نسخة أرشيفية معاد بناؤها من النص المفهرس - RECONSTRUCTED_FROM_INDEXED_TEXT

<PARSED TEXT FOR PAGE: 1 / 40>
MOUSSAIF UNIFIED AGI RESEARCH PROGRAM / 05 September 2026
نظرية هندسة السؤال-الفرق والاقتصاد المعرفي متعدد
العملات
Moussaif Question-Difference Geometry & Resource Rationality (M-QDGR) - 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-C / M-QDGR / PAGE 002 / 05 September 2026
المشكلة المركزية وحدود الادعاء - Moussaif Question-Difference Geometry &
(Resource Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد
متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالمشكلة المركزية وحدود الادعاء.
1. الكائن النظري والدقة الافتراضية
region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors
Q يحدد decision surface؛ Δ يحدد of leverage
currencies substitutable-non. التركيز هنا هو المشكلة المركزية وحدود الادعاء، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(Route* = argmax_r E[VerifiedUtility(r|Q,Δ)] subject to B=(tokens,time,cost,risk,privacy,authority,verification
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، question
تجربة صمم matched-budget حول intervention study
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 لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوي heuristic بسيط نتائج
QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل، branch، DifferenceField، QuestionGraph
تعريف entropy، information gain، shadow prices
محور المشكلة المركزية وحدود الادعاء، مع schema typed وid_trace وpointers evidence/source 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-C / M-QDGR / PAGE 003 / 05 September 2026
MUT/MUG/MOUSSAIXT - Moussaif Question-Difference إلى النسب
(Geometry & Resource Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط
السؤال الصحيح فضاء الفروع ويوزع موارد
متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالنسب إلى MOUSSAIXT/MUG/MUT.
1. الكائن النظري والدقة الافتراضية
region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors
Q يحدد decision surface؛ Δ يحدد of leverage
وnon-substitutable currencies. التركيز هنا هو النسب إلى MOUSSAIXT/MUG/MUT، مع منع الخلط بين schema نظري ونتيجة
تجريبية.
ΔState = State_after − State_before ; VerifiedSuccess ⇔ Postcondition(State_after)=true
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
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، branch contraction
measurement 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 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 عامًا.
5. شرط الهدم
إذا لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل QuestionGraph، DifferenceField، branch، تعريف entropy، information gain، shadow prices
rollback receipt. وsource/evidence pointers وtrace_id مع typed schema، MUT/MUG/MOUSSAIXT إلى النسب محور
حالة هذا البناء: Candidate theoretical formalization. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•

Moussaif Question-Difference Geometry & - والكائنات الحالات قاموس
(Resource Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد

متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بقاموس الحالات والكائنات.
1. الكائن النظري والدقة الافتراضية
region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors
Q يحدد decision surface؛ Δ يحدد of leverage
currencies substitutable-non. 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 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
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 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
مع التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ MCD versus
تجربة صمم matched-budget حول heuristic routing
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 لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph؛

تعريف entropy، information gain، shadow prices
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.
•
•
•
•
•
4
<PARSED TEXT FOR PAGE: 5 / 40>
P14-C / M-QDGR / PAGE 005 / 05 September 2026
الدنيا المسلمات - Moussaif Question-Difference Geometry & Resource
(Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالمسلمات الدنيا.
1. الكائن النظري والدقة الافتراضية
يفرض M-MCBOS hard floors of leverage، Δ يحدد decision surface، Q يحدد region، MCD scores local differences، QMCD conditions scores on governing question.
التركيز هنا هو المسلمات الدنيا، مع منع الخلط بين schema نظري ونتيجة تجريبية. وcurrencies substitutable-non.
RiskAdjustedUtility = ExpectedVerifiedUtility − λ_risk·Risk − λ_cost·Cost
المعادلة هنا formal research object 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
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 نفياً 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

budget shadow-price، جمد model/tool/context، سجل raw trajectory، افصل process score عن outcome score، التجربة أعد ثم
تجربة صمم matched-budget حول calibration
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 لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوي heuristic بسيط نتائج
QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph
تعريف entropy، information gain، shadow prices
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.
•
•
•
•
•
5
<PARSED TEXT FOR PAGE: 6 / 40>
الأولى - P14-C / M-QDGR / PAGE 006 / 05 September 2026 Moussaif Question-Difference Geometry & Resource
الرياضية الصياغة
(Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط
السؤال الصحيح فضاء الفروع ويوزع موارد
متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالصياغة الرياضية الأولى.
1. الكائن النظري والدقة الافتراضية
region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors
Q يحدد decision surface، Δ يحدد of leverage
وcurrencies substitutable-non. التركيز هنا هو الصياغة الرياضية الأولى، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(Generalization = Conservation(Invariants | model,domain,tool,context change
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
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 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 ،retrieval/tool تجربة صمم matched-budget حول avoidance benefit
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 لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوي heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل ،branch ،DifferenceField ،QuestionGraph تعريف entropy، information gain، shadow prices
مع ،الأولى الرياضية الصياغة محور 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.
•
•
•
•
•
6
<PARSED TEXT FOR PAGE: 7 / 40>
P14-C / M-QDGR / PAGE 007 / 05 September 2026
القياس قابلية وidentifiability - Moussaif Question-Difference Geometry & (Resource Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بقابلية القياس وidentifiability.
1. الكائن النظري والدقة الافتراضية
region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors Q يحدد decision surface؛ Δ يحدد of leverage
وnon-substitutable currencies. التركيز هنا هو قابلية القياس وidentifiability، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(Route* = argmax_r E[VerifiedUtility(r|Q,Δ)] subject to B=(tokens,time,cost,risk,privacy,authority,verification
المعادلة هنا formal research object وليست رقمًا تجريبيًا جاهزًا. كل متغير يجب أن يحصل على operational definition، نطاق، وحدة أو

.runtime، وطريقة قياس قبل استعماله في قرار scale
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
أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ question quality transfer
تجربة صمم matched-budget حول across domains
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 لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.
.6 ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph
تعريف entropy، information gain، shadow prices
القياس قابلية محور 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.
•
•
•

.
.
7
<PARSED TEXT FOR PAGE: 8 / 40>
- P14-C / M-QDGR / PAGE 008 / 05 September 2026 Moussaif Question-Difference Geometry & Resource
الأنطولوجية الحدود
(Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد
متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالحدود الأنطولوجية.
1. الكائن النظري والدقة الافتراضية
region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors of leverage يحدد Δ ،decision surface يحدد Q
currencies substitutable-non. 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 formal research object 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
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 نفياً محتملاً.
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؛ question
تجربة صمم matched-budget حول intervention study
حذف المكوّن المستهدف. استخدم cases failure كبيانات أساسية لا كضوضاء.
مقاييس الحد الأدنى
Task/world-effect success تحت budget مجمد.
ا من القرار.

ً

حيث تكون الثقة جزء error confidenceو Calibration
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
عندما يكون الادعاء عامًا. environment/domain/model عبر Transfer
5. شرط الهدم
إذا لم يغير question-governing graph search intervention أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph، تعريف entropy، information gain، shadow prices
مع الأنطولوجية الحدود محور، 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.
•
•
•
•
•
8
<PARSED TEXT FOR PAGE: 9 / 40>
P14-C / M-QDGR / PAGE 009 / 05 September 2026
MCD - Moussaif Question-Difference Geometry & Resource بـ العلاقة
(M-QDGR Rationality)
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـMCD.
1. الكائن النظري والدقة الافتراضية
M-MCBOS يفرض hard floors region، MCD scores local differences، QMCD conditions scores on governing question، Q يحدد decision surface، Δ يحدد of leverage
التركيز هنا هو العلاقة بـMCD، مع منع الخلط بين schema نظري ونتيجة تجريبية. currencies substitutable-non
(EvidenceStrength_eff = Σ w_i E_i − CorrelationPenalty(E_1…E_n
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
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 تُدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملاً .

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
branch contraction، جمد model/tool/context، سجل raw trajectory، افصل process score عن outcome score، التجربة أعد ثم
measurement 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

error confidence وCalibration حيث تكون الثقة جزء
Recovery / rollback / failure localization.
Cost, latency, tool calls, context load.
environment/domain/model عبر Transfer عندما يكون الادعاء عامًا.
5. شرط الهدم
إذا لم يغير question-governing بسيط نتائج heuristic، أو ساوي graph search intervention أو outcome تحت budget matched،
QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph
تعريف entropy، information gain، shadow prices
بـMCD، 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.
•
•
•
•
•
9
<PARSED TEXT FOR PAGE: 10 / 40>
P14-C / M-QDGR / PAGE 010 / 05 September 2026
QMCD - Moussaif Question-Difference Geometry & Resource بـالعلاقة
(Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط
السؤال الصحيح فضاء الفروع ويوزع موارد
متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـQMCD.
1. الكائن النظري والدقة الافتراضية
region، MCD scores local differences، QMCD conditions scores on governing question، M-MCBOS يفرض hard floors
Q يحدد decision surface، Δ يحدد of leverage
currencies substitutable-non. التركيز هنا هو العلاقة بـQMCD، مع منع الخلط بين schema نظري ونتيجة تجريبية.
RiskAdjustedUtility = ExpectedVerifiedUtility − λ_risk·Risk − λ_cost·Cost
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
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

مع التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ MCD versus تجربة صمم matched-budget حول heuristic routing

حذف 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 عامًا.

5. شرط الهدم

إذا لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.

6. ممر التطبيق

في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph تعريف entropy، information gain، shadow prices

بـ QMCD العلاقة محور، 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.

•
•
•
•
•

(Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد
متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـMCAL.
1. الكائن النظري والدقة الافتراضية
region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors
Q يحدد decision surface؛ Δ يحدد of leverage
currencies substitutable-non. 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 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
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
التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ budget shadow-price calibration 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 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 لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج

QMCD/MCD، تَخفض الصياغة الهندسية.
6. ممر التطبيق
QuestionGraph، DifferenceField، branch، وتشغيل router calibration على APB/ToolSandbox-like internal tasks. في تعريف entropy، information gain، shadow prices
MCAI، مع typed schema وtrace_id وsource/evidence pointers وrollback receipt.
العلاقة محور بـ
حالة هذا البناء: formalization theoretical Candidate. لا claim AGI؛ لا 4E دون baseline frozen؛ لا 5E دون replication.
•
•
•
•
•
11
<PARSED TEXT FOR PAGE: 12 / 40>
P14-C / M-QDGR / PAGE 012 / 05 September 2026
HAKIM-MCD - Moussaif Question-Difference Geometry &
العلاقة بـ
(Resource Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد
متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـMCD-HAKIM.
1. الكائن النظري والدقة الافتراضية
region، MCD scores local differences، QMCD conditions scores on governing question، M-MCBOS يفرض hard floors
Q يحدد decision surface، Δ يحدد of leverage
currencies substitutable-non. التركيز هنا هو العلاقة بـMCD-HAKIM، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(Route* = argmax_r E[VerifiedUtility(r|Q,Δ)] subject to B=(tokens,time,cost,risk,privacy,authority,verification
المعادلة هنا formal research object 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
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
retrieval/tool، جمد model/tool/context، سجل raw trajectory، افصل process score عن outcome score، مع التجربة أعد ثم
تجربة صمم matched-budget حول avoidance benefit
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 لم يغير graph search intervention question-governing أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج
QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph
تعريف entropy، information gain، shadow prices
بـ العلاقة محور HAKIM-MCD، 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.
•
•
•
•
•
12
<PARSED TEXT FOR PAGE: 13 / 40>
P14-C / M-QDGR / PAGE 013 / 05 September 2026
M-HKOS - Moussaif Question-Difference Geometry & Resource بـ العلاقة
(Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط
السؤال الصحيح فضاء الفروع ويوزع موارد
متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـ HKOS-M.
1. الكائن النظري والدقة الافتراضية
region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors
Q يحدد decision surface؛ Δ يحدد of leverage
currencies substitutable-non. التركيز هنا هو العلاقة بـ HKOS-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
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 نفياً 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
أعد ثم outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ question quality transfer
تجربة صمم matched-budget حول across domains
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 عاماً.
5. شرط الهدم
إذا لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph
تعريف entropy، information gain، shadow prices
بـ العلاقة محور M-HKOS، 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.
•
•
•
•
•

HyperSkills - Moussaif Question-Difference Geometry & Resource بـ العلاقة
(Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـHyperSkills.
1. الكائن النظري والدقة الافتراضية
region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors Q يحدد decision surface؛ Δ يحدد of leverage
وnon-substitutable currencies. التركيز هنا هو العلاقة بـHyperSkills، 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 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
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 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
مع التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ question
تجربة صمم matched-budget حول intervention study
حذف 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 عامًا.
5. شرط الهدم
إذا لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى بسيط heuristic نتائج، تُخفض الصياغة الهندسية. QMCD/MCD
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل، branch، DifferenceField، QuestionGraph
تعريف entropy، information gain، shadow prices
العلاقة محور HyperSkills، 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.
.

.
.
.
.
14
<PARSED TEXT FOR PAGE: 15 / 40>
P14-C / M-QDGR / PAGE 015 / 05 September 2026
HyperAbilities - Moussaif Question-Difference Geometry & Resource Rationality (M-QDGR) بـ العلاقة
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـHyperAbilities.
1. الكائن النظري والدقة الافتراضية
يفرض M-MCBOS hard floors وQMCD conditions scores on governing question، MCD scores local differences، region يحدد Δ of leverage، يحدد Q decision surface
التركيز هنا هو العلاقة بـHyperAbilities، مع منع الخلط بين schema نظري ونتيجة تجريبية. currencies substitutable-non
RiskAdjustedUtility = ExpectedVerifiedUtility − λ_risk·Risk − λ_cost·Cost
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
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 محتملاً.
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 branch contraction؛ تجربة صمم matched-budget حول measurement
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 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 لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.
6. ممر التطبيق
في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph
تعريف entropy، information gain، shadow prices
العلاقة محور بـHyperAbilities، 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.
•
•
•
•
•

الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـHyperLoops Bounded.
1. الكائن النظري والدقة الافتراضية
يفرض M-MCBOS hard floors؛ QMCD conditions scores on governing question؛ MCD scores local differences region؛ يحدد Δ of leverage، يحدد Q decision surface
وcurrencies substitutable-non. التركيز هنا هو العلاقة بـHyperLoops Bounded، مع منع الخلط بين schema نظري ونتيجة تجريبية.
(Generalization = Conservation(Invariants | model,domain,tool,context change
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
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 تَدمج ورقة خارجية داخل MUT بوصفها «دليل صحة»؛ بل بوصفها مرجع آلية أو معيار اختبار أو نفيًا محتملًا.

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؛ MCD versus تجربة صمم matched-budget حول heuristic routing

حذف 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 عامًا.

5. شرط الهدم

إذا لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.

6. ممر التطبيق

في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph

تعريف entropy، information gain، shadow prices

بـ العلاقة محور Bounded HyperLoops، 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.

•

•

•

•

•

Agent Mesh - Moussaif Question-Difference Geometry & Resource بـ العلاقة
(Rationality (M-QDGR

الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـMesh Agent.

1. الكائن النظري والدقة الافتراضية

region؛ MCD scores local differences؛ QMCD conditions scores on governing question؛ M-MCBOS يفرض hard floors Q يحدد decision surface؛ Δ يحدد of leverage

وnon-substitutable currencies. التركيز هنا هو العلاقة بـMesh Agent، مع منع الخلط بين schema نظري ونتيجة تجريبية.

(Route* = argmax_r E[VerifiedUtility(r|Q,Δ)] subject to B=(tokens,time,cost,risk,privacy,authority,verification

المعادلة هنا formal object research وليست رقمًا تجريبيًا جاهزًا. 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

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 تُدمج 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

التجربة أعد ثم، outcome score عن process score افصل، raw trajectory سجل، model/tool/context جمد؛ budget shadow-price حول matched-budget calibration 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 لم يغير question-governing intervention graph search أو outcome تحت budget matched، أو ساوى heuristic بسيط نتائج QMCD/MCD، تُخفض الصياغة الهندسية.

6. ممر التطبيق

في APB/ToolSandbox-like internal tasks. على router calibration وتشغيل branch، DifferenceField، QuestionGraph
تعريف entropy، information gain، shadow prices
بـ العلاقة محور Agent Mesh، 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: 18 / 40>

M-EAPOS - Moussaif Question-Difference Geometry & Resource بـ العلاقة
(Rationality (M-QDGR
الأطروحة: تحويل Density Question وQMCD/MCD من heuristics منفصلة إلى هندسة space-search واقتصاد قرار يقيس كيف يضغط السؤال الصحيح فضاء الفروع ويوزع موارد متعددة العملات. في هذه الصفحة يختبر التركيز الخاص بالعلاقة بـEAPOS-M.
1. الكائن النظري والدقة الافتراضية