تجربة صمم

المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كغضوضاء.

مقاييس الحد الأدنى

Task/world-effect success تحت budget مجمد.

ا من القرارٍ.

Calibration وerror confidence حيث تكون الثقة جزء

Recovery / rollback / failure localization.

Cost, latency, tool calls, context load.

Transfer عبر environment/domain/model عندما يكون الادعاء عامًا.

5. شرط الهدم

إذا كان skill منفرد أو baseline planner يحقق نفس verification/recovery/reachability بموارد أقل، تُدمج loop/ability ولا تعتبر capability مستقلة.

6. ممر التطبيق

Agent بـ العلاقة محور في. transfer envelope، failure envelope، proof bundle، constraints، post-state class، pre-state class:

إنشاء Capability Closure Certificate لكل قدرة:

Mesh، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.

حالة هذا البناء: Candidate formalization theoretical. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.

•

•

•

•

•

M-EAPOS - Moussaif Agentic Capability Closure Theory)M-ACCT(بـ العلاقة

الأطروحة: تعريف القدرة بما يمكن للنظام نقله فعليًا بين أصناف الحالات تحت قيود، لا بعدد skills أو loops أو أسماء abilities. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـ EAPOS-M.

1. الكائن النظري والدقة الافتراضية

تعريف يعاد. الإغلاق معيار هو World Effect، Agent jurisdictional executor، HyperLoop bounded controller، HyperAbility، HyperSkill primitive، compositional policy

registry capability نكريطة reachability مثبتة. التركيز هنا هو العلاقة بـ EAPOS-M، مع منع الخلط بين schema نظري ونتيجة تجريبية.

ΔState = State_after − State_before ; VerifiedSuccess ⇔ Postcondition(State_after)=true

المعادلة هنا object research formal وليست رقمًا تجريبيًا جاهزًا. كل متغير يجب أن يحصل على definition operational، نطاق، وحدة أو scale، وطريقة قياس قبل استعماله في قرار runtime.

الغاية هي منع انتقال المشروع من غنى اصطلاحي إلى دقة زائفة.

2. ضغط المجتمع العلمي

PaperBench 2025

Research replication benchmark: 20 ICML papers and 8,316 gradable outcomes; stresses full research lifecycle competence.
OpenAI 2025
ToolSandbox 2024
Stateful conversational tool-use evaluation with dynamic milestone checks and state dependencies.
arXiv:2408.04682
BrowseComp-Plus 2025
Fixed-corpus deep-research benchmark for retriever/citation/context-engineering disentanglement.
arXiv:2508.06600
Agentic Misalignment 2025