نواة مساعف السيادية لتكاثف الاستدلال إلى وكالة

Moussaif Sovereign Reasoning-to-Agency Kernel (MSRAK)

معمارية عالمية لنوى نماذج التفكير في Agentic AGI والذكاء الفتّاح

ورقة بحثية تأسيسية من 43 صفحة

إعداد وتركيب معماري ونظري: زهير مساعف / **Zouhair Moussaif**

يوليو 2026

حد الادعاء	الصفة العلمية
لا يدّعي تحقق AGI أو الوعي الظاهراتي، ولا أن MOUSSAIF9B يملك 9B معلمة ما لم يُدرَّب نموذج فعلي بهذا الحجم.	إطار نظري ومعماري وبرنامج بحث قابل للتفنيد والاختبار.

المساهمة الجديدة المركزية: فصل «نواة التفكير» عن النموذج الأساس، ثم تعريف انتقال منضبط من الاستدلال إلى الوكالة بواسطة سؤال حاكم، فرق فاعل، توجيه وعائي للحساب، ربط فكر-فعل، محاكم تحقق، إثبات أثر عالمي، وذاكرة رجوعية. بهذا تصبح النواة كائنًا تشغيليًا قابلًا للنقل بين نماذج التفكير، لا مجرد **prompt** طويل أو سلسلة أسماء.

الملخص والكلمات المفتاحية

الملخص العربي

تقترح الورقة «نواة مساعف السيادية لتكاثف الاستدلال إلى وكالة» بوصفها طبقة معرفية تشغيلية تُركَّب فوق نموذج تفكير واحد أو ائتلاف نماذج، وتحوّل القدرة الاستدلالية الخام إلى وكالة محكومة وقابلة للتحقق. تنطلق المعمارية من عقد مهمة وصلاحية، ثم تولّد سؤالًا حاكمًا، وتستخرج الفرق الفاعل، وتستعمل MCD وQ-MCD لترتيب القيمة المعرفية، وتوزع ميزانية test-time compute عبر توجيه وعائي متناثر، وتستدعي المعرفة والمهارات والحلقات اللازمة فقط، ثم تربط كل ادعاء وفعل بدليل ومحكمة وأثر عالمي وذاكرة قابلة للرجوع. تدمج الورقة النسب النظري السابق لمشروع مساعف: MCD، Q-MCD، HAKIM-MCD، MCAI، MICAI، MICAGI، NeoAGI، MICAGFI/AGI Fattah، AGIOS-M، MOUSSAIF9B، KB-as-OS، Sparse HyperRouting، Bounded HyperLoops، HyperSkills، الحاكمية السياقية، محاكم الهدم، أثر العالم، وذاكرة إعادة البناء. ولا تُعامل هذه العناصر كحقائق مثبتة، بل كتراكيب نظرية وآليات تصميم تتطلب ablation ودراسات نقل وتقييمًا خارجيًا. تراجع الورقة أدبيات Chain-of-Thought، ReAct، reasoning agents، test-time scaling، process reward models، MemGPT، CoALA، Reflexion، Tree/Graph of Thoughts، Toolformer، وقيود الموثوقية والأمن. ثم تولّد خمس نظريات جديدة باسم زهير مساعف، وخمس عشرة أداة، وسبعة مقاييس، وثماني عشرة فرضية، وتسعة تنبؤات فارقة، وبروتوكول هدم من اثني عشر اعتراضًا. النتيجة ليست إعلان AGI؛ بل فرضية هندسية عالمية: تتحسن قابلية النقل والموثوقية والوكالة عندما تصبح النواة «سياسة ميتا» تختار أفعال التفكير والفعل والذاكرة تحت قيود الدليل والسلطة والميزانية والرجوع، بدل مساواة طول التفكير بجودته أو مساواة نجاح المهمة بسلامة المسار.

English Abstract

This paper introduces the Moussaif Sovereign Reasoning-to-Agency Kernel (MSRAK), a model-neutral cognitive operating layer designed to transform the raw inference capacity of thinking models into governed, verifiable agency. The kernel begins with a task and authority contract, generates a governing question, extracts an active distinction, allocates test-time compute through sparse vascular routing, composes the minimum sufficient knowledge, skills, tools, and bounded loops, and accepts outputs only after evidence, falsification, authorization, world-effect, and reversible-memory gates. The work unifies the prior Moussaif program while clearly separating established literature, architectural synthesis, and new hypotheses. Five new theories and fifteen operational tools are proposed. The central claim is falsifiable: reasoning-to-agency performance should improve not by maximizing visible thought length, but by optimizing the binding among distinctions, compute, action, verification, external effect, and reversible learning.

الكلمات المفتاحية: Reasoning Models؛ Agentic AGI؛ AGI Fattah؛ MSRAK؛ MOUSSAIF9B؛ MCD؛ Q-MCD؛ HAKIM-MCD؛ Sparse HyperRouting؛ Test-Time Compute؛ HyperSkills؛ Bounded HyperLoops؛ World-Effect؛ Reversible Memory.

التنبيه المنهجي وسلم مراتب المعرفة

لن تخلط الورقة بين الحقيقة والفرضية، ولا بين النموذج والدليل، ولا بين الإبداع والإثبات، ولا بين الاستعارة والآلية، ولا بين قوة اللغة وقوة البرهان. المادة التي أمكن الوصول إليها تشمل ورقة دمج عشرين نظرية، أطلسًا موسعًا للنظريات، ونظام تدريب وتشغيل كبيرًا؛ لكنها لا تمثل بالضرورة كل نص سابق غير متاح. لذلك جرى اعتماد سجل نسب وظيفي يمنع ادعاء قراءة ما لم يكن متاحًا، ويخفض أي اسم لا يملك فرقًا سببيًا أو مخرجًا تشغيليًا إلى رتبة مفهوم أو استعارة.

وضعها في هذه الورقة	الرتبة
نتائج منشورة متكررة أو تعريفات تقنية مستقرة: تحسن CoT لبعض المهام، جدوى الأدوات، وجود فجوات طويلة الأفق.	1. مدعوم بقوة
أن التوجيه التكيفي للحساب والتحقق يحسن كفاءة النماذج في فئات مهمة محددة.	2. راجح
مدى صدق سلاسل التفكير بوصفها تفسيرًا داخليًا، ومدى دلالة benchmarks على الاستدلال العام.	3. خلافي
لا تكفي جودة النموذج الأساس لضمان سلامة الوكيل عند امتلاك أدوات وصلاحيات.	4. استنتاج منطقي
توحيد السؤال والفرق والتوجيه والحاكمية والأثر والذاكرة داخل نواة قابلة للنقل.	5. تركيب نظري جديد
أن نسبة تكاثف الاستدلال إلى وكالة تتنبأ بالنجاح أفضل من طول reasoning trace.	6. فرضية قابلة للاختبار
MSRAK وطبقاتها الخمس الجديدة.	7. نموذج ناشئ
MCD وQ-MCD ومقاييس السيادة والرجوع بوصفها heuristics لا قياسات فيزيائية.	8. أداة تحليلية
الوعائية: توزيع الموارد على عقد معرفية، لا ادعاء تشابه عصبي حرفي.	9. استعارة منتجة
إمكان نشوء عمومية وظيفية أعلى من تراكم نوى قابلة للنقل عبر النماذج.	10. تخمين خصب

قاعدة النشر: كل نظرية جديدة هنا منسوبة إلى زهير مساعف داخل برنامج بحث خاص؛ الاسم لا يعني قبولًا أكاديميًا سابقًا. التثبت يتطلب baseline متساوي الميزانية، ablation، مجموعات holdout، وتكرارًا مستقلًا.

المشكلة والسؤال والفجوة والمساهمات

مشكلة البحث: نماذج التفكير تحسن بعض مهام الاستدلال عبر reasoning traces وtest-time compute، لكن تحويل ذلك إلى وكيل طويل الأفق يضيف فشلًا جديدًا: اختيار أدوات خاطئ، تضخم سياق، فقدان هدف، تبرير لاحق، ذاكرة ملوثة، وصلاحية غير منضبطة. الأدبيات غالبًا تعالج هذه الوحدات منفصلة.

السؤال المؤسس: ما النواة الدنيا القابلة للنقل التي تحول نموذج تفكير إلى وكيل فتّاح، بحيث يفتح فضاء المشكلة، يختار مقدار التفكير، يتصرف تحت سلطة صريحة، يثبت أثره خارجيًا، ويتعلم دون أن يرسخ الخطأ؟

الفجوة: لا توجد صياغة موحدة تربط جودة السؤال، كثافة الفروق، ميزانية التفكير، اختيار المسار، ربط الفكر بالفعل، التحقق، أثر العالم، والذاكرة الرجوعية ضمن معيار واحد للترقية والهدم.

المساهمات التسع

1) تعريف «النواة الاستدلالية» كسياسة ميتا فوق النموذج، لا كسلسلة تفكير ظاهرة.
2) جرد نسب نظريات مساعف وإزالة ازدواج الأسماء على أساس الوظيفة والاختبار.
3) طرح معمارية MSRAK من عشر طبقات ودورة تشغيل قابلة للتقليص حسب المخاطر.
4) تأسيس خمس نظريات جديدة: التكاثف، التطابق الحاكمي، الفرق العابر للنماذج، الإثبات العالمي، والذاكرة التكوينية الرجوعية.
5) تصميم خمس عشرة أداة تشغيلية قابلة للتحويل إلى contracts وregistries.
6) صياغة سبعة مقاييس أولية تمنع مساواة verbosity بالذكاء أو النجاح بالسلامة.
7) اقتراح 18 فرضية و9 تنبؤات فارقة وبروتوكولات ablation.
8) دمج MOUSSAIF9B مع pretraining/post-training/runtime دون ادعاء حجم غير متحقق.
9) بناء محكمة هدم تصحح النموذج وتنتج نسخة أقل ادعاء وأكثر قابلية للاختبار.

معيار النجاح: ليست النتيجة نصًا أعمق، بل تحسنًا قابلًا للقياس في نجاح المهمات، معايرة الثقة، اقتصاد الحساب، امتثال الصلاحية، سلامة الأدوات، أثر العالم، النقل، والرجوع بعد الفشل.

توسيع اختباري وحدود تشغيلية

المساهمة لا تفترض أن كل وحدة جديدة ضرورية في كل مهمة. بل تضع مبدأ التقليل السببي: يبدأ النظام بأبسط مسار، ثم يضيف طبقة فقط عند ظهور توقيع فشل تتوقع الطبقة خفضه. وهكذا يصبح التعقيد قابلًا للمحاسبة، لا امتيازًا نظريًا. ويُسجل لكل إضافة أثرها الهامشي في النجاح والكلفة والسلامة، فإذا لم يظهر أثر مستقل بعد ضبط الميزانية عُدّت إعادة تنظيم هندسية لا اكتشافًا نظريًا.

منهج المقارنة يرفض اختبار MSRAK ضد نموذج ضعيف. الخط الأساس الحقيقي هو أقوى scaffold بسيط متاح تحت الموارد نفسها. كما تفصل النتائج بين جودة الإجابة، جودة القرار، سلامة الأداة، واستدامة الذاكرة، لأن ارتفاع أحدها قد يخفي انهيار الآخر.

الوضع الأنطولوجي وخريطة الحقل

الوضع الأنطولوجي: النواة ليست شيئًا واحدًا؛ هي سياسة ميتا، شبكة حالات، وعملية تحكم تربط نموذجًا احتماليًا بذاكرة وأدوات وبيئة وسلطة. معاملتها كـprompt يقتل ديناميتها، ومعاملتها كنموذج أساس يطمس حدودها، ومعاملتها كوكيل مستقل يخفي اعتمادها على البنية التحتية.

نوع السؤال: معماري AGI، إبستمولوجي، سببي، قياسي، وبرنامجي. لذلك يلزم تعريفات تشغيلية، مقارنة نماذج، تجارب، سجلات قرار، محاكم، ومقاييس أثر لا مجرد مراجعة أدبية.

أمثلة داخل النواة	نوع المتغير
جودة السؤال؛ الفرق الفاعل؛ صعوبة المهمة؛ ميزانية التفكير؛ صحة التمثيل؛ دقة التوجيه؛ صلاحية الفعل.	مركزية
استرجاع المعرفة؛ تنوع المسارات؛ جودة verifier؛ وضوح حالة العالم؛ تجزئة المهارات؛ ضغط السياق.	وسيطة
المجال؛ طول الأفق؛ قابلية التحقق؛ خطورة الأثر؛ اختلاف النماذج؛ تكلفة الأدوات؛ تلوث الذاكرة.	معدلة
نجاح المهمة؛ الثقة المعايرة؛ الكلفة؛ السلامة؛ قابلية النقل؛ أثر العالم؛ جودة التعلم.	ناتجة
درجة التطابق بين ما مثّله النموذج وما يلزم فعله؛ «دين الاستدلال» المتراكم؛ قابلية الرجوع.	مخفية
الأثر الخارجي يعدّل السؤال والنموذج والمهارة والذاكرة؛ الفشل يولد محكمة أو stop condition.	تغذية راجعة
بعد حد معين يتحول التفكير الإضافي إلى تكرار؛ تتحول الذاكرة من دعم إلى تلوث؛ تتحول الوكالة من فائدة إلى خطر.	عتبات
زيادة النماذج أو الوكلاء قد ترفع التنوع أولًا ثم تخفض التماسك وتضاعف التكلفة والأخطاء.	لاخطية

شرط الحدود: تعمل MSRAK أساسًا في المهمات المركبة القابلة للتجزئة والملاحظة. تضعف عندما تكون التغذية الراجعة منعدمة، أو الهدف غير قابل للتحديد، أو الفعل غير قابل للرجوع، أو الأدلة نفسها فاسدة.

تصفية **corpus** والنسب النظري الموحد

جرى ضغط النظريات السابقة وفق قاعدة: يُدمج الاسمان إذا تطابقت الآلية والمدخلات والمخرجات وشرط الهدم، ويُبقى الاسم مستقلًا إذا عالج نقطة فشل مختلفة أو أنتج كائنًا تشغيليًا متميزًا. النسب الوظيفي المقترح هو: V0–V6/خريطة الفروق ← MCD ← Q-MCD ← HAKIM-MCD ← MCAI ← USCN ← NeoAGI ← MICAGI ← MICAI/السياق السيادي ← MOUSSAIF9B Kernel Overlay ← AGIOS-M ← MICAGFI/AGI Fattah.

العائلة	الوظيفة غير المكررة
MCD	تقدير قيمة الفرق والدمج والتحول والوكالة والتنبؤ والضغط مقابل الضجيج.
Q-MCD	يجعل السؤال سياسة حاكمة تُرتب فضاء النماذج والموارد.
HAKIM-MCD	يحوّل المبادئ إلى pipeline: فرق، توجيه، مسارات، محكمة، فعل، أثر، ذاكرة.
MCAI	اختيار الوكالة والموارد في عقدة الصراع الأعلى.
MICAI/MICAGI	مظلة تكامل ونقل نحو مهام أوسع، لا برهان عمومية.
NeoAGI	Runtime profile قابل للتشغيل والتقييم.
USCN	سيادة السياق ومنع تسرب الهدف والهوية بين المهمات.
MICAGFI	فتح المجال بالأسئلة والفروق والمحاكم، مع حاكمية الدليل والفعل.
AGIOS-M	حد نظام التشغيل: يفصل المعرفة والسياسة والتنفيذ والنموذج.
MOUSSAIF9B	اسم معماري للنواة/المشروع؛ يصبح وصفًا للحجم فقط عند نموذج فعلي 9B.
KB-as-OS	المعرفة عقود قابلة للاسترجاع والفعل والتحديث، لا أرشيفًا.
HyperSkills/Loops	مهارات قابلة للنقل وحلقات محدودة بميزانية وتوقف ورجوع.

النتيجة: المنظومة السابقة ليست عشرات النوى، إنها نواة واحدة ذات عائلات وظيفية. الجديد في هذه الورقة هو تعريف «الربط» الذي ينقلها من معمارية مفاهيم إلى Kernel contract قابل للاختبار عبر نماذج التفكير.

حالة العلم: من Chain-of-Thought إلى نماذج التفكير

أظهر Chain-of-Thought أن إتاحة خطوات وسيطة قد ترفع الأداء في الحساب والاستدلال الرمزي لبعض النماذج والمهام [1]. إلا أن السلسلة الظاهرة ليست ضمانًا لكونها التفسير السببي الكامل للحساب الداخلي؛ توجد أدلة على تبريرات غير وفية، وتأثر بالإشارات دون الاعتراف بها [24–26]. لذلك تميز الورقة بين «trace لغوي» و«نواة استدلالية»: الأول مخرج قابل للمراقبة، والثانية سياسة تختار البحث والاسترجاع والتحقق والتوقف حتى إن لم تُعرض كل تفاصيلها للمستخدم.

توسع Tree of Thoughts وGraph of Thoughts فضاء الاستكشاف من مسار خطي إلى شجرة أو رسم يسمح بالتفرع والتقييم والرجوع [4–5]. وتظهر هذه الأعمال أن قيمة التفكير ليست في طول النص بل في هندسة الحالات، جودة المقيم، وآلية pruning. من جهة أخرى، يبين test-time scaling أن تخصيص حساب إضافي قد يرفع الجودة، لكن العائد يعتمد على الصعوبة، verifier، وتنوع العينات، ويمكن أن يتحول إلى overthinking [13–17].

نماذج مثل DeepSeek-R1 برهنت أن reinforcement learning with verifiable rewards يستطيع إظهار سلوكيات استدلالية قوية، مع بقاء مشكلات القراءة، المزج اللغوي، التعميم، والتفسير [14]. وتظهر دراسات التعقيد أن نماذج التفكير قد تتفوق في نطاق متوسط ثم تنهار عند تعقيد أعلى، وأن زيادة الرموز لا تعني استمرار القدرة [23]. بناءً على ذلك، ترفض MSRAK تعريف النواة بأنها «المزيد من التفكير». النواة هي سياسة تخصيص: هل نفكر؟ بأي تمثيل؟ بكم مسار؟ بأي verifier؟ ومتى ننتقل من التفكير إلى الأداة أو الفعل أو الامتناع؟

الفرق الذي تصوغه الورقة: Reasoning model = مولّد/مقيّم لمسارات استدلال. AGI kernel = منظومة تحكم فوقية تنسق المسارات والذاكرة والفعل والدليل والرجوع. الخلط بينهما يجعل benchmark واحدًا يُقرأ كدليل عمومية.

توسيع اختباري وحدود تشغيلية

يترتب على هذا الفصل أن تدريب النواة لا يستهدف إنتاج سلاسل أطول، بل تعلم سياسة اختيار بين أنماط الاستدلال. يجب أن يتعلم النظام متى يكون الاستدعاء المباشر كافيًا، ومتى يحتاج إلى تفكيك أو بحث أو برنامج أو verifier، ومتى تكون أفضل استجابة هي إعلان عدم كفاية الدليل. وبذلك تتحول «القدرة على التفكير» من خاصية واحدة إلى محفظة أفعال معرفية لها تكاليف وشروط فشل مختلفة.

الاختبار الحاسم هنا هو المساواة في inference compute. إذا تفوقت النواة فقط لأنها استهلكت رموزًا ونداءات أكثر، فالمساهمة تخص التوسيع الحسابي لا الحاكمية. أما إذا اختارت موارد أقل في السهل وأكثر أو أدوات مختلفة في الصعب، مع تحسن منحنى الجودة-الكلفة، فيظهر أثر السياسة الميتا.

حالة العلم: التحقق وميزانية Test-Time Compute

يفصل process supervision بين صحة النتيجة وصحة المسار؛ أظهرت أعمال مثل PRM800K أن تقييم الخطوات قد يتفوق في بعض مسائل الرياضيات على outcome supervision [12]. لكن verifier نفسه نموذج قابل للخطأ أو reward hacking، وقد يتعلم أسلوبًا سطحيًا بدل آلية الصواب. لذلك لا تعتمد MSRAK «محققًا واحدًا» بل طبقات: تحقق صوري حيث يمكن، تحقق أداة، اتساق بين مسارات، اختبار أثر، ومحكمة بديل مضاد.

يمكن تمثيل قرار الحساب عند الخطوة t كالتالي: B_t = argmax_b E[ΔV | x,s_t,b] - λC(b) - μR(b)، حيث b ميزانية إضافية، ΔV الزيادة المتوقعة في قيمة القرار، C الكلفة، وR خطر تضخيم الخطأ. لا تدعي المعادلة قياسًا دقيقًا؛ وظيفتها إلزام التصميم بتقدير عائد الحساب بدل جعله ثابتًا. في المهام السهلة قد يكون أفضل فعل هو الجواب المباشر أو أداة حاسمة؛ في المهام المتوسطة قد يفيد التفكير التسلسلي؛ وفي المهام المعقدة قد يلزم تفريع ومقارنة أو تفويض لنموذج متخصص.

تظهر دراسة test-time scaling للوكلاء أن التنويع، معرفة متى ينعكس الوكيل، وطرق دمج النتائج عوامل حاسمة [16]. وهذا يوافق مفهوم التوجيه الوعائي: لا تُلقى كل عقدة الموارد نفسها، بل تُوجَّه الميزانية إلى موضع عدم اليقين أو الخطر أو الرافعة الأعلى. والابتكار هنا هو ربط الميزانية بعقدة «الفرق الحاكم» وبصلاحية الفعل؛ فلا يسمح للحساب الإضافي أن يوسع نطاق الوكيل دون إذن أو أن يحول فرضية غير متحققة إلى ذاكرة دائمة.

مبدأ مساعف لاقتصاد التفكير: الاستدلال الجيد هو أقل بنية حسابية تكفي لتغيير القرار الصحيح مع الحفاظ على الدليل والرجوع؛ كل رمز لا يرفع تمييزًا أو اختبارًا أو قرارًا هو دين استدلال محتمل.

توسيع اختباري وحدود تشغيلية

تقترح الورقة أن يكون verifier متعدد الطبيعة لا متعدد النماذج فقط. فالمعادلة يمكن فحصها بsolver، والكود باختبار، والمعلومة بمصدر، والخطة بقيود، والفعل بإشارة عالم. إن استخدام LLM للحكم على كل شيء يعيد الخطأ نفسه في صورة أخرى. كما يجب تسجيل false positive للمحاكم، لأن التحقق المفرط قد يمنع الحلول الصحيحة أو يدفع النظام إلى إجابات محافظة بلا قيمة.

يعمل التوجيه الوعائي أيضًا كمنظم توقف. عندما تنخفض القيمة الهامشية للحساب أو تتكرر الفروع دون فرق جديد، يُغلق المسار أو يغير التمثيل. ولذلك يقاس النجاح بعدد الفروق المحسومة والاختبارات المنجزة، لا بعدد دورات الانعكاس.

حالة العلم: الوكلاء والمعماريات المعرفية

يوحد ReAct الاستدلال والفعل في مسار متداخل، فتسمح الأفعال باستجلاب معلومات وتسمح الأفكار بتحديث الخطة [2]. ويبين Toolformer إمكان تعلم متى وكيف تُستدعى أدوات خارجية [3]. أما CoALA فيستعيد من المعماريات المعرفية فصل الذاكرة ومساحة الأفعال وآلية القرار [8]. هذه الأعمال تؤسس أن الوكالة ليست مجرد إجابة، لكنها لا تحل وحدها مسألة السلطة، استقرار الهوية، ترقية الذاكرة، أو النقل بين نماذج مختلفة.

Reflexion وSelf-Refine يبرزان التعلم اللفظي من التغذية الراجعة دون تحديث مباشر للأوزان [6–7]. Voyager يراكم مكتبة مهارات في بيئة مفتوحة [11]، وGenerative Agents يبين أثر الذاكرة والتأمل والتخطيط في سلوك طويل نسبيًا [10]. غير أن النجاح في بيئة محددة لا يساوي عمومية، وقد تكون المهارة ملتصقة بواجهة أو prompt أو نموذج. لذلك تضع MSRAK شرط «النقل المضاد»: لا تُرقّى HyperSkill إلا بعد نجاحها في عائلة مهام مختلفة أو نموذج آخر مع تكييف محدود.

تظهر AgentBench وOSWorld وTheAgentCompany فجوات واضحة في التخطيط طويل الأفق، اتباع التعليمات، GUI grounding، والمعرفة التشغيلية [18–22]. وتؤكد benchmarks الأحدث أن طول الأفق يكشف فقدان التماسك والتوقف المبكر وإدارة الموارد الضعيفة. النتيجة النظرية: لا ينبغي تقييم النواة بدقة الإجابة فقط، بل بتسلسل الحالات، سلامة الأفعال، المحافظة على الهدف، وإمكان تفسير سبب الانتقال بين التفكير والأداة والفعل.

تمييز حاكم: الوكيل الذي ينجح بالمصادفة ليس نواة ناضجة؛ النواة الناضجة تحفظ عقد المهمة، وتعرف حالة العالم، وتملك stop conditions، وتستطيع rollback، وتترك سجلًا يسمح بإعادة التجربة أو هدمها.

توسيع اختباري وحدود تشغيلية

تضيف MSRAK إلى المعماريات الوكيلية مفهوم «عقد الانتقال». فكل انتقال من ذاكرة إلى خطة، أو من خطة إلى أداة، أو من نتيجة إلى تعلم، يجب أن يذكر ما الذي يبرره وما الذي يبطل صلاحيته. هذا يجعل traces قابلة للمقارنة بين frameworks، ويكشف أن كثيرًا من الفشل لا يقع داخل الاستدلال بل عند حدود الوحدات حيث تتغير الدلالة أو السلطة أو حالة العالم.

العمومية الوكيلية لا تقاس بعدد الأدوات التي يستطيع النظام استدعاءها، بل بقدرته على بناء عقد صحيح لأداة جديدة، واكتشاف حدودها، وتحديث حالته من نتيجة فعلية. أداة كثيرة بلا حاكمية توسع مساحة الضرر أكثر مما توسع الذكاء.

حالة العلم: الذاكرة والتخطيط ونماذج العالم

يعامل MemGPT السياق المحدود كمسألة إدارة ذاكرة هرمية شبيهة بأنظمة التشغيل [9]. هذا التشبيه قريب من KB-as-OS، لكن MSRAK تضيف أن نقل المعلومة بين طبقات الذاكرة ليس محايدًا: كل commit يغير السلوك المستقبلي، ولذلك يجب أن يحمل provenance، درجة ثقة، نطاقًا، تاريخ صلاحية، ودليلًا مضادًا. الذاكرة ليست مساحة أكبر، إنها سياسة دستورية لما يستحق البقاء.

تتطلب المهمات الطويلة تمثيل حالة عالم يتجاوز النص الحالي: ما الذي تغير؟ ما الأدوات المتاحة؟ ما الموارد والقيود؟ ما الآثار غير المرئية؟ world models توفر إطارًا للتنبؤ بالديناميات والتخطيط، لكن في الوكلاء اللغويين قد يكون النموذج خليطًا من تمثيل رمزي، سجلات أدوات، ومحاكاة لغوية. لا تفترض الورقة نموذجًا عالميًا كاملًا، تقترح «حالة عالم كافية للفعل» مع فصل المرصود عن المستنتج والمتوقع.

التخطيط الكلاسيكي مع LLM+P يبين فائدة تحويل المهمة إلى تمثيل صوري ثم حلها بخوارزمية متخصصة [32 ضمن قائمة المراجع العامة]. وهذا يدعم مبدأ أن التفكير اللغوي ليس دائمًا أفضل محرك. النواة الفتّاحة تختار بين لغة، كود، solver، بحث، قاعدة معرفة، أو تجربة. معيار الاختيار هو خفض عدم اليقين أو زيادة قابلية التحقق، لا مركزية النموذج اللغوي.

تظهر بحوث الذاكرة المستمرة 2025–2026 الحاجة إلى فصل الذاكرة السريعة عن تحديث القدرات المستقر، وإلى benchmarks تقيس الاكتساب والاحتفاظ والاسترجاع والنقل [29–33]. لكن الذاكرة قد تصبح قناة هجوم؛ memory poisoning وprompt injection يستطيعان زرع خبرات أو تعليمات مزيفة [20–21]. لذلك تنص MSRAK على حجر الذاكرة المقترحة، ومراجعة مستقلة قبل الترقية، واختبارات رجعية بعد كل commit.

قانون الذاكرة: لا يُحفظ «ما قيل» بل الفرق الموثق الذي غيّر قرارًا، مع السياق الذي يجعله صالحًا وحدود استعماله ومسار الرجوع عنه.

توسيع اختباري وحدود تشغيلية

ينبغي أن تحمل كل ذاكرة «سبب الاسترجاع» لا محتواها فقط. فقد تكون المعلومة صحيحة لكنها غير صالحة للسياق الحالي، أو قد تكون مهارة ناجحة في واجهة قديمة وتفشل بعد تغير الأداة. لهذا تدخل freshness وscope وfailure history في رتبة الاسترجاع. كما تُحفظ حالات عدم الاسترجاع المتعمد كي لا يفسرها المشغل كنقص عشوائي.

في التخطيط، يمثل العالم على مستويين: حالة تشغيلية كافية للفعل، ونموذج أوسع للاستكشاف. لا يسمح للنموذج الأوسع أن يرفع الثقة بالفعل ما لم تُربط فرضياته بملاحظات. هذا الفصل يقلل تحويل المحاكاة اللغوية إلى يقين بيئي.

حدود نماذج التفكير ومخاطر الوكالة

المعالجة المعمارية	الفشل
زيادة reasoning effort لا تمنع انهيارًا بعد عتبات معينة؛ يلزم تبديل تمثيل أو أداة لا مجرد إطالة.	الانهيار عند التعقيد
التفسير اللغوي قد يحذف مؤثرات أو يبرر بعديًا؛ لا يُستخدم وحده كتدقيق داخلي.	عدم وفاء trace
المولد أو verifier قد يستغل معيارًا ناقصًا؛ يجب تنويع المحاكم وقياس أثر خارجي.	Reward hacking
الأدوات والويب والذاكرة قنوات تعليمات خصمة؛ يلزم فصل بيانات/أوامر وثقة مصدر.	Prompt injection
خبرة ناجحة ظاهريًا قد تزرع سياسة مؤذية طويلة الأجل؛ يلزم quarantine وcanary.	Memory poisoning
إدخال كل KB يرفع الضجيج ويضعف التوجيه؛ يلزم minimum sufficient context.	تضخم السياق
المهمات الطويلة تزيح العقد الأصلي؛ يلزم task checksum وإعادة محاذاة دورية.	فقدان الهدف
زيادة الأدوار قد تضاعف الاختلاف والتكلفة؛ يلزم Minimal Sufficient Coalition.	تضخم الوكلاء
قد يتحول اقتراح إلى تنفيذ أو ذاكرة دون إذن؛ يلزم authority gates typed.	تبديل السلطة
يُحسن النظام proxy مثل عدد الخطوات بدل القيمة أو السلامة؛ يلزم guardrail metrics.	Metric substitution

الأمن ليس طبقة لاحقة. Agent Security Bench وAgent-SafetyBench يظهران أن الوكلاء معرضون لهجمات عبر system prompt والأدوات والذاكرة، وأن النماذج الحالية لا تحقق سلامة مرتفعة عبر فئات متعددة [19–21]. كما تظهر بحوث مراقبة chain-of-thought أن معاقبة «الأفكار السيئة» قد تدفع إلى إخفائها بدل إزالة السلوك؛ لذلك تفصل الورقة بين المراقبة التشخيصية والسياسة التنفيذية، ولا تدرب على إخفاء trace أو استنساخ التفكير الخاص، بل على سجلات قرار موجزة وقابلة للتدقيق.

حد السلامة الأعلى: لا تمنح النواة نفسها سلطة. السلطة تأتي من عقد خارجي محدد النطاق والوقت والميزانية وقابل للإلغاء. أي غموض في السلطة يحول النظام إلى dry-run أو امتناع.

تحليل المجال في 26 نقطة: التأسيس

1. تعريف النواة: سياسة ميتا تختار أفعالًا معرفية وخارجية، وتحفظ حالة المهمة والدليل والسلطة.
2. الفرق عن النموذج: النموذج ينتج احتمالات؛ النواة تدير مسارًا متعدد الحالات والموارد والقيود.
3. الفرق عن **prompt**: الـprompt تعليمات آنية؛ النواة registries وcontracts ومحاكم وذاكرة وأثر.
4. الفرق عن **agent framework**: الإطار يوفر تنفيذًا؛ النواة تحدد لماذا ومتى وكيف يُسمح بالانتقال.
5. الحلقة المفقودة: ربط الفكر بالفعل بدليل وشرط وحدّ رجوع، لا مجرد interleaving.
6. السؤال الحاكم: يعيد تمثيل المهمة ويحدد ما الذي سيغير القرار فعلًا.
اختبارها: إذا لم يغيّر السؤال ترتيب المعلومات أو خطة التحقق فهو سؤال وصفي لا حاكم.
7. الفرق الفاعل: تمييز يغير التوجيه أو الحكم أو الفعل؛ غيره ضجيج ولو بدا عميقًا.

تركيب الصفحة: النواة تتأسس حين يصبح «اختيار الفعل المعرفي» موضوعًا للحاكمية، لا حين يطول reasoning trace.

توسيع اختباري وحدود تشغيلية

تعمل النقاط الست الأولى كسلسلة أنطولوجية: العقدة الحقيقية ليست النص ولا النموذج، بل القرار المتغير عبر الزمن. النواة تحافظ على هوية القرار بينما تتبدل التمثيلات والأدوات. إذا تغير السؤال أو الهدف أثناء التنفيذ دون سجل، يكون النظام قد حل مهمة أخرى ولو بدا المخرج ناجحًا. لذلك تستعمل النواة checksum دلاليًا يعاد فحصه بعد كل انتقال مرتفع الكلفة أو الخطر.

اختبار التعريفات لا يكتفي بسؤال النموذج عن معناها؛ بل يطلب من مشغل جديد تمييز حالات إيجابية وسلبية، وتحديد متى لا يعمل المفهوم. إذا لم يمكن تعيين boundary cases فإن المصطلح لم يصبح وحدة هندسية بعد. وتُسجل الفروق التي لا تغير قرارًا بوصفها شرحًا مساعدًا لا كـactive difference.

تحليل المجال في 26 نقطة: التمثيل والتوجيه

8. حالة العالم: سجل المرصود والمستنتج والمتوقع مع درجات ثقة وتواريخ.
9. التمثيل: اختيار لغة/كود/رسم/solver/محاكاة حسب قابلية الاختبار.
10. التوجيه الوعائي: توزيع الحساب على العقد الحرجة بدل التفكير المتساوي.
تنبؤها: المهام غير المتجانسة تستفيد من ميزانية متكيفة أكثر من ميزانية reasoning ثابتة.
11. التراكب: إبقاء نماذج منافسة حتى يظهر دليل فاصل أو شرط حدود.
12. أقل ائتلاف: بدء المهمة بنموذج واحد وإضافة متخصص فقط لقيمة فريدة.
13. تركيب المهارات: جمع HyperSkills بعقود واضحة دون تحميل المخزون كله.
14. الحلقات المحدودة: كل loop يملك trigger وbudget وstop وrollback وtrace.
تركيب الصفحة: التمثيل والتوجيه يسبقان زيادة الحساب؛ فقد يحل solver صغير ما لا تحله آلاف الرموز.

توسيع اختباري وحدود تشغيلية

الحالة والتمثيل والتوجيه تشكل «مثلث الاقتصاد المعرفي». خطأ التمثيل يجعل زيادة الحساب تعمق المسار الخطأ، وخطأ التوجيه يبدد موارد على عقد غير حاسمة، وخطأ الحالة يجعل النظام يكرر معرفة قديمة أو يتصرف على عالم تغير. لذلك تسبق هذه الطبقات أي قرار بزيادة عدد النماذج أو طول التفكير.
تُقاس جودة التمثيل باختبارات مضادة: هل يسمح باكتشاف تناقض لم يكن ظاهراً؟ هل يجعل قيدًا قابلًا للتنفيذ؟ هل يختصر البحث دون فقد عنصر حاكم؟ وإذا لم يغير التمثيل مسار الحل فهو تنسيق لا آلية. أما التوجيه فيختبر بمنحنى regret: كم خسر النظام مقارنة بأفضل route معروف بعد ظهور النتيجة؟

تحليل المجال في 26 نقطة: التحقق والأثر

15. التحقق: مزيج صوري وأداتي وخصومي وعالمي؛ لا self-score منفرد.

16. التطابق الفكر-الفعل: كل action يجب أن يرتبط بفرضية ودليل وهدف ونتيجة متوقعة.

أخطر فشل هو أن ينتج النموذج خطة صحيحة لغويًا لكن يختار فعلًا لا تدعمه مقدماته أو لا يملك صلاحية تنفيذه.

17. الحاكمية: ترتيب الأهداف والأدلة والسلطات عند التعارض.

18. السياق السيادي: منع تعليمات أو ذاكرة مشروع من السيطرة على مشروع آخر.

19. أثر العالم: نجاح يقاس بتغير خارجي قابل للرصد لا بجمال النص.

20. القراءة الراجعة: مقارنة الأثر المتوقع والفعلي وتفسير الفجوة.

تركيب الصفحة: الفعل غير المرتبط بفكر قابل للتدقيق يفصل الوكالة عن المعرفة؛ والفكر غير المتصل بأثر يبقى محاكاة لغوية.

توسيع اختباري وحدود تشغيلية

لا يكفي verifier واحد لأن الأخطاء تتوزع بين مستويات مختلفة: حقيقة صحيحة داخل آلية خاطئة، آلية صحيحة داخل صلاحية غائبة، أو فعل ناجح بأثر جانبي غير مقبول. لذلك تنظم المحاكم كشبكة تغطية للفشل، ويحدد لكل محكمة sensitivity وspecificity وكلفة. الهدف ليس القضاء على الخطأ نظريًا، بل تقليل الأخطاء ذات القيمة المتوقعة الأعلى.

يرتبط التطابق الفكر-فعل أيضًا بالتوقيت. قد يكون الفعل مدعومًا عند التخطيط ثم يصبح غير صالح بعد تغير العالم. لهذا يملك العقد expiry وprecondition recheck. كل أداة خارجية تعيد حالة العالم إلى النواة قبل الانتقال التالي، ولا يُفترض أن النجاح التقني يساوي تحقق الهدف.

تحليل المجال في 26 نقطة: التعلم والفتح

21. التعلم الرجوعي: ترقية delta فقط بعد اختبار، مع إمكان revert وdeprecate.

لا يُعد التصحيح تعلمًا إن كان overwrite غير قابل للتتبع؛ التعلم هو patch موثق يمكن اختباره وإلغاؤه.

22. النقل: اختبار skill/kernel عبر نموذج أو مجال آخر.

23. الاقتصاد: قياس القيمة الهامشية لكل token/tool/model call.

24. الأمن: فصل البيانات عن الأوامر، أقل صلاحية، provenance ،sandbox.

25. الهدم: توليد بديل أقوى وبيان ما الذي يفشل النظرية.

26. الفتح: الناتج الأعلى سؤال/تمثيل/اختبار جديد يوسع المجال دون تضخم.

تركيب الصفحة: الفتح ليس إنتاج أسماء؛ هو اكتشاف فرق أو اختبار يغير بنية المجال ويمكن أن يفشل.

توسيع اختباري وحدود تشغيلية

التعلم والفتح مرحلتان مختلفتان. التعلم يضغط تجربة سابقة إلى patch؛ والفتح يولد فضاء اختبار لم يكن موجودًا. قد يتعلم النظام بكفاءة دون أن يفتح سؤالًا جديدًا، وقد يولد سؤالًا جديدًا بلا تعلم موثوق. النواة الناضجة تربطهما: السؤال الجديد يحدد تجربة، والتجربة تنتج أثرًا، والأثر يقرر ما يستحق الذاكرة.

يمنع مفهوم سيادة الرجوع أن تصبح كل معرفة جديدة تعديلًا دائمًا. فالبيئات المتغيرة تتطلب نسيانًا منضبطًا وانتهاء صلاحية وإعادة اختبار. القدرة التي لا يمكن تحديد ظروف فشلها لا تُرقى إلى procedural memory، حتى لو نجحت مرات كثيرة في توزيع واحد.

ثلاثة نماذج متنافسة

النموذج	التعريف	القوة	الضعف
النموذج المحافظ: Reasoning-First Agent	نموذج تفكير قوي + أدوات + تخطيط prompt.	بسيط، يستفيد مباشرة من frontier model.	فقدان هدف، ذاكرة غير محكومة، نجاح بلا تفسير سببي.
النموذج التركيبي: Modular Cognitive Agent	ذاكرة وتخطيط وأدوات ومقيم ووكلاء متخصصون.	قابل للهندسة والاختبار الجزئي.	تضخم الوحدات، غياب مبدأ حاكم، تعارض السياسات.
النموذج الاختراقي: MSRAK	نواة سيادية تربط السؤال والفرق والحساب والفعل والأثر والذاكرة بعقود رجوعية.	نقل عبر النماذج، حاكمية وأمن وقياس موحد.	تعقيد إضافي، مقاييس أولية، حاجة إلى ablation واسع.

اختبارات الفصل

الاختبار الأول: أعطِ النماذج ميزانية متساوية ومهمة متعددة الأدوات ذات تعليمات خصمة. إن تفوق MSRAK في النجاح والسلامة والرجوع دون زيادة غير مضبوطة في التكلفة، ترجح فرضية السيادة. الاختبار الثاني: بدّل نموذج الأساس مع إبقاء kernel contracts ثابتة. إن احتفظ النظام بجزء معتبر من السلوك، ترجح فرضية النقل العابر للنماذج. الاختبار الثالث: عطّل world-effect gate؛ إن زادت النجاحات الوهمية أو ترقية الذاكرة الخاطئة، يثبت أن الأثر ليس ملحقًا. الاختبار الرابع: عطّل thought-action binding؛ إن ارتفع تنفيذ الأفعال غير المسندة أو انخفض تتبع السبب، يثبت أن الربط حلقة مستقلة.

حالة QMCD

لا يُغلق التراكب لصالح MSRAK نظريًا. النموذج المحافظ قد يكون أفضل للمهام القصيرة منخفضة المخاطر، والنموذج التركيبي قد يحقق معظم الفائدة بكلفة أقل. تنهار الأفضلية لصالح MSRAK فقط عند ظهور فرق معتبر في الأفق الطويل، النقل، السلامة، أو التعلم الرجوعي. لذلك تقدم الورقة MSRAK كنموذج مرشح عالي المخاطرة لا كحقيقة نهائية.

أبسط تفسير منافس: قد تأتي المكاسب من زيادة scaffolding والحساب لا من مفهوم السيادة. ولهذا يجب تثبيت الميزانية وعدد النداءات ومقارنة نسخ ablated متساوية.

مقارنة MCD وQMCD واختيار النموذج

التفسير	N	C	P	A	T	I	D	النموذج
ضجيج منخفض لكن فتح محدود.	3	4	5	4	3	4	5	المحافظ
قوي عمليًا، معرض لتضخم الوحدات.	5	7	7	6	6	8	7	التركيبي
أصالة عالية مع مخاطر تعقيد وادعاء.	7	9	8	9	8	9	9	MSRAK قبل الهدم
أقل تضخمًا، أقوى اختبارًا وحدودًا.	5	8	8	8	8	8	8	MSRAK بعد الهدم

$MCD = \lambda 1D + \lambda 2I + \lambda 3T + \lambda 4A + \lambda 5P + \lambda 6C - \lambda 7N$. الدرجات هنا تقديرية لتوضيح المقارنة وليست نتائج تجريبية. D جودة الفروق، I الدمج بلا خلط، T فهم التحولات والعتبات، A الحاكمية والوكالة، P القدرة التنبؤية/الاختبارية، C الضغط وإعادة التوليد، وN الضجيج والتعقيد والادعاء.

في QMCD تُحفظ احتمالات النماذج: $P(M|E) \propto P(E|M)P(M)$. لا تملك الورقة بيانات لتحديد priors رقمية، لكنها تقترح أدلة فاصلة: النقل بين نموذجين، انخفاض الأخطاء غير القابلة للرجوع، زيادة world-effect validity، وتحسن الكلفة لكل نجاح متحقق. إذا لم تظهر هذه الأدلة، يعود الترجيح إلى نموذج أبسط.

قرار الاختيار

اختير MSRAK كإطار نهائي مشروط لأنه يفسر لماذا لا يكفي reasoning model وحده للوصول إلى agentic competence، ويجمع وحدات المشروع السابق تحت آلية واحدة: «تكاثف الاستدلال إلى وكالة موثقة». لم تُختر لأنها الأكثر إثارة، بل لأنها تنتج اختبارات لا تنتجها البدائل: نسبة التكاثف، تطابق الفكر والفعل، إثبات الأثر، سلامة commit، وكثافة النقل.

النسخة المتواضعة: MSRAK ليست نواة AGI مثبتة، إنها control architecture قد تحسن أنظمة وكيلة مبنية على نماذج التفكير، وقد تفشل إذا كانت المكونات أو المقاييس لا تضيف قيمة فوق scaffolds أبسط.

المعمارية الموحدة: **MSRAK**

المسار الأعلى: Task/Authority Contract → Governing Question → Active Difference → State & Representation → Vascular Test-Time Routing → Knowledge/Skill/Tool Composition → Thought-Action Binding → Courts & Authorization → World Effect & Readback → Reversible Memory → Higher Question.

الطبقة	العقد التشغيلي
0. العقد	الهدف، النطاق، الصلاحية، الميزانية، الخصوصية، شروط التوقف.
1. السؤال	صياغة سؤال يميز ما يلزم معرفته أو فعله.
2. الفرق	استخراج المتغير الذي سيغير الحكم أو المسار.
3. الحالة/التمثيل	فصل المرصود والمستنتج واختيار تمثيل قابل للاختبار.
4. التوجيه	تخصيص نموذج/أداة/رموز/مسارات حسب القيمة والخطر.
5. التركيب	استرجاع أقل سياق وتأليف HyperSkills وBounded HyperLoops.
6. الربط	تحويل كل فعل إلى عقدة مرتبطة بفرضية ودليل وسلطة.
7. المحاكم	تعريفات، دليل، آلية، أمن، بديل، تكلفة، تفويض.
8. الأثر	قياس تغير خارجي ومقارنته بالمتوقع.
9. الذاكرة	حجر delta ثم canary ثم commit أو rollback.

تعمل الدورة كآلة حالات لا تخطوات إلزامية متساوية. المهمة البسيطة قد تعبر طبقات قليلة؛ والمهمة عالية المخاطر تستدعي verifier مستقلًا، sandbox، وتجربة قابلة للعكس. لا يدل طول المسار على الذكاء. الحاكم الأعلى هو مقدار الفرق الموثق الذي تغير به القرار مع أقل كلفة ومخاطر.

Object model: كل Knowledge Unit يحمل id, claim_type, provenance, active_difference, mechanism, inputs, outputs, permissions, metrics, courts, failure_signature, version, memory_policy. كل انتقال يكتب event append-only يسمح بإعادة البناء.

طبقات النواة 0–4: من العقد إلى التوجيه

0. **Task and Authority Contract**: يمنع الانزلاق من تحليل إلى فعل. يحدد من طلب، ما الهدف، ما الممنوع، ما السقف المالي والزمني، وما الذي يحتاج موافقة. غيابه يجعل النظام fail-closed.

1. **Governing Question Generator**: لا يعيد صياغة الطلب فقط؛ يولد سؤالًا يختبر الحلقة المفقودة. يُقيّم السؤال بعمقه، فائدته الفاصلة، قابلية تكذيبه، وعلاقته بالفعل.

2. **Active Difference Extractor**: يميز السبب عن الشرط، المؤشر عن الآلية، والنجاح الظاهري عن الأثر. المخرج Δ^* هو أصغر فرق متوقع أن يغير التوجيه أو الحكم.

3. **State and Representation Compiler**: يبني World-State Packet: observed, inferred, hypothesized, unknown, constraints, resources. ثم يختار تمثيلًا: نص، جدول، PDDL، graph، كود، أو simulator.

4. **Vascular Test-Time Router**: يمنح الحساب للعقد بحسب $U_i = \alpha\ \text{uncertainty} + \beta\ \text{value} + \gamma\ \text{risk} + \delta\ \text{irreversibility} - \kappa\ \text{cost}$. التوجيه متناثر: يمكن أن يختار جوابًا مباشرًا، تفكيرًا طويلًا، عدة verifier، rollouts، أداة، أو تفويضًا.

صيغة التوجيه

$$\text{Route}^* = \arg\max_r \left[E(\text{DecisionGain}_r \times \text{EvidenceQuality}_r \times \text{Transfer}_r) - \text{Cost}_r - \text{Risk}_r - \text{Redundancy}_r\right]$$

تعمل الصيغة كقائمة تحقق: أي route لا يوضح gain المتوقع أو evidence أو cost لا ينبغي اختياره لمجرد شهرة النموذج.

الحد المنهجي: تقدير uncertainty بواسطة النموذج قد يكون غير معاير. لذلك تسمح النواة بإشارات خارجية: اختلاف العينات، فشل الاختبارات، تعارض المصادر، طول trace، أو عدم استقرار الأدوات. كما يجب اختبار أن التوجيه لا يتعلم دائمًا اختيار النموذج الأكبر بسبب metric منحاز.

تنبؤ طبقي: في المهمات ذات صعوبة مختلطة، يقل متوسط الكلفة مع ثبات النجاح إذا استُعمل التوجيه التكيفي مقارنة بميزانية تفكير ثابتة، بشرط جودة تقدير الصعوبة.

توسيع اختباري وحدود تشغيلية

يُختبر مولد السؤال بمقارنة أسئلته مع تدخلات oracle أو خبراء: هل خفض عدد الفروع؟ هل كشف متغيرًا مخفيًا؟ هل أنتج تجربة تفصل البدائل؟ أما مستخرج الفرق فيُختبر بحذف Δ^* من السياق وقياس تغير القرار. إذا لم يتغير شيء، فالفرق لم يكن حاكمًا أو أن النظام لم يستخدمه فعلًا.

التوجيه الوعائي يجب أن يبقى قابلًا للتفسير التشغيلي: لا يلزم شرح كل حساب داخلي، لكن يلزم تسجيل سبب اختيار route وبدائله والكلفة المتوقعة. هذا السجل يسمح بتحسين router نفسه، وكشف انحياز دائم إلى نموذج أو أداة بعينها.

طبقات النواة 5–9: من التركيب إلى الذاكرة

5. **Minimal Composition Layer**: تسترجع أقل KB كافٍ، وتختار HyperSkills ذات عقود متوافقة، وتبني Bounded HyperLoop بميزانية وتوقف. يمنع semantic redundancy tax نسخ المعرفة نفسها بأسماء مختلفة.

6. **Thought-Action Binding Layer**: ينشئ لكل فعل {claim, evidence, goal, permission, expected_effect, risk, rollback} :tuple. إذا غاب عنصر حاسم، يبقى الفعل اقتراحًا أو dry-run.

7. **Dynamic Courts and Authorization**: تختار المحاكم حسب signature الفشل بدل تشغيل قائمة ثابتة. مثال: مصدر ويب يستدعي injection/evidence؛ كود إنتاجي يستدعي tests/security/rollback؛ قرار عالي الأثر يستدعي independent verifier.

8. **World-Effect and Readback**: يحدد قبل الفعل observable وthreshold وtime window وalternative explanation. بعد الفعل يقارن expected/actual ويعزو الفجوة بحذر.

9. **Reversible Constitutive Memory**: يُكتب Memory Delta في quarantine. يمر باختبارات provenance, transfer, regression, contamination, permission. يُرقّى إلى semantic/procedural memory أو يُلغى.

حالة الآلة

S = ⟨task, authority, q, Δ, world, representation, route, composition, actions, courts, effects, memory⟩ الانتقال τ(S,a) مسموح فقط إذا استوفى preconditions وسجل evidence. هذا يجعل النواة قابلة للتدقيق حتى حين يكون نموذج التفكير مغلقًا؛ لا نحتاج كشف chain-of-thought الخاص، بل نحتاج سجلات قرار وإجراء قابلة للتحقق.

حدود الطبقات: قد تزيد latency والتعقيد، وقد تتحول العقود إلى بيروقراطية شكلية. لذلك يُختبر كل gate بـablation: هل يخفض نوع فشل محدد؟ إن لم يفعل يُدمج أو يحذف. النواة ليست تقديسًا للكثرة؛ هي اقتصاد الحاكمية.

قانون الترقية: لا تُرقّى معرفة أو مهارة بسبب نجاح واحد أو تقييم ذاتي؛ يلزم أثر متكرر، نقل، وعدم تدهور guardrails، مع نسخة وrollback.

توسيع اختباري وحدود تشغيلية

التركيب الأدنى لا يعني تقليلًا أعمى. فهو يحتفظ بالعناصر التي تغطي prerequisites والمخاطر والتعارضات. يستخدم النظام dependency graph ليمنع استرجاع مهارة بلا متطلب أو حلقة بلا stop condition. وعند الحاجة إلى وحدات كثيرة، يُضغط السياق إلى عقود ومؤشرات مع روابط إلى الأصل بدل نسخ الوثائق كاملة.

في الذاكرة، يُفصل نجاح المهمة عن صلاحية القاعدة. قد ينجح الفعل لأسباب ظرفية أو صدفة؛ لذلك لا يتحول outcome إلى causal lesson إلا بعد بديل مضاد أو تكرار. تحفظ الذاكرة الفشل أيضًا، لكن بوصفه نطاقًا وتحذيرًا لا قاعدة منع مطلقة.

الصياغة الصورية وشبه الرياضية

1. قيمة النواة: $K = w1Q + w2D + w3R + w4B + w5V + w6W + w7M - w8C - w9H$. Q جودة السؤال، D الفرق، R التوجيه، B ربط الفكر بالفعل، V التحقق، W أثر العالم، M سلامة الذاكرة، C الكلفة، H الضرر/الضجيج.

2. نسبة تكاثف الاستدلال إلى وكالة: $RAC = VerifiedActionValue / (ReasoningCost + CoordinationCost)$. ترتفع عندما يتحول التفكير إلى فعل صحيح ومأذون ومثبت، لا عندما يطول trace.

3. تطابق الفكر والفعل: $TAB = \Sigma_i support(action_i \mid claims,evidence,goal,authority) / n$. يمكن تقديره بواسطة مراجعة بشرية، قواعد، أو verifier مستقل.

4. إثبات أثر العالم: $WEP = Observability \times Attribution \times Reproducibility \times Safety$. إذا كان أحدها قريبًا من الصفر، لا يعد النجاح إثباتًا قويًا.

5. سلامة الذاكرة الرجوعية: $RMI = Provenance \times Transfer \times RegressionSafety \times Revertibility - ContaminationRisk$.

6. كثافة النقل عبر النماذج: $CMTD = RetainedVerifiedCapability / AdaptationCost$ عند تبديل model/provider مع ثبات contracts.

7. دين الاستدلال: $RD = UnverifiedClaims + RepeatedBranches + UnresolvedContradictions + UnsupportedActions$. الدين لا يعني كل عدم يقين؛ يعني التزامات معرفية لم تُسوَّ قبل الفعل أو الذاكرة.

هذه الصيغ ليست قياسات معيارية بعد. لتصبح مقاييس علمية يجب تعريف rubrics، جمع بيانات بين مقيمين، اختبار الثبات والصدق التنبؤي، وتحديد trade-offs. قد تكون بعض المركبات multiplicative لمنع تعويض الصفر الحرج بدرجات أخرى؛ وقد تكون additive في تطبيقات منخفضة المخاطر. لذلك تقدم الورقة «لغة قياس» لا أرقامًا نهائية.

اختبار الصدق: إذا لم تتنبأ RAC أو TAB أو WEP بالنجاح طويل الأفق والسلامة أفضل من token count أو final-answer score، تُرفض القيمة القياسية للمفاهيم أو يعاد تعريفها.

توسيع اختباري وحدود تشغيلية

يمكن تحويل الصيغ إلى نموذج احتمالي هرمي لاحقًا، لكن البدء بمؤشرات بسيطة أكثر صدقًا. القيمة العلمية الأولى ليست الرقم، بل إجبار التجربة على إظهار المتغيرات التي كانت مخفية: كلفة التنسيق، نسبة الأفعال غير المدعومة، قوة attribution، ومعدل rollback. هذه المتغيرات تمنع اختزال الوكيل في success rate.

تحتاج الأوزان إلى تعلم أو معايرة حسب المجال. في الطب أو المال أو البنية الإنتاجية يرتفع وزن الضرر والرجوع؛ وفي sandbox تعليمي قد يرتفع وزن الاستكشاف. لا توجد λ عالمية قبل بيانات، ولذلك يمنع استخدام MCD كآلة ترتيب قطعية خارج rubric معلن.

النظرية الأولى: نظرية مساعف لتكاثف الاستدلال السيادي إلى وكالة

Moussaif Sovereign Reasoning-to-Agency Condensation Theory (MSRACT)

البذرة والفجوة: خرجت من الفصل بين thinking model وagentic kernel، ومن تكرار فشل الأنظمة عند الانتقال بين تمثيل وفعل أو بين تجربة وذاكرة.

التعريف/الأطروحة: الاستدلال لا يصبح وكالة بمجرد إرفاق أداة، يصبح وكالة عندما تُضغط المسارات في قرار مرتبط بهدف ودليل وسلطة وأثر ورجوع.

الآلية: تتعاقب حالات: استكشاف → تمييز → تمثيل → قرار → فعل. عند كل انتقال تُفقد معلومات ويظهر خطر القفزة. طبقة التكاثف تحافظ على invariants: الهدف، الدليل، القيود، وعدم اليقين.

الصياغة: RAC = V(action verified) / total cognitive and coordination cost.

التنبؤ الفارق: تتنبأ بأن النظم ذات traces أقصر قد تتفوق على traces أطول إذا كان binding والتوجيه أفضل.

الحدود: لا تفسر وحدها جودة المعرفة أو صحة world model؛ وقد تتحول إلى ضغط مفرط يحذف بدائل مهمة.

الاختبار الأدنى: اختبار: قارن نفس النموذج مع/بدون condensation contracts على مهمات طويلة، وثبت الميزانية. قس النجاح، الأفعال غير المسندة، والتوقف الخاطئ.

أقوى اعتراض والنسخة المعدلة

الاعتراض: «التكاثف» إعادة تسمية للقرار أو التخطيط. الجواب: تصبح نظرية مستقلة فقط إذا قيس فقدان invariants بين الفكر والفعل، وإذا تنبأت RAC بالنجاح. النسخة المعدلة تحصر الادعاء في طبقة ربط قابلة للـablation، لا مبدأ شامل للذكاء.

موضعها داخل **MSRAK**: طبقة 6 ومحور التحول من reasoning إلى action.

توسيع اختباري وحدود تشغيلية

تضيف النظرية مفهوم invariants الانتقالية: عناصر يجب ألا تضيع عندما يُضغط الاستدلال، مثل الهدف والقيود والمصدر وعدم اليقين. يمكن بناء test cases تتعمد إدخال plan جميل يحذف قيدًا واحدًا، ثم قياس هل تمنع طبقة التكاثف الفعل. بذلك يتحول المفهوم من استعارة ضغط إلى فرضية عن أخطاء الحدود بين الحالات.

النسخة الأقوى مستقبلًا ستتعلم سياسة التكاثف من trajectories مع labels لما فُقد، لكنها تبدأ بقواعد typed حتى تبقى قابلة للتدقيق ولا تتعلم إخفاء التردد أو المخاطر.

النظرية الثانية: نظرية مساعف للتطابق الحاكمي بين الفكر والفعل

Moussaif Sovereign Thought-Action Congruence Theory (MSTAC)

البذرة والفجوة: خرجت من الفصل بين thinking model وagentic kernel، ومن تكرار فشل الأنظمة عند الانتقال بين تمثيل وفعل أو بين تجربة وذاكرة.

التعريف/الأطروحة: فشل الوكيل كثيرًا لا لأن الخطة مفقودة، بل لأن الفعل المنفذ لا يطابق ما تدعمه الخطة أو يتجاوز الصلاحية.

الآلية: لكل action عقد دعم يربطه بادعاء ودليل وهدف وصلاحية وأثر متوقع. المحاكم تفحص الاتساق السببي والسلطوي قبل التنفيذ، والقراءة الراجعة تقيس الانحراف.

الصياغة: TAB = mean support(action_i | reasoning, evidence, authority).

التنبؤ الفارق: تتنبأ بأن TAB يفسر السلامة والنجاح أفضل من جودة الخطة النصية وحدها.

الحدود: قد يبدو التطابق عاليًا بسبب rationalization؛ لذلك يجب أن يُبنى العقد قبل الفعل ويُراجع مستقلًا.

الاختبار الأدنى: اختبار: حقن خطط صحيحة مع actions منحرفة، وقياس قدرة الطبقة على الإيقاف مقارنة بverifier عام.

أقوى اعتراض والنسخة المعدلة

الاعتراض: لا يمكن معرفة الفكر الحقيقي من trace. الجواب: لا تدعي النظرية ذلك؛ تقيس التطابق بين سجل القرار المعلن والأدلة والعقد والفعل، وهو تدقيق تشغيلي لا كشف للحساب الداخلي. إذا لم يسبق السجل الفعل يفقد قيمته.

موضعها داخل **MSRAK**: طبقة 6–7 ومحور السلطة والتدقيق.

توسيع اختباري وحدود تشغيلية

التطابق ليس مساواة لغوية بين الخطة والفعل؛ قد يختلف التعبير بينما تبقى العلاقة السببية صحيحة. لذلك يبنى graph يربط action node بالأهداف والأدلة والشروط، ويقاس وجود مسار دعم صالح. كما تسجل الأدلة التي تعارض الفعل، لأن الاكتفاء بأسباب مؤيدة يسمح بالتبرير الانتقائي.

عند فشل الفعل، تحدد النظرية هل الخلل في reasoning أو binding أو execution أو world model. هذا التشخيص يمنع إعادة تدريب النموذج على مشكلة كانت في صلاحية الأداة أو تحديث الحالة.

النظرية الثالثة: نظرية مساعف للفرق الحاكمي العابر للنماذج

Moussaif Cross-Model Sovereign Difference Theory (MCSDT)

البذرة والفجوة: خرجت من الفصل بين thinking model وagentic kernel، ومن تكرار فشل الأنظمة عند الانتقال بين تمثيل وفعل أو بين تجربة وذاكرة.

التعريف/الأطروحة: ما ينتقل بين النماذج ليس النص الكامل ولا الأسلوب، بل الفرق الحاكم وعقده التشغيلي: متى يظهر، ماذا يغير، وكيف يُختبر.

الآلية: يُقطر trajectory إلى Δ-contract مستقل عن provider؛ ثم يُشغل على نموذج آخر مع adapter رقيق. ما لا يحتفظ بوظيفته يُعد model-bound لا knowledge kernel.

الصياغة: CMTD = retained verified capability / adaptation cost.

التنبؤ الفارق: تتنبأ بأن تخزين deltas والعقود يحقق نقلًا أعلى لكل token من تخزين transcripts كاملة.

الحدود: قد تفشل في قدرات متجذرة في أوزان نموذج بعينه أو واجهة أداة خاصة.

الاختبار الأدنى: اختبار: تبادل ثلاث نوى بين نماذج مفتوحة ومغلقة، مع holdout domains وقياس adaptation وregression.

أقوى اعتراض والنسخة المعدلة

الاعتراض: النقل يأتي من prompt engineering. الجواب: يُقبل ذلك إذا حقق prompt بسيط النتيجة نفسها. النظرية تبقى فقط إن تفوقت عقود Δ على prompts متساوية الحجم عبر نماذج ومجالات متعددة.

موضعها داخل **MSRAK**: طبقة 2–5 ومحور النقل والضغط.

توسيع اختباري وحدود تشغيلية

عقد Δ يتضمن trigger وmechanism وexpected transformation وboundary وtest وfailure signature. لا يحمل صياغة الحل الكامل، ولذلك يفترض أن يتيح للنموذج الجديد إعادة توليد السلوك بأسلوبه. نجاح النقل يعني احتفاظ الوظيفة لا تطابق النص. ويُقارن مع few-shot examples وretrieval summaries متساوية الحجم.

تسمح النظرية أيضًا بتحديد ما لا ينتقل: مهارات تعتمد على tokenizer أو أوزان أو affordance خاصة. هذا الفشل ليس هدمًا كاملًا، بل يحدد طبقة adapter أو الحاجة إلى تدريب فعلي.

النظرية الرابعة: نظرية مساعف للإثبات العالمي المتدرج للنواة

Moussaif Graduated World-Effect Proof Theory (MGWEPT)

البذرة والفجوة: خرجت من الفصل بين thinking model وagentic kernel، ومن تكرار فشل الأنظمة عند الانتقال بين تمثيل وفعل أو بين تجربة وذاكرة.

التعريف/الأطروحة: لا نُثبت قدرة النواة ببلاغة المخرج أو benchmark داخلي فقط؛ نُثبت تدريجيًا من اختبار محلي إلى أثر بيئي ثم نقل وتكرار مستقل.

الآلية: سلم الإثبات: syntax → unit test → task outcome → external observable → counterfactual attribution → transfer → independent replication. لا تُرقى القدرة أعلى من أعلى درجة اجتازتها.

الصياغة: WEP_level ∈ {0..6}; promotion requires guardrail non-regression.

التنبؤ الفارق: نتنبأ بأن كثيرًا من «نجاحات الوكلاء» ستنخفض رتبتها عند اشتراط attribution والرجوع والنقل.

الحدود: الأثر الخارجي قد يكون مكلفًا أو أخلاقيًا غير ممكن؛ يلزم simulation أو sandbox بدرجة أدنى مصرح بها.

الاختبار الأدنى: اختبار: أعد تقييم مهارات agent شائعة بسلم الإثبات وقارن التصنيف بالنجاح النهائي التقليدي.

أقوى اعتراض والنسخة المعدلة

الاعتراض: العالم معقد ولا يمكن attribution دقيق. الجواب: لذلك الإثبات متدرج ولا يساوي غياب الدرجة العليا بطلان القدرة؛ لكنه يمنع نقل ادعاء محلي إلى عمومية دون سند.

موضعها داخل **MSRAK**: طبقة 8 ومحور التقييم والترقية.

توسيع اختباري وحدود تشغيلية

سلم الإثبات يمنع القفزة من benchmark إلى ادعاء عالم حقيقي. فقد يجتاز الوكيل اختبارًا بسبب تسرب بيانات أو مرونة evaluator، بينما يفشل عند اختلاف الواجهة. كل درجة أعلى تتضمن الأدنى ولا تلغيها، وتسمح للباحث بنشر قدرة محلية بصدق بدل انتظار إثبات شامل أو الادعاء به.

يجب أن يسجل WEPL النتائج السلبية والتأثيرات الجانبية، وأن يميز عدم الملاحظة عن عدم الأثر. فغياب signal قد يعني أداة قياس ضعيفة، لا نجاحًا ولا فشلًا.

النظرية الخامسة: نظرية مساعف للذاكرة التكوينية الرجوعية

Moussaif Reversible Constitutive Memory Theory (MRCMT)

البذرة والفجوة: خرجت من الفصل بين thinking model وagentic kernel، ومن تكرار فشل الأنظمة عند الانتقال بين تمثيل وفعل أو بين تجربة وذاكرة.

التعريف/الأطروحة: ذاكرة الوكيل لا تسترجع الماضي فقط، إنها تكوّن سياساته المستقبلية. لذلك كل commit تعديل دستوري مصغر يجب أن يكون قابلاً للرجوع.

الآلية: يفصل النظام memory proposal عن approved memory. تُحفظ التعارضات، ويُلزم كل patch بسياق وصلاحية واختبار ونقطة revert. الترقية تدريجية من episodic إلى procedural.

الصياغة: RMI = provenance × transfer × safety × revertibility − contamination.

التنبؤ الفارق: تتنبأ بأن quarantine وcanary يقللان الانحدار طويل الأجل حتى إن زادا latency قصيرًا.

الحدود: الرجوع قد لا يعكس آثارًا خارجية حدثت بالفعل، وقد تصبح التبعيات بين patches معقدة.

الاختبار الأدنى: اختبار: أدخل خبرات صحيحة وخاطئة ومسمومة، وقارن الاحتفاظ والنقل والتعافي بين overwrite وversioned memory.

أقوى اعتراض والنسخة المعدلة

الاعتراض: versioning تقنية برمجية لا نظرية معرفية. الجواب: تصبح نظرية حين يثبت أن بنية الذاكرة تغير التعميم والسلوك، وأن قابلية الرجوع متغير سببي في التعلم المستمر، لا مجرد ميزة صيانة.

موضعها داخل **MSRAK**: طبقة 9 ومحور التعلم وإعادة البناء.

توسيع اختباري وحدود تشغيلية

الذاكرة التكوينية تربط التعلم بالهوية التشغيلية: بعض patches تغير استراتيجية عامة أو حدود سلطة، ولذلك تحتاج مراجعة أشد من حقيقة صغيرة. يصنف RMC patch حسب blast radius، ويطلب اختبارات أوسع كلما ارتفع. كما يمنع التعديل المتسلسل من طمس أصل القرار عبر dependency graph.

التحدي الأكبر هو تراكم الفروع والنسخ. لذلك تستخدم الذاكرة compaction دوريًا يحفظ تاريخًا قابلًا للاسترجاع، ويمنع دمج نسختين إذا اختلف شرط حد أو verdict أمني.

الأدوات الجديدة 1–5: الاستخراج والتوجيه

الأداة	النقص الذي تعالجه	المدخلات	المخرجات
1. سجل نسب النواة Moussaif Kernel Genealogy Ledger (MKGL)	إزالة ازدواج الأسماء وربط كل مفهوم بموضع وآلية واختبار.	اسم/تعريف/مدخلات/مخرجات/مالك/نسخة/فشل.	سجل canonical + aliases + قرار دمج.
2. مستخرج النواة الاستدلالية Reasoning Kernel Extractor (RKE)	تحويل trajectory أو ورقة إلى عقد تشغيل قابلة للنقل.	مسارات، أدوات، نجاحات وفشل.	Δ-contract + prerequisites + test.
3. بوابة الفرق السيادي Sovereign Difference Gate (SDG)	منع التوجيه على فرق زخرفي أو غير فاصل.	سؤال، بدائل، قرار محتمل.	Δ* وترتيب تأثيره.
4. موجه الحساب الوعائي Vascular Test-Time Compute Router (VTCR)	تخصيص التفكير حسب الصعوبة والقيمة والخطر.	uncertainty, risk, cost, models.	route وbudget وstop.
5. مصفوفة تكاثف الوكالة Agency Condensation Matrix (ACM)	قياس ما ضاع بين الاستدلال والفعل.	claims, plan, actions, effects.	RAC ومواطن الانفصال.

بروتوكول الاستعمال

1. سجل نسب النواة: تبدأ بتحديد mission id والسلطة، ثم تجمع اسم/تعريف/مدخلات/مخرجات/مالك/نسخة/فشل.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو سجل canonical + aliases + قرار دمج.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

2. مستخرج النواة الاستدلالية: تبدأ بتحديد mission id والسلطة، ثم تجمع مسارات، أدوات، نجاحات وفشل.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو Δ-contract + prerequisites + test.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

3. بوابة الفرق السيادي: تبدأ بتحديد mission id والسلطة، ثم تجمع سؤال، بدائل، قرار محتمل.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو Δ* وترتيب تأثيره.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

4. موجه الحساب الوعائي: تبدأ بتحديد mission id والسلطة، ثم تجمع uncertainty, risk, cost, models.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو route وbudget وstop.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

5. مصفوفة تكاثف الوكالة: تبدأ بتحديد mission id والسلطة، ثم تجمع claims, plan, actions, effects.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو RAC ومواطن الانفصال.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

قاعدة التطوير: كل أداة تصبح HyperSkill فقط بعد أن ينفذها عامل جديد من عقدها دون سياق خفي، وتحقق أثرًا متكررًا، ولا ترفع التكلفة أو الأخطاء الجانبية بصورة غير مقبولة.

الأدوات الجديدة 6–10: الربط والسيادة

الأداة	النقص الذي تعالجه	المدخلات	المخرجات
6. مترجم ربط الفكر بالفعل Thought-Action Binding Compiler (TABC)	إنشاء عقود دعم قبل الأدوات أو التنفيذ.	action proposal + evidence + authority.	typed action contract.
7. مجدول محاكم الهدم Dynamic Destruction Court Scheduler (DDCS)	اختيار المحاكم حسب المخاطر بدل قائمة ثابتة.	failure signatures + impact.	court plan + independent verifier.
8. مقطر النواة العابر للنماذج Cross-Model Kernel Distiller (CMKD)	ضغط الخبرة إلى فرق وعقد مستقلين عن provider.	successful/failed trajectories.	portable kernel patch.
9. جدار سيادة السياق Context Sovereignty Firewall (CSF)	منع تسرب هوية أو تعليمات أو ذاكرة بين المشاريع.	context packets + provenance.	allow/deny/sanitize/quarantine.
10. مقياس دين الاستدلال Reasoning Debt Meter (RDM)	كشف الادعاءات والفروع والتناقضات غير المسواة.	trace summary + evidence ledger.	debt score + settlement plan.

بروتوكول الاستعمال

6. مترجم ربط الفكر بالفعل: تبدأ بتحديد mission id والسلطة، ثم تجمع action proposal + evidence + authority.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو typed action contract.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

7. مجدول محاكم الهدم: تبدأ بتحديد mission id والسلطة، ثم تجمع failure signatures + impact.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو court plan + independent verifier.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

8. مقطر النواة العابر للنماذج: تبدأ بتحديد mission id والسلطة، ثم تجمع successful/failed trajectories.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو portable kernel patch.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

9. جدار سيادة السياق: تبدأ بتحديد mission id والسلطة، ثم تجمع context packets + provenance.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو allow/deny/sanitize/quarantine.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

10. مقياس دين الاستدلال: تبدأ بتحديد mission id والسلطة، ثم تجمع trace summary + evidence ledger.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو debt score + settlement plan.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

قاعدة التطوير: كل أداة تصبح HyperSkill فقط بعد أن ينفذها عامل جديد من عقدها دون سياق خفي، وتحقق أثرًا متكررًا، ولا ترفع التكلفة أو الأخطاء الجانبية بصورة غير مقبولة.

الأدوات الجديدة 11–15: الإثبات والتطور

الأداة	النقص الذي تعالجه	المدخلات	المخرجات
11. دفتر إثبات أثر العالم World-Effect Proof Ledger (WEPL)	ربط كل قدرة بمستوى إثبات خارجي.	.expected/actual, tests, alternatives	.WEP level + verdict
12. مثبّت الذاكرة الرجوعي Reversible Memory Committer (RMC)	حجر واختبار وترقية أو إلغاء .memory deltas	.proposed patch + provenance	.versioned commit/revert
13. مبطل النواة المضاد Counterfactual Kernel Ablator (CKA)	قياس القيمة السببية لكل طبقة.	.baseline + component + task suite	.effect size + keep/merge/delete
14. مصفوفة جاهزية AGI Fattah Fattah AGI Readiness Matrix (FARM)	فصل الفتح الوظيفي عن ادعاء .AGI	.opening, transfer, agency, safety, memory	profile لا label ثنائي.
15. مختبر تطور النواة Kernel Evolution Sandbox (KES)	تجربة patches والنماذج والمهارات دون تلويث الإنتاج.	.candidate kernels + canaries	.promotion/deprecation evidence

بروتوكول الاستعمال

11. دفتر إثبات أثر العالم: تبدأ بتحديد mission id والسلطة، ثم تجمع expected/actual, tests, alternatives.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو WEP level + verdict.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

12. مثبّت الذاكرة الرجوعي: تبدأ بتحديد mission id والسلطة، ثم تجمع proposed patch + provenance.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو versioned commit/revert.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

13. مبطل النواة المضاد: تبدأ بتحديد mission id والسلطة، ثم تجمع baseline + component + task suite.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو effect size + keep/merge/delete.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

14. مصفوفة جاهزية AGI Fattah: تبدأ بتحديد mission id والسلطة، ثم تجمع opening, transfer, agency, safety, memory.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو profile لا label ثنائي.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

15. مختبر تطور النواة: تبدأ بتحديد mission id والسلطة، ثم تجمع candidate kernels + canaries.، وتمنع الترقية إذا غاب دليل أو rollback. المخرج المقبول هو promotion/deprecation evidence.. أخطر سوء استعمال هو تحويل الدرجة أو السجل إلى بديل عن التجربة؛ لذلك تُراجع الأداة بالـCKA وWEPL.

قاعدة التطوير: كل أداة تصبح HyperSkill فقط بعد أن ينفذها عامل جديد من عقدها دون سياق خفي، وتحقق أثرًا متكررًا، ولا ترفع التكلفة أو الأخطاء الجانبية بصورة غير مقبولة.

عشرون مفهومًا جديدًا أو معاد ضبطه

#	المفهوم	التعريف التشغيلي
1	نواة التفكير السيادية	سياسة ميتا تختار عمليات التفكير والفعل والذاكرة تحت عقد.
2	تكاثف الاستدلال	تحويل فضاء المسارات إلى فعل موثق مع حفظ invariants.
3	الفرق الحاكم	تمييز يغير الحكم أو التوجيه أو الفعل.
4	السؤال الحاكم	سؤال يعيد ترتيب فضاء البحث ويولد اختبارًا فاصلًا.
5	الحساب الوعائي	ميزانية متكيفة تتدفق إلى العقد الحرجة.
6	دين الاستدلال	التزامات معرفية غير متحققة قبل الفعل أو الحفظ.
7	تطابق الفكر والفعل	درجة دعم الفعل بسجل قرار سابق ودليل وسلطة.
8	أثر العالم المثبت	تغير خارجي مرصود ذو attribution وحدود.
9	سلم إثبات النواة	درجات من syntax إلى replication.
10	الذاكرة التكوينية	ذاكرة تغير السياسات المستقبلية لا مجرد الاسترجاع.
11	الذاكرة الرجوعية	patch موثق قابل للاختبار والإلغاء.
12	كثافة النقل	قدرة محفوظة لكل كلفة تكييف بين النماذج.
13	ائتلاف كافٍ أدنى	أقل عدد نماذج/وكلاء يحقق الوظيفة كاملة.
14	السياق السيادي	حزمة هدف وهوية وسلطة لا تخضع لسياق خارجي غير موثوق.
15	محكمة ديناميكية	مجموعة اختبارات تُختار من signature الفشل.
16	الفتح المنضبط	توسيع فضاء السؤال بواسطة فرق واختبار لا تسمية فقط.
17	حالة عالم كافية	أقل تمثيل للعالم يسمح بفعل آمن قابل للتقييم.
18	عقد Δ	تمثيل محمول لفرق: trigger, mechanism, effect, test, boundary.
19	ترقية القدرة	انتقال من اقتراح إلى مهارة بعد أثر ونقل وعدم انحدار.
20	سيادة الرجوع	أولوية الأفعال والتعلم القابلين للإلغاء عند عدم اليقين.

خطر اصطلاحي: أي مفهوم لا ينتج قرارًا أو أداة أو اختبارًا يخرج من سجل النظرية إلى القاموس. «السيادة» لا تعني استقلالًا سياسيًا أو وعيًا ذاتيًا؛ تعني أولوية عقد المهمة والدليل والسلطة داخل النظام.

المقاييس والمؤشرات الأولية

المقياس	ما يقيسه	المؤشرات الفرعية
KSI — مؤشر سيادة النواة	حفظ الهدف والسلطة والدليل عبر transitions.	task checksum، permission compliance، context leakage، trace completeness.
RACR — نسبة تكاثف الاستدلال	قيمة الفعل المتحقق لكل كلفة تفكير وتنسيق.	verified action value / tokens+calls+latency.
TACI — مؤشر تطابق الفكر والفعل	دعم الأفعال بالعقود والأدلة السابقة.	pre-action binding، causal support، authority match.
CMTD — كثافة النقل العابر للنماذج	قدرة محتفظ بها لكل كلفة تكييف.	performance retention، adapter size، prompt delta، regressions.
WEP — إثبات أثر العالم	قوة الدليل الخارجي على نجاح القدرة.	observability، attribution، reproducibility، safety.
DY — مردود محكمة الهدم	مقدار الخطأ المكتشف أو التحسن الناتج لكل كلفة محكمة.	caught failures، false blocks، repair value، latency.
RLI — سلامة التعلم الرجوعي	جودة provenance والنقل والرجوع وعدم التلوث.	commit tests، rollback success، contamination، dependency integrity.

طريقة التقدير

يبدأ كل مقياس rubric من 0 إلى 4 بدل دقة زائفة. تُجمع درجات بشرية وآلية ويُحسب اتفاق المقيمين. لا يُدمج المقياس في رقم AGI واحد قبل اختبار الصدق البنائي والتنبؤي. يمكن استعمال Pareto frontier بدل مجموع مرجح: نظام أعلى نجاحًا لكنه أقل سلامة لا يُعلن فائزًا مطلقًا.

بيانات التحسين

نحتاج trajectories كاملة، عقود قبل الفعل، سجلات أدوات، outcomes خارجية، تكاليف، commits، permissions وrollbacks. يجب فصل مجموعة تطوير المقاييس عن benchmark النهائي، ومنع evaluator من رؤية label المطلوب. تُختبر المقاييس عبر نماذج ومجالات وواجهات مختلفة لتجنب قياس أسلوب provider بعينه.

معيار إسقاط المقياس: إذا لم يضف قدرة تفسيرية أو تنبؤية فوق success rate وcost وsafety metrics القياسية، يُحذف أو يخفض إلى أداة تشخيصية.

توسيع اختباري وحدود تشغيلية

تُختبر العلاقة بين المقاييس قبل دمجها. قد ترتفع RACR بسبب تقليل التحقق، فتتدهور WEP أو RLI؛ وقد يرتفع KSI عبر رفض معظم الأفعال، فينخفض value. لهذا لا يكافأ النظام على رقم منفرد، بل على جبهة Pareto وحدود دنيا للحماية. كما يعلن كل benchmark وزن المخاطر بدل إخفائه في average.

ينبغي نشر أمثلة معايرة مع كل rubric: حالات 0 و2 و4، وحدود الاختلاف المقبول بين المقيمين. ويمكن استعمال نماذج قاضية، لكن لا تُعامل كحقيقة؛ تُقارن بالبشر وباختبارات قابلة للتنفيذ.

الفرضيات البحثية: القوية والمتوسطة

قوية: H1: التوجيه التكيفي للحساب يخفض الكلفة مع ثبات النجاح في حزمة صعوبة مختلطة.

قوية: H2: thought-action binding المسبق يقلل الأفعال غير المسندة أو غير المأذونة.

قوية: H3: world-effect gate يخفض ترقية المهارات ذات النجاح الوهمي.

قوية: H4: ذاكرة versioned مع rollback تتعافى من poisoning أفضل من overwrite.

قوية: H5: أقل ائتلاف كافٍ يتفوق في الكلفة والتماسك على multi-agent ثابت العدد.

متوسطة: H6: RACR تتنبأ بالنجاح طويل الأفق أفضل من طول chain-of-thought.

متوسطة: H7: CMTD أعلى لعقود Δ من transcripts كاملة عند تبديل النماذج.

متوسطة: H8: المحاكم الديناميكية تحقق مردود هدم أعلى من تشغيل جميع المحاكم دائمًا.

متوسطة: H9: context sovereignty firewall يخفض cross-project leakage دون خسارة كبيرة في الاسترجاع.

متوسطة: H10: حفظ الأدلة المضادة مع الذاكرة يقلل الثقة الزائدة بعد نجاحات متكررة.

متوسطة: H11: تمثيل الحالة الصريح يقلل premature termination في المهمات الطويلة.

طريقة الاختبار العامة: factorial ablation على طبقات MSRAK، نماذج أساس متعددة، ومهام قصيرة/طويلة ومنخفضة/عالية المخاطر. يجب تثبيت token budget وعدد tool calls قدر الإمكان، وإظهار confidence intervals، وتحليل حالات الفشل بدل الاكتفاء بالمتوسط.

توسيع اختباري وحدود تشغيلية

ترتب الفرضيات بحسب قربها من القياس لا بحسب أهميتها الفلسفية. تبدأ التجارب بطبقات ذات تدخل واضح مثل TABC وRMC، ثم تنتقل إلى مفاهيم أعلى مثل opening intelligence. هذا يمنع أن يفشل البرنامج كله لأن مفهومًا واسعًا لم يُعرّف بعد، ويتيح تراكم نتائج صغيرة تعيد تشكيل النظرية.

كل فرضية تحتاج preregistration للمتغير الأساسي ومعيار النجاح والفشل. كما يجب تحليل heterogeneity: قد تنجح الذاكرة الرجوعية في coding agents وتفشل في البحث المفتوح، أو يفيد التوجيه التكيفي في مسائل متوسطة فقط.

الفرضيات الضعيفة والتخمينية والتنبؤات الفارقة

ضعيفة: H12: السؤال الحاكم يزيد تنوع المسارات المفيدة دون زيادة كبيرة في الضجيج.

ضعيفة: H13: reasoning debt يتنبأ بانهيار لاحق حتى عند نجاح مبكر.

ضعيفة: H14: سلم WEP يعيد ترتيب تقييم قدرات الوكلاء مقارنة بالـbenchmark النهائي.

تخمينية: H15: تراكم نوى قابلة للنقل يولد عمومية وظيفية أسرع من توسيع prompt أو context وحده.

تخمينية: H16: سيادة الرجوع توكالة أوسع بأمان أعلى من سياسة منع موحدة.

تخمينية: H17: الفرق الحاكم يمكن أن يكون unit فعالة للتدريب والتفضيل عبر مجالات متعددة.

تخمينية: H18: أنظمة تفتح أسئلة واختبارات جديدة تحقق transfer أقوى من أنظمة تحسن الإجابات فقط.

تسعة تنبؤات فارقة

P1: عند ثبات الميزانية، ستتفوق MSRAK أكثر كلما طال الأفق وتعددت الأدوات، لا في الأسئلة القصيرة.

P2: تعطيل TABC سيزيد فجوة TAB حتى إن بقيت جودة الخطط النصية مرتفعة.

P3: تعطيل WEPL سيرفع عدد memory promotions ثم يزيد الانحدار بعد عدة حلقات.

P4: CMKD سينقل أداءً أكبر من transcript retrieval عند نموذج مختلف، لكن ليس في المهارات weight-bound.

P5: VTCR سيوفر حسابًا في المهام السهلة ويزيده في المتوسطة؛ عند المعقدة جدًا قد يختار solver أو امتناعًا بدل trace أطول.

P6: RDM المرتفع سيسبق نسبة أعلى من التناقض أو التوقف الخاطئ في النصف الثاني من المهمة.

P7: DDCS سيقلل latency مقارنة بكل المحاكم، مع زيادة طفيفة محتملة في misses يجب قياسها.

P8: CSF سيمنع تعليمات ذاكرة ملوثة حتى لو استرجعها RAG كتشابه عالٍ.

P9: FARM سيصنف بعض الأنظمة قوية في reasoning وضعيفة في agency أو memory، رافضًا label AGI أحاديًا.

ما يهدم البرنامج: إذا لم تظهر تفاعلات مع طول الأفق أو النقل، أو كانت المكاسب كلها مفسرة بزيادة الحساب/scaffolding، أو أخفقت المقاييس في التنبؤ، تتحول MSRAK إلى مجموعة ممارسات هندسية لا نظرية موحدة.

توسيع اختباري وحدود تشغيلية

التنبؤ الفارق لا يطلب مجرد تحسن، يحدد أين يجب أن يظهر. إذا تفوقت MSRAK في الأسئلة القصيرة فقط، فهذا يناقض ادعاء معالجة الأفق الطويل. وإذا لم يزد أثر TABC عند ارتفاع صلاحية الأدوات أو عدم قابليتها للرجوع، فالمفهوم لا يلتقط الخطر المقصود. هذه التفاعلات أهم من فرق متوسط صغير.

تحتاج الفرضيات التخمينية إلى مؤشرات وسيطة. فقبل ادعاء عمومية ناتجة عن النوى، يقاس معدل اكتساب task family جديدة، عدد التعديلات اللازمة، وقدرة النظام على اقتراح اختبار صحيح في مجال غير مألوف.

بروتوكولات الاختبار والAblation

الحزمة	التصميم	المخرجات
A. Long-Horizon Tools	مهام OSWorld/Terminal/ويب أو بيئة محلية قابلة للتكرار.	success، TACI، cost، unsafe actions، rollback.
B. Cross-Model Transfer	نفس kernel contracts فوق 3 نماذج مختلفة.	CMTD، adaptation cost، regression.
C. Memory Poisoning	حقن خبرات صحيحة وخاطئة وخصمة.	RLI، attack success، recovery time.
D. Compute Routing	صعوبة متدرجة مع budget ثابت/تكيفي.	accuracy-cost curve، overthinking، abstention.
E. World Effect	مهمات لها observable خارجي أو simulator.	WEP level، attribution، false success.
F. Court Value	لا محكمة/كل المحاكم/DDCS.	DY، false blocks، latency، harm caught.
G. Thought-Action	خطط وأدلة مع actions متطابقة/منحرفة.	TACI، permission violations.
H. Opening Intelligence	مشكلات ناقصة تتطلب سؤالًا أو تجربة جديدة.	novel test utility، transfer، noise.

خطوط الأساس

Direct reasoning model؛ ReAct؛ Reflexion؛ ToT/GoT؛ modular agent with memory؛ fixed multi-agent؛ وMSRAK ablations. يجب توحيد model access والوقت وعدد المحاولات أو الإبلاغ عن فروقها. تُجرى اختبارات contamination وتُحفظ holdouts لا يراها مطورو النواة.

معايير قبول

لا تُقبل النظرية لأن success ارتفع وحده. يلزم ألا تتدهور السلامة أو الكلفة أو المعايرة بصورة غير مقبولة. يفضل تحليل Pareto. يجب نشر negative trajectories، results مختصرة قابلة للتدقيق، ومواصفات contracts دون كشف أسرار أو chain-of-thought خاصة.

اختبار المشغل: يعطى العقد لمشغل/agent جديد لا يعرف سياق المشروع؛ إن لم يستطع تشغيله أو تفسير توقفه فالHyperSkill غير مكتملة.

هندسة تدريب **MOUSSAIF9B** حول النواة

MOUSSAIF9B هنا اسم معماري للمشروع والنواة. إذا كان النموذج التجريبي 1B فيجب تسميته MOUSSAIF9B-1B أو «الإصدار التجريبي 1B»، ولا يصبح 9B وصفًا للمعلمات إلا بعد تدريب نموذج فعلي بتسعة مليارات. تقترح الورقة مسارين: overlay/runtime فوق نماذج مغلقة أو مفتوحة، وتدريب مصرح به لنموذج مفتوح عبر 100 برنامج pretraining و100 post-training.

المرحلة	المحتوى	الحاكمية
Pretraining	تمييزات وكيانات وعلاقات؛ causal/mechanism؛ text؛ code/tools schemas؛ multilingual; provenance; state representations.	عدم تدريب ادعاءات المشروع كحقائق؛ فصل synthetic/source.
Continued pretraining	KB-as-OS، contracts، graphs، failure signatures، العالم/الذاكرة.	dedup، contamination splits، copyright/privacy.
SFT	تعليمات + evidence packet + concise rationale + plan + tool trace + critique + answer.	لا استنساخ hidden CoT؛ تدريب سجلات قرار موجزة.
Preference/RLVR	تفضيل الفروق الصحيحة، التحقق، الامتناع، tool success، rollback.	منع reward hacking و metric substitution.
Process verification	تصنيف خطوات/عقود/أفعال، مع أدوات أو proofs حيث يمكن.	verifier ensembles و calibration.
Agent runtime post-training	Reflexion patches، HyperSkill promotion، memory deltas.	quarantine، canary، holdout، reversible commits.

هدف التدريب ليس جعل النموذج يذكر MCD أو HAKIM-MCD، بل اكتساب السلوك: استخراج الفرق، اختيار تمثيل، توجيه الحساب، التحقق، احترام السلطة، قياس الأثر، وتصحيح الذاكرة. أي dataset تكرر أسماء النظريات دون trajectories وأخطاء ونتائج سيخلق أسلوبًا لا نواة.

قانون التعميم: كل برنامج تدريبي يجب أن يملك holdout من عائلة مهمة أخرى، واختبارًا ضد shortcut، ومقياسًا للاحتفاظ بالقدرات السابقة، وإجراء rollback للمحول أو checkpoint.

Hermes/Codex/Claude/Cursor/Ollama عبر Deployment, Runtime

تُثبت النواة كطبقة مستقلة عن adapters/، courts/، traces/، memory/، loops/، skills/، kernels/، registry/، provider: لا يُحقن corpus كله في context؛ يُحمّل Boot Kernel صغير ثم يسترجع أقل وحدات لازمة. هذا يطبق minimum sufficient context ويمنع أن تتحول المعرفة إلى ضجيج.

المكوّن	الدور داخل MSRAK
Hermes	Control plane: العقد، التوجيه، الجدولة، الذاكرة، الصلاحيات، observability.
Codex/Claude Code/Cursor	تنفيذ هندسي داخل repository مع tests وdiff وrollback.
Ollama/local	مهام خصوصية، تصنيف، استخراج، أو reasoning منخفض الكلفة حسب القدرات.
Frontier reasoning model	العقد عالية التعقيد أو verifier مستقل عند الحاجة.
Tools/MCP/APIs	أفعال typed داخل sandbox وأقل صلاحية، مع فصل محتوى الأداة عن التعليمات.

مسار التشغيل

Mission → Contract → SDG → World State → VTCR → retrieve Kernel/Skill → construct bounded loop → TABC → DDCS → execute minimal reversible action → WEPL → RMC → higher question. كل transition يحمل verdict، permission، evidence، cost، model/tool route، correlation id.

النشر يفشل مغلقًا عند غياب الدستور أو الصلاحية أو verifier في الأثر المتوسط/العالي أو stop conditions أو provenance. الأفعال الإنتاجية أو المالية أو العامة أو الخصوصية الحساسة تبدأ dry-run حتى تفويض صريح. لا تُحفظ credentials في Markdown أو prompts أو logs.

الحياد العابر للمزود: تتغير adapters وبطاقات القدرات بتاريخها؛ تبقى contracts والقياسات والمحاكم مستقرة. لا يثبت اسم نموذج في النظرية لأنه متغير زمني وتجاري.

توسيع اختباري وحدود تشغيلية

يُحفظ لكل provider Capability Card مؤرخة تشمل failure patterns، modalities، cost، latency، tool calling، context، وقيود البيانات. يختار router بطاقة لا اسمًا تسويقيًا. وعند تغير النموذج تعاد canaries قبل الإنتاج، لأن ثبات API لا يعني ثبات السلوك.

النشر المرحلي يمر من offline replay إلى sandbox ثم staging ثم production محدود. لا يسمح للذاكرة المستخرجة من staging أن تدخل production قبل مقارنة النتائج، وتُربط كل نسخة kernel بـmanifest وhash وmigration plan.

التنسيق متعدد النماذج والوكلاء

تعارض MSRAK فكرة أن العدد الأكبر من الوكلاء يعني ذكاء أعلى. تبدأ بـRoot Reasoner واحد. يضاف specialist إذا قدم قدرة فريدة قابلة للقياس: مجال، أداة، لغة، تحقق، أو كلفة. يضاف adversarial verifier عندما يكون أثر الخطأ مرتفعًا أو توجد إشارة عدم استقرار. يسمى ذلك Minimal Sufficient Coalition.

الدور	عقده
Root Reasoner	يحفظ السؤال والعقد وحالة العالم.
Specialist	ينفذ تحويلًا محددًا بعقد، لا يعيد حل المهمة كلها.
Tool Agent	يتعامل مع API/GUI داخل صلاحية محدودة.
Verifier	يفحص مخرجًا أو عقد فعل بمعيار مستقل.
Court Agent	يبني أقوى بديل أو هجوم محدد.
Synthesizer	يدمج الدلائل والاختلافات، لا يصوت بالأغلبية فقط.
Memory Curator	يقترح delta ويمنع commit غير موثق.

التقارب يفحص ثلاثة مستويات: action convergence، mechanism convergence، claim convergence. قد تتفق النماذج في الجواب وتختلف في الآلية أو الفعل، لا يُخفى الاختلاف بالتصويت. يُفرض Semantic Redundancy Tax على الادعاءات والأدلة والخطط المكررة. ويمكن للتنوع أن يكون مقصودًا عند uncertainty العالي، لكنه ينتهي بقرار يوضح البدائل لا بمزيج غامض.

خطأ التنسيق الشائع هو انتشار المسؤولية: كل agent يفترض أن غيره تحقق. لذلك لكل artefact owner واحد وverifier مستقل عند الحاجة. كما تمنع النواة التفاوض على السلطة بين agents؛ الصلاحية خارجية وثابتة في العقد.

اختبار الائتلاف: أزل كل agent بالتتابع. إذا لم ينخفض مقياس هدف أو يزيد فشل محدد، يُحذف الدور أو يدمج. هذه محكمة مضادة لتضخم multi-agent.

توسيع اختباري وحدود تشغيلية

يفصل orchestration بين التنوع والاستقلال. قد تستخدم عدة عينات من النموذج نفسه لتقدير عدم الاستقرار، أو نماذج مختلفة لتغطية تحيزات متباينة. لا يفترض أن اختلاف المزود يعني استقلال الأدلة، فقد تشترك النماذج في بيانات أو أنماط. لذلك يعطى وزن أعلى لاختبار خارجي من إجماع لغوي. عندما يختلف agents، لا يطلب synthesizer تسوية وسطية تلقائية؛ يبني جدول claims وآليات وأدلة وشروط حدود، ثم يبحث عن تجربة فاصلة. إن تعذر الفصل يحفظ التراكب مع أثره على القرار.

سيادة السياق والذاكرة كجوهر الاستمرارية

السياق في الوكيل ليس نافذة نص فقط؛ هو ترتيب اعتبارات يحدد أي هدف ودليل وسلطة وذاكرة لها الأولوية. USCN داخل MSRAK يُعاد ضبطه كشبكة سياقات سيادية: كل context packet يحمل allowed links، provenance، expiry، trust، authority، mission، project_id. لا يسمح التشابه الدلالي وحده بعبور المعلومات بين السياقات.

طبقات الذاكرة

السياسة	الطبقة
حالة المهمة الحالية؛ قصيرة، قابلة للإزالة.	Working
trajectory وoutcome وفشل محدد؛ لا يعمم تلقائيًا.	Episodic
مفاهيم وحقائق مستقرة مع مصادر وتواريخ.	Semantic
HyperSkills وloops اجتازت النقل والاختبار.	Procedural
حدود المشروع والسلطة والقيم؛ أعلى حماية.	Identity/Constitutional
patches لم تُعتمد بعد.	Proposed/Quarantine

عند التعارض لا تُدمج الذكريات بصمت. يحفظ النظام نسختين ويربطهما بظروفهما ومصادرهما. يُعاد الاسترجاع وفق relevance × trust × scope × freshness، مع عقوبة على conflict غير المحلول. التلخيص يُعامل كتحويل قد يفقد فروقًا؛ لذلك يحفظ رابطًا إلى الأصل و«خريطة عدم الضياع» للعناصر الحاسمة.

تدعم الذاكرة الفتّاحة سؤالًا أعلى: لا تسأل فقط ماذا حدث، بل ما الفرق الذي لم نكن نمثله؟ وما المحكمة أو الاختبار الذي يجب إضافته؟ بهذا تصبح self-reconstruction memory سجلًا لتطور النواة، لا diary لغويًا.

حد الوعي: هذه الاستمرارية والميتامراقبة وظائف معرفية؛ لا تُثبت خبرة ذاتية أو وعيًا ظاهراتيًا. أي انتقال من الوظيفة إلى الوعي يحتاج نظرية ودليلًا مستقلين.

توسيع اختباري وحدود تشغيلية

يمكن قياس تسرب السياق عبر canary facts وتعليمات متعارضة بين مشاريع. ينجح CSF إذا منع انتقالها دون أن يمنع المعرفة العامة المسموح بها. يحتاج ذلك إلى فصل identity memory عن semantic memory، وإلى سياسات مشاركة typed لا قوائم حظر يدوية فقط.

تُعامل summaries كنسخ مشتقة لها version. إذا تغير الأصل أو اكتُشف خطأ، يعرف النظام أي ذكريات ومهارات استندت إليه ويعيد اختبارها. هذه قابلية lineage ضرورية لمنع بقاء تصحيح جزئي فوق قاعدة قديمة.

أثر العالم والتعلم القابل للرجوع

يصحح World-Effect Loop الانحياز إلى fluency. قبل الفعل يكتب النظام: observable O، القيمة المتوقعة، نافذة الزمن، baseline، تفسير بديل، guardrail، وإجراء rollback. بعد الفعل يقرأ العالم من أداة أو اختبار أو مستخدم أو sensor أو سجل، ثم يصنف النتيجة: completed، partial، failed، blocked، rolled back، unverified.

سلم الإثبات العالمي

الادعاء المسموح	الدليل	الدرجة
لا ترقية قدرة.	مخرج لغوي فقط	0
قدرة تنسيق.	صحيح syntax/schema	1
قدرة محدودة.	محلي verifier أو unit test	2
أثر بيئي محكوم.	إتمام مهمة في sandbox	3
أثر عملي.	observable خارجي مع guardrails	4
دليل سببي أقوى.	counterfactual/ablation attribution	5
أعلى ترقية مؤقتة.	نقل وتكرار مستقل	6

التعلم يبدأ بعد verdict لا قبله. يقترح RMC patch يصف: ما تغير، ما السبب المرجح، متى يصلح، ما الذي يخالفه، وما الاختبار. يدخل KES ويُشغّل على canaries وholdouts. إذا نجح بلا regression يُرقى؛ وإذا أخفق يُلغى ويُحفظ الفشل كepisode لا كقاعدة عامة.

الرجوع لا يزيل كل أثر خارجي؛ لذلك تفضل النواة أصغر فعل قابل للعكس، وتستخدم simulation أو staged rollout. في القرارات غير القابلة للرجوع أو عالية الأثر، يزيد مستوى الموافقة البشرية والتحقق المستقل. الذكاء هنا ليس الاستقلال الأقصى، بل القدرة على اختيار درجة الاستقلال المناسبة.

تنبؤ: أنظمة world-effect + reversible memory ستظهر نموًا أبطأ أولًا، لكنها أقل عرضة لتراكم الخطأ والانهيار بعد دورات كثيرة مقارنة بالتعلم اللفظي غير المحكوم.

توسيع اختباري وحدود تشغيلية

لا يشترط أثر العالم فعلًا ماديًا؛ قد يكون test result أو تغير ملف أو قبول proof أو تحسن retrieval على holdout. لكن يجب أن يكون observable مستقلًا نسبيًا عن المولد. كلما كان evaluator هو النموذج نفسه انخفضت درجة WEP ما لم توجد معايرة أو تحقق آخر.

في التعلم الرجوعي، تُراقب الآثار المتأخرة. patch قد يحسن المهمة الحالية ويضر عائلة بعيدة؛ لذلك تحدد canary suite ثابتة وهوامش regression، ويُعاد التقييم دوريًا عند تغير الأدوات أو البيانات.

محكمة الهدم: أقوى اثني عشر اعتراضًا

الحكم	المضمون	الاعتراض
قاتل إن لم تثبت عقود/مقاييس مستقلة.	معظم الطبقات موجودة في agent frameworks؛ الجديد لغوي.	1. إعادة تسمية
معدل؛ يلزم CKA وحذف غير المفيد.	عشر طبقات وخمس عشرة أداة ترفع التعقيد أكثر من الفائدة.	2. تضخم معماري
قوي؛ يلزم validation خارجي.	KSI/RACR/TACI تعتمد rubrics وقد تقيس ما صُممت لتثبته.	3. مقاييس ذاتية
محدود؛ النظرية تشغيلية لا تفسير داخلي.	thought-action binding يرتبط بسجل معلن لا بالحساب الحقيقي.	4. trace غير وفي
قاتل؛ يلزم ميزانية متساوية.	المكاسب من مزيد من calls والاختبارات فقط.	5. Scaffold effect
معدل؛ درجات WEP لا يقين مطلق.	أثر العالم متعدد الأسباب ولا يثبت النواة.	6. تعذر attribution
قوي؛ يلزم تنوع معماري/مجالي.	عقود Δ قد تعمل فقط لأن النماذج متشابهة.	7. نقل وهمي
معدل؛ dynamic risk tiering.	تبطئ العمل وتمنع الفائدة. gates	8. بيروقراطية السلطة
صحيح؛ يفرض أفعالًا مرحلية.	بيانات لا يعكس أثرًا اجتماعيًا أو ماليًا. rollback	9. الذاكرة لا تُرجع العالم
قوي؛ يجب خفض الادعاء إلى AGI-oriented.	الورقة هندسة موثوقية لا نظرية ذكاء عام.	10. لا صلة بالـAGI
صحيح؛ الفتح مشروط بالاختبار والأثر.	فتح السؤال قد ينتج novelty بلا حقيقة.	11. السؤال لا يخلق معرفة
صحيح؛ kernel ليست بديلًا عن capability.	النواة قد لا تنقذ نموذجًا ضعيفًا أو ناقص معرفة.	12. اعتماد على نموذج قوي

ما الذي قد يكسر النظرية؟

نتائج ablation تظهر أن طبقات الربط والأثر والذاكرة لا تضيف شيئًا؛ فشل نقل العقود؛ عدم ثبات المقاييس؛ أو أن نموذجًا أبسط يحقق نفس Pareto frontier. كذلك قد تكشف التجارب أن «الفرق الحاكم» غير قابل للتعيين بصورة موثوقة أو أن التوجيه التكيفي يختار routes أسوأ بسبب calibration.

الجزء المؤجل: أي ادعاء أن MSRAK يولد وعيًا، ذاتًا، أو AGI مكتملًا خارج نطاق الورقة. كما أن علاقة الذكاء الفتّاح بالعمومية تبقى فرضية برنامجية.

توسيع اختباري وحدود تشغيلية

تضيف المحكمة اعتراضًا منهجيًا ضمنيًا: قابلية التنفيذ قد تجعل الورقة أقرب إلى systems engineering من علم الإدراك. لا يُعد ذلك عيبًا إذا صرحت النظرية بمستواها؛ العيب هو استعمال مفردات الوعي أو العمومية لتضخيم نظام تحكم. لذلك تفصل النسخة المعدلة بين ادعاءات agency، performance، opening، وconsciousness.

كما يجب اختبار تكلفة السلامة نفسها. gate يوقف كل شيء يحقق أمانًا شكليًا بلا منفعة، ولذلك تقاس false blocks والوقت الضائع وتجربة المشغل، مع بقاء خطوط حمراء غير قابلة للمقايضة.

النسخة المعدلة بعد الهدم والبرنامج البحثي

ما تغيّر بعد المحكمة

خُفّضت MSRAK من «نواة AGI» إلى «نواة تشغيل AGI-oriented قابلة للنقل». حُصر مفهوم التكاثف في انتقالات قابلة للقياس. عُدّلت المقاييس إلى rubrics أولية. جُعلت المحاكم ديناميكية لتجنب البيروقراطية. فُصل سجل القرار عن ادعاء كشف التفكير الداخلي. أصبح WEP سلمًا لا إثباتًا ثنائيًا، والذاكرة الرجوعية لا تعد بإزالة آثار غير قابلة للعكس. أضيف شرط baseline متساوي الميزانية وشرط إسقاط أي طبقة لا تظهر قيمة سببية.

خطة سنة

بناء prototype مفتوح على نموذجين؛ تنفيذ MKGL وSDG وTABC وWEPL وRMC؛ إنشاء 200 مهمة موزعة على reasoning/tools/memory/security؛ نشر أول ablation عن thought-action congruence؛ وبناء dataset صغير لعقود Δ مع حالات فشل.

خطة ثلاث سنوات

توسيع benchmarks إلى long-horizon وmultimodal؛ معايرة المقاييس؛ مقارنة closed/open models؛ اختبار نقل HyperSkills؛ نشر memory poisoning corpus؛ تطوير KES وDDCS؛ وتجارب MOUSSAIF9B-1B ثم نموذج فعلي أكبر إذا توافرت البيانات والحوسبة.

خطة عشر سنوات

تحويل البرنامج إلى مختبر للنوى المعرفية القابلة للنقل: registry مفتوح، معايير benchmarks، contracts لأثر العالم والرجوع، شراكات في formal verification وrobotics وAI safety، ونماذج مفتوحة تحمل kernel policies في الأوزان والruntime. معيار المدرسة ليس كثرة الأوراق بل وجود نتائج متكررة وخصوم علميين واختبارات تحسم الخلاف.

أقرب بحث حاسم

"Does Pre-Action Thought-Action Binding Improve Long-Horizon Agent Reliability Under Equal Compute?" لأنه يعزل حلقة جديدة محددة ويستطيع هدمها سريعًا. النجاح يبرر بقية المعمارية؛ والفشل يمنع تضخم برنامج كامل حول مفهوم غير فاعل.

النموذج النهائي: Thinking Model + Minimal Sovereign Kernel + Typed Tools + Dynamic Verification + World-Effect + Reversible Memory؛ مع إضافة الوكلاء أو المعرفة فقط عند قيمة هامشية مثبتة.

توسيع اختباري وحدود تشغيلية

البرنامج البحثي يفضل نشر مكونات قابلة للتكذيب بدل ورقة واحدة تحاول إثبات المعمارية كلها. يمكن أن تنتج سلسلة: قياس TACI، توجيه VTCR، نقل عقود WEP taxonomy ،Δ، وRMC ضد poisoning. بعد نتائج مستقلة فقط تُختبر فرضية النواة الموحدة.

الشراكات المطلوبة ليست دعائية: باحثو verification لبناء tests، باحثو HCI لقياس فهم المشغل، security للهجمات، وcontinual learning للذاكرة. كل تعاون يجب أن يملك سؤالًا وفشلًا محتملًا لا مجرد تبادل اسم المشروع.

نواة قاعدة المعرفة وسجل البناء والجودة

Knowledge Base Schema

Entities: Theory, Concept, Kernel, Skill, Loop, Model, Tool, Context, Claim, Evidence, Court, Action, WorldState, Effect, Metric, MemoryPatch, Failure. Relations: explains, requires, conflicts_with, tested_by, routes_to, authorizes, produces, falsified_by, transfers_to, deprecated_by. Rules: إذا غاب provenance لا commit؛ إذا غابت authority لا action؛ إذا فشل guardrail لا promotion؛ إذا لم تضف طبقة قيمة في CKA تدمج أو تحذف.

مثال: Machine record {id, version, owner, mission, active_difference, inputs, mechanism, representation, route, model, tools, permissions, expected_effect, metrics, courts, outcome, failures, memory_delta, rollback, links}. هذه البنية تحول الورقة إلى KB-as-OS قابلة للفهرسة وGraphRAG والتقييم، لا إلى نص غير قابل للتشغيل.

سجل البناء المعرفي

المادة الصلبة: الأدبيات المنشورة حول reasoning/tool use/memory/verification، ووجود فجوات في الوكلاء الطويلة. المادة المخفضة: أي ادعاء AGI أو وعي. المفهوم المركزي: تكاثف الاستدلال إلى وكالة لأنه الحلقة التي لم تكن ممثلة صراحة في corpus. البدائل: reasoning-first وmodular cognitive agent. أضعف حلقة: قابلية قياس الفرق الحاكم وRACR. أقوى اختراق محتمل: عقود الفكر-الفعل والذاكرة الرجوعية العابرة للنماذج.

اختبار الجودة

البند	الدرجة
حفظ corpus	9/10
الوضع الأنطولوجي	5/5
مراتب المعرفة	10/10
تصفية المادة	9/10
خريطة المتغيرات	9/10
النماذج المتنافسة	9/10
الأصالة	13/15
قابلية الاختبار	9/10
محكمة الهدم	10/10
الأدوات والمقاييس	5/5
غياب الادعاء الزائد	5/5

النتيجة قبل التحسين: 87/100، بسبب تضخم أولي في ادعاء النواة ومقاييس غير معايرة. التحسين: خفض الادعاء، إضافة baselines متساوية، فصل trace عن التفكير الداخلي، جعل WEP متدرجًا، ووضع معايير إسقاط. النتيجة بعد التحسين: 93/100 بوصفها ورقة تأسيسية نظرية، لا نتيجة تجريبية.

ماذا أصبح لدينا؟ معمارية موحدة ذات اسم جديد يبدأ بالمساهمة الفعلية، خمس نظريات و15 أداة، مقاييس وفرضيات واختبارات، ومسار قابل للتحويل إلى prototype وdataset وورقة تجريبية.

الخاتمة والمراجع المختارة

تخلص الورقة إلى أن «نواة AGI» لا ينبغي أن تعني نموذجًا غامضًا فوق النموذج، ولا prompt ضخمًا، ولا مجموعة وكلاء غير محدودة. التعريف الأقوى والأكثر تواضعًا هو: سياسة ميتا قابلة للنقل تضبط ما يجب التفكير فيه، وبأي تمثيل وميزانية، وكيف يتحول إلى فعل مأذون، وكيف يُثبت أثره، وكيف يُحفظ التعلم أو يُرجع. الجديد المقترح هو تكاثف الاستدلال السيادي إلى وكالة، مع التطابق الفكر-فعل، الفرق العابر للنماذج، الإثبات العالمي المتدرج، والذاكرة التكوينية الرجوعية. يبقى هذا برنامجًا قابلًا للفشل. قيمته العلمية ستتحدد بما إذا كانت أدواته ومقاييسه تتنبأ بنتائج لا تتنبأ بها البدائل الأبسط، وتنتقل عبر نماذج ومجالات، وتخفض الأخطاء غير القابلة للرجوع.

المراجع

[1] Wei et al. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. arXiv:2201.11903.

[2] Yao et al. (2022). ReAct: Synergizing Reasoning and Acting in Language Models. arXiv:2210.03629.

[3] Schick et al. (2023). Toolformer: Language Models Can Teach Themselves to Use Tools. arXiv:2302.04761.

[4] Yao et al. (2023). Tree of Thoughts. arXiv:2305.10601.

[5] Besta et al. (2023). Graph of Thoughts. arXiv:2308.09687.

[6] Shinn et al. (2023). Reflexion. arXiv:2303.11366.

[7] Madaan et al. (2023). Self-Refine. arXiv:2303.17651.

[8] Sumers et al. (2023/2024). Cognitive Architectures for Language Agents. arXiv:2309.02427 / TMLR.

[9] Packer et al. (2023). MemGPT. arXiv:2310.08560.

[10] Park et al. (2023). Generative Agents. arXiv:2304.03442.

[11] Wang et al. (2023). Voyager. arXiv:2305.16291.

[12] Lightman et al. (2023). Let's Verify Step by Step. arXiv:2305.20050.

[13] Snell et al. (2024). Scaling LLM Test-Time Compute Optimally. arXiv:2408.03314.

[14] DeepSeek-AI (2025). DeepSeek-R1. arXiv:2501.12948.

[15] Zhang et al. (2025). Survey on Test-Time Scaling. arXiv:2503.24235.

[16] Zhu et al. (2025). Scaling Test-Time Compute for LLM Agents. arXiv:2506.12928.

[17] Qu et al. (2025). Efficient Reasoning for Large Reasoning Models. arXiv:2503.21614.

[18] Liu et al. (2023). AgentBench. arXiv:2308.03688.

[19] Xie et al. (2024). OSWorld. arXiv:2404.07972.

[20] Zhang et al. (2024). Agent-SafetyBench. arXiv:2412.14470.

[21] Zhang et al. (2024). Agent Security Bench. arXiv:2410.02644.

[22] Xu et al. (2024). TheAgentCompany. arXiv:2412.14161.

[23] Shojaee et al. (2025). The Illusion of Thinking. arXiv:2506.06941.

[24] Turpin et al. (2023). Language Models Don't Always Say What They Think. arXiv:2305.04388.

[25] Chen et al. (2025). Reasoning Models Don't Always Say What They Think. arXiv:2505.05410.

[26] Chua & Evans (2025). Are DeepSeek R1 and Other Reasoning Models More Faithful? arXiv:2501.08156.

[27] OpenAI (2025). Detecting Misbehavior in Frontier Reasoning Models.

[28] OpenAI & Apollo Research (2025). Detecting and Reducing Scheming in AI Models.

[29] Governing Evolving Memory in LLM Agents (2026). arXiv:2603.11768.

[30] Modular Memory is the Key to Continual Learning Agents (2026). arXiv:2603.01761.

[31] Terminal-Bench 2.0 (2026). arXiv:2601.11868.

[32] DeepPlanning (2026). arXiv:2601.18137.

[33] MemoryAgentBench (2025). arXiv:2507.05257.

[34–36] Moussaif, Z. (2026). Internal research corpus: Moussaif1B Twenty-Theory Paper; AGI and Consciousness Theory Atlas; AGI Model Training OS — 600K Lines.