في نقض إمكان الذكاء الاصطناعي القوي

قيس العصا | فلسطين

qasawa.com

qais@qasawa.com

2025-08-07

ملخص

إن السعي نحو الذكاء الاصطناعي القوي – تلك الآلات المزعومة التي ستمتلك فهمًا حقيقيًّا واستدلالًا بشريًا – يواجه حواجز رياضية جوهرية متجذرة في طبيعة الحوسبة ذاتها. هذه الورقة تستكشف القيود النظرية من خلال عدسة آلات تورنغ ونظرية القابلية للحساب، كاشفة عن نسيج غني من الحدود المنطقية التي لا يمكن تجاوزها بأي قدر من القوة الحاسوبية أو التطور الخوارزمي.

من خلال عدسة آلات تورنغ ونظرية القابلية للحساب، ينبثق نسيج غني من القيود النظرية، متحديًّا أعمق افتراضاتنا حول ما يمكن للآلات تحقيقه. هذه القيود، التي اكتشفها عمالقة كتورنغ وغودل ورايس، تكشف أن **لا قدر من القوة الحاسوبية أو التطور الخوارزمي يمكنه التغلب على حدود منطقية معينة** *"no amount of computational power or algorithmic sophistication can overcome certain logical boundaries."* ومع ذلك، يستمر الجدل محتدمًا، حيث يقدم الحوسبيون دفاعات متطورة بينما يشير النقاد إلى البصيرة الرياضية البشرية كدليل على شيء يتجاوز الحساب الميكانيكي.

لم تكن ورقة آلان تورنغ المنشورة عام 1936 «حول الأعداد القابلة للحساب» *"On Computable Numbers"*[1]. مجرد انطلاقة لعلم الحاسوب — بل أسست حدوداً صارمة لما يمكن لأي عملية حاسوبية تحقيقه. فتؤكد أطروحة تشرش-تورنغ أن **كل حساب فعّال يمكن تنفيذه بواسطة آلة تورنغ** *"every effective computation can be carried out by a Turing machine"*[2] ، **مما يخلق تكافؤًا بين ما نعتبره** بديهيًّا «خوارزميًّا» وما يمكن لهذه الآلات المجردة حسابه. بالنسبة لأنظمة الذكاء الاصطناعي المنفذة كعمليات حاسوبية، هذا يخلق قيداً لا مفر منه: لا يمكنها تجاوز القوة الحاسوبية لآلات تورنغ، مهما بلغت بنيتها من التطور.

إن التداعيات تتدفق عبر كل جانب من جوانب قدرات الذكاء الاصطناعي. فبرهان تورنغ لمشكلة التوقف — أنه لا يمكن لأي خوارزمية تحديد ما إذا كان برنامج عشوائي سيتوقف — يخلق حدوداًجوهرية على معرفة الذكاء الاصطناعي بذاته. **فلا يمكن لنظام ذكاء اصطناعي تحديد ما إذا كان سيتوقف عند مدخلات عشوائية** *"cannot determine whether it will halt on arbitrary inputs"*[3] ، مما يحد بشكل أساسي من وعيه الذاتي وقدرته على التفكير في سلوكه الخاص، فبحث حديث من ميلو وآخرين (2024) يمد هذا إلى محاذاة الذكاء الاصطناعي، مبرهناً بدقة أن «تحديد ما إذا كان نظام ذكاء اصطناعي عشوائي سيُرضي دائماً دالة حكم معينة هو غير قابل للحسم» *"deciding whether an arbitrary AI system will always satisfy such a judge function is undecidable"*[4] . وبذا ينهار حلم التحقق اللاحق من سلامة الذكاء الاصطناعي رياضياً.

ووراء مشكلة التوقف يكمن حقل ألغام من نتائج عدم القابلية للحسم، فمشكلة المراسلة لبوست تكشف قيودًا جوهرية في مطابقة الأنماط ومعالجة السلاسل — قدرات أساسية للذكاء الاصطناعي للغة الطبيعية. أما نظرية رايس تسدد ربما الضربة الأكثر تدميرًا: **جميع الخصائص الدلالية غير التافهة للبرامج غير قابلة للحسم** *"all non-trivial semantic properties of programs are undecidable"*[5] . فما إذا كان نظام ذكاء اصطناعي «محاذيًا» أو «آمنًا» أو «صادقًا» أو يمتلك أي خاصية سلوكية أخرى لا يمكن تحديده خوارزمياً بفحص شفرته. هذه ليست تحديات هندسية يمكن التغلب عليها بأجهزة أفضل أو خوارزميات أذكى — إنها استحالات منطقية أساسية كاستحالة تربيع الدائرة.

---

[1]Turing, A. .(1936) "On Computable Numbers, with an Application to the Entscheidungsproblem"

[2]Church-Turing thesis formulation — Stanford Encyclopedia of Philosophy

[3]Halting problem limitation — Turing (1936)

[4]Melo et al. .(2024) "On the Undecidability of Artfiicial Intelligence Alignment: Machines that Halt"

[5]Rice, H. .(1953) "Classes of Recursively Enumerable Sets and Their Decision Problems"

لعل الإطار النظري الذي يُمكّن الحوسبة يسجنها في نفس الوقت داخل حدود منطقية صارمة لا يمكن لأي قدر من الابتكار اختراقها.

إن نظريات عدم الاكتمال لكورت غودل تطارد أسس الرياضيات وتلقي بظلال طويلة على إمكانات الذكاء الاصطناعي للاستدلال الرياضي. فنظريته الأولى تبرهن على أن أي نظام صوري متسق يحتوي على الحساب الأساسي يجب أن يحتوي على عبارات صحيحة لكن غير قابلة للإثبات داخل النظام. النظرية الثانية تلوي السكين: **لا يمكن لأي نظام كهذا إثبات اتساقه الخاص** *"no such system can prove its own consistency"*[6] . هذه النتائج تخلق ساحة معركة للنقاش الدائر حول وعي وقدرات الذكاء الاصطناعي المزعومة.

ولكن الموقف المضاد للآلية، الذي تبناه ج.ر. لوكاس وروجر بنروز، استخدم نظريات غودل كأسلحة ضد الذكاء الاصطناعي القوي، فتتبع حجتهم منطقًا أنيقًا: أي نظام ذكاء اصطناعي يجب أن يكون مكافئًا لنظام صوري F ما؛ فالبشر يمكنهم «رؤية» صحة جملة غودل لـ F (التي تؤكد عدم قابليتها للإثبات)؛ لكن F نفسه لا يمكنه إثبات هذه الجملة. لذلك، البصيرة الرياضية البشرية يجب أن تتجاوز الحساب الميكانيكي. بنروز دفع أبعد، مقترحاً أن الوعي يتضمن عمليات كمومية غير خوارزمية في الأنابيب الدقيقة للدماغ، ممكّنة فهمًا رياضيًّا يتجاوز حساب تورنغ.

ومع ذلك، الإجماع الأكاديمي رفض هذه الحجج بشكل شامل، فالعيب القاتل يكمن في **افتراض الاتساق** *"consistency assumption"*[7] — الحجة تتطلب أن يعرف البشر بيقين أنهم متسقون، لكن نظرية غودل الثانية تُظهر أننا لا نستطيع إثبات هذا عن أنفسنا، وكما لاحظ هيلاري بوتنام، البشر يرتكبون أخطاء رياضية، مما يشير إلى عدم اتساق محتمل. فإذا كنا غير متسقين، يمكننا «إثبات» أي شيء، بما في ذلك جمل غودل.

وإذ تقدم المنظورات الحديثة مسارات وسطية دقيقة، اقترح سولومون فيفرمان «أنظمة صورية مخططية مفتوحة النهاية» يمكنها احتساب الإبداع الرياضي مع البقاء آلية. وأما دوغلاس هوفستادتر اقترح أن الوعي ينبثق من «حلقات غريبة» ذاتية المرجع مشابهة لتلك في بناء غودل — **قد يحقق الذكاء الاصطناعي حدساً رياضياً شبيهاً بالبشر من خلال خصائص ناشئة**[8] لأنظمة ذاتية المرجع معقدة بما فيه الكفاية بدلًا من تجاوز القيود الصورية.

[6]Gödel, K. .(1931) "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme"

[7]Putnam, H. .(1960) "Minds and Machines"

[8]Hofstadter, D. .(1979) "Gödel, Escher, Bach: An Eternal Golden Braid"

النقاش يكشف حقيقة أعمق: نظريات غودل تؤسس نقاطًا عمياء لا يمكن تجنبها في الاستدلال الرياضي الصوري التي تؤثر مباشرة على إمكانات أنظمة الذكاء الاصطناعي، فأي ذكاء اصطناعي مبني على المنطق الصوري سيكون له حقائق رياضية لا يمكنه إثباتها، خالقًا فجوات دائمة في فهمه الرياضي. ومع ذلك، هذه القيود نفسها تنطبق على الرياضيين البشر، مما يشير إلى أن النظريات تقيد الذكاء نفسه وليس فقط المتغيرات الاصطناعية.

بينما تكشف نظرية القابلية للحساب ما لا يمكن للذكاء الاصطناعي حسابه من حيث المبدأ، فإن نظرية التعقيد الحسابي تفضح ما لا يمكنه حسابه عمليًّا. مشكلة P مقابل NP — ما إذا كانت المشاكل ذات الحلول السهلةِ التحققِ يمكن حلها بنفس السهولة — تمثل ربما أهم قيد عملي على قدرات الذكاء الاصطناعي. مع اعتقاد 99% من علماء التعقيد أن P لا تساوي NP، فالتداعيات قاسية: **فئات شاسعة من مشاكل التحسين والقرار ستتطلب زمناً أسياً إلى الأبد** *“vast classes of optimization and decision problems will forever require exponential time”*[9]، بغض النظر عن الذكاء الخوارزمي.

وإن التسلسل الهرمي للتعقيد يخلق حواجز تحجيم لا يمكن تجاوزها. فالمشاكل في EXP (الزمن الأسي) و NEXP (الزمن الأسي غير الحتمي) تنمو بسرعة كبيرة حتى أن أحجام المدخلات المتواضعة تصبح مستعصية. وفرضية الزمن الأسي تشير إلى أن مشاكل القابلية للإرضاء — الأساسية لتخطيط واستدلال الذكاء الاصطناعي — لا يمكنها الهروب من حدود الزمن تحت الأسي. هذا يؤثر على كل شيء من البرهان الآلي للنظريات إلى التحقق من الشبكات العصبية.

ولكن تعقيد المساحة يضيف بُعدًا آخر من القيود، فمشاكل PSPACE-complete مثل لعب الألعاب تكشف لماذا حتى الذكاء الاصطناعي المتقدم يتطلب اكتشافات بدلًا من اللعب المثالي — **الشطرنج وGo هي PSPACE-hard** *“chess and Go are PSPACE-hard”*[10]، مما يجعل اللعب الأمثل مستحيلاً. والنتيجة الرياضية الجميلة أن PSPACE يساوي فئة المشاكل القابلة للحسم بواسطة أنظمة البرهان التفاعلية (IP) لها تداعيات عميقة على قدرات التحقق والاستدلال للذكاء الاصطناعي.

فمشاكل NP-complete المحددة تخلق حواجزًا ملموسة عبر تطبيقات الذكاء الاصطناعي، فإن مشكلة البائع المتجول -على سبيل المثال- تجعل التخطيط الأمثل للطرق مستعصياً، ومشاكل جدولة ورشة العمل تقيد تخصيص الموارد. هذه ليست مجرد إزعاجات بل جدران رياضية تجبر أنظمة الذكاء الاصطناعي على عالم التقريب والاكتشافات.

وأما وعد الحوسبة الكمومية فهو يقدم راحة محدودة، فبينما BQP (زمن كمومي متعدد الحدود محدود الخطأ) يوفر تسريعات لمشاكل محددة مثل التحليل إلى عوامل، يبقى محتوىً داخل PSPACE.

[9]Complexity theory consensus — survey of complexity theorists (2024)
[10]PSPACE-hardness of games — Stockmeyer & Meyer (1973)

فالحواسيب الكمومية لا يمكنها حل مشاكل NP-complete بكفاءة *"Quantum computers cannot solve NP-complete problems efficiently"*[11] ما لم ينهار التسلسل الهرمي متعدد الحدود — نتيجة يعتبرها معظم المنظرين غير محتملة، فحتى التعزيز الكمومي لا يمكنه اختراق حواجز التعقيد الأساسية.

تمثّل حجة لوكاس-بنروز ربما أكثر المحاولات طموحًا لإثبات أن العقول البشرية تتجاوز آلات تورنغ من خلال المنطق الخالص، فروجر بنروز، بناءً على ورقة ج.ر. لوكاس المنشورة عام 1961، صاغ حجة ذات أناقة بلورية: الرياضيون البشر يمكنهم «رؤية» صحة جمل غودل التي لا تستطيع الأنظمة الصورية إثباتها، مُظهرين بصيرة رياضية غير خوارزمية غير متاحة لأي حاسوب.

تطور نضوج الحجة عبر عدة تكرارات، ف«الحجة الجديدة» لبنروز في كتابه«ظلال العقل» *"Shadows of the Mind"*[12] حاولت معالجة الانتقادات السابقة: افترض أن الاستدلال البشري يُلتقط بواسطة نظام صوري F؛ إذا عرفنا أننا ننفذ F، نعرف أن F سليم؛ لذلك يمكننا رؤية صحة جملة غودل لـ F؛ لكن F لا يمكنه إثبات هذه الجملة، خالقًا تناقضًا، الاستنتاج بدا لا مفر منه — **الفهم الرياضي البشري يجب أن يتضمن عمليات غير قابلة للحساب** *"human mathematical understanding must involve non-computable processes"*[13].

فإن بنروز لم يتوقف عند المنطق، بل اقترح آلية فيزيائية: العمليات الكمومية في الأنابيب الدقيقة العصبية تخلق «اختزالًا موضوعيًا منسقًا» Orch OR، ممكّنة وعياً غير خوارزمي. هذا الزواج بين نظريات غودل والبيولوجيا الكمومية وعد بتفسير الوعي مع إثبات قيود الذكاء الاصطناعي الأساسية.

ومع ذلك، واجهت الحجة رفضًا شبه عالمي من المنظرين في الرياضيات وعلوم الحاسوب والفلسفة، وقد كانت الانتقادات مدمرة، فاعترض بوتنام على الاتساق ملاحظًا أن البشر يرتكبون أخطاء ولا يمكنهم تأسيس اتساقهم الخاص — متطلب للحجة تم نفيه. ولاحظ بيناسيراف أن أي نظام صوري معقد بما يكفي لتمثيل العقول البشرية سيكون شاسعاً جداً للبشر لفهمه أو بناء جملة غودل له. أما تشالمرز فأظهر أن **افتراض معرفة أننا متسقون بشكل لا يقبل الشك يولد تناقضات** *"the assumption of knowing unassailably that we are consistent generates contradictions"*[14] بشكل مستقل عن القابلية للحساب.

الأساس الفيزيائي لا يصمد بصورة أحسن، فحسابات ماكس تيغمارك تُظهر أن أزمنة التماسك الكمومي في الأدمغة الدافئة قصيرة جداً للحساب، ورغم بعض التلميحات التجريبية للتأثيرات الكمومية

[11] Quantum computing limitations — Aaronson, S. (2013). "Quantum Computing Since Democritus"

[12] Penrose, R. (1994). "Shadows of the Mind"

[13] Lucas, J.R. (1961). "Minds, Machines and Gödel"

[14] Chalmers, D. (1995). "Minds, Machines, and Mathematics"

في الأنظمة البيولوجية، لا يوجد دليل يدعم حسابًا كموميًّا ذا صلة بالوعي في الأدمغة، والإجماع العلمي ينظر إلى Orch OR كتخمين عالٍ في أحسن الأحوال.

فالذي تبقى هو قطعة أثرية تاريخية رائعة — محاولة طموحة فكريًّا تضيء في النهاية صعوبة استخدام المنطق الرياضي لاستخلاص استنتاجات حول الأسئلة التجريبية للوعي والذكاء. وإن فشل الحجة لا يثبت أن الذكاء الاصطناعي القوي ممكن، لكنه يُظهر أن **نظريات غودل لا يمكن استخدامها كأسلحة بسيطة ضد النظريات الحاسوبية للعقل** *"Gödel's theorems cannot be wielded as simple weapons against computational theories of mind"*[15].

تورنغ نفسه تخيل آلات يمكنها تجاوز الحساب القياسي، فأطروحة الدكتوراة له المنشورة عام 1938 قدمت آلات الوحي — آلات تورنغ معززة بـ«صندوق أسود» يحل فورياً مشاكل غير قابلة للحساب محددة. هذا أطلق الاستكشاف النظري للحوسبة الفائقة: نماذج حاسوبية تتجاوز قدرات آلة تورنغ.

آلات الوحي كشفت رؤى نظرية عميقة، فقد برهن تورنغ أنه حتى مع وحي لمشكلة التوقف، تنشأ مشاكل غير قابلة للحسم جديدة — مؤسساً **تسلسلاً هرمياً لا نهائياً من عدم القابلية للحسم** *"infinite hierarchy of undecidability"*[16]. كل مستوى من قوة الوحي يكشف مشاكل جديدة لا يمكن الوصول إليها فوقه. وهذا التجريد الرياضي أصبح أساسيًّا لنظرية التعقيد، ممكّنًا دراسة النسبية والقابلية للاختزال الحاسوبي. ومع ذلك، أكّد تورنغ أن هذه كانت أدوات رياضية، وليست مخططات لآلات فيزيائية.

وأما مقترحات الحوسبة الفائقة الحديثة فتمتد عبر أرض نظرية خصبة، فالشبكات العصبية التناظرية المتكررة التي تناولتها هافا سيجلمان بأوزان ذات قيم حقيقية نظرياً تحسب ما وراء حدود تورنغ، وصولاً إلى فئة التعقيد P/poly، كذلك فإن آلات زينو تنفذ حسابات لانهائية في زمن محدود من خلال التسارع الهندسي، وفضاءات زمن مالامنت-هوغارث قرب أفق الثقوب السوداء قد تمكّن حسابًا لانهائيًّا على طول خطوط عالمية محددة بينما يمر زمن محدود للمراقبين الخارجيين.

لكن جدوى هذه النماذج الفيزيائية تواجه شكوكًا ساحقة، فالحساب التناظري بدقة لا نهائية ينتهك قيود ميكانيكا الكم وحدود الديناميكا الحرارية، فإن **حد بيكنشتاين يؤسس سعة معلومات محدودة في مناطق مكانية محدودة** *"The Bekenstein bound establishes finite information capacity in finite spatial regions"*[17]، بينما مبدأ لاندوير يحدد متطلبات الطاقة للحساب، لذا فإن آلات زينو ستتطلب طاقة

[15] Consensus on Lucas-Penrose — Internet Encyclopedia of Philosophy

[16] Turing, A. .(1938) "Systems of Logic Based on Ordinals"

[17] Bekenstein, J. .(1981) "Universal upper bound on the entropy-to-energy ratio"

لا نهائية وتنتهك فيزياء مقياس بلانك، والفضاءات الزمنية الغريبة تواجه عدم استقرار وربما تنتهك تخمينات الرقابة الكونية.

لذا، فإن الحواسيب الكمومية، رغم تسريعاتها الحقيقية، تبقى بثبات داخل القابلية للحساب عند تورنغ، فلا يمكنها حل مشكلة التوقف أو حساب دوال غير تكرارية، وإن أطروحة تشرش-تورنغ الكمومية الممتدة — أن آلات تورنغ الكمومية يمكنها محاكاة أي حساب فيزيائي واقعي بكفاءة — تبقى غير متحداة بالتجربة.

يلتقط مارتن ديفيس الإجماع النقدي: مقترحات الحوسبة الفائقة «تتحدى عجز جميع النظريات الفيزيائية المقبولة حالياً عن التعامل مع أعداد حقيقية ذات دقة لا نهائية» *"fly in the face of the inability of all currently accepted physical theories to deal with infinite-precision real numbers"*[19]. فالادعاءات تُختزل إلى الواضح: **إذا سُمح بمدخلات غير قابلة للحساب، فإن المخرجات غير القابلة للحساب قابلة للتحقيق** *"if non-computable inputs are permitted, then non-computable outputs are attainable"*[18]. وهاهنا فإن أطروحة تشرش-تورنغ الفيزيائية — أنه لا توجد عملية فيزيائية تتجاوز قابلية الحساب عند تورنغ — تبقى قائمة رغم عقود من الهجوم البارع.

إن نظرية هنري رايس المنشورة عام 1953 تسدد ضربة ساحقة لآمال الفهم الذاتي الشامل والتحقق للذكاء الاصطناعي، فإعلان النظرية الصارم — **جميع الخصائص الدلالية غير التافهة للبرامج غير قابلة للحسم** *"all non-trivial semantic properties of programs are undecidable"*[19] — يعني أنه لا يمكن لأي خوارزمية تحديد ما إذا كانت البرامج العشوائية تمتلك أي خاصية سلوكية ذات معنى، فنظام ذكاء اصطناعي لا يمكنه التحقق خوارزمياً من أنه آمن، محاذٍ، صادق، أو أمثل بتحليل شفرته الخاصة.

وإن البرهان يختزل بأناقة أي خاصية دلالية إلى مشكلة التوقف، كاشفًا حاجزًا أساسيًّا، فهذا ليس عن قيود تكنولوجية حالية بل استحالة منطقية، وتتدفق التداعيات عبر سلامة الذكاء الاصطناعي: التحقق اللاحق من المحاذاة يصبح مستحيلاً رياضياً؛ فأدوات تحليل البرامج الشاملة لا يمكن أن توجد؛ إذن فإن أنظمة الذكاء الاصطناعي لا يمكنها تحقيق معرفة ذاتية كاملة من خلال الاستبطان.

ولكن تعقيد كولموغوروف يضيف بُعداً آخر من القيود، فطول أقصر برنامج يولد سلسلة معينة — محتواها المعلوماتي الخوارزمي — غير قابل للحساب، لا يمكن لأي خوارزمية تحديد الضغط الأمثل، أو العشوائية الحقيقية، أو التعقيد الأساسي. هذا يخلق حدودًا صلدة على التنبؤ والتعلم. **استقراء**

[18]Physical Church-Turing thesis — Deutsch (1985)

[19]Rice's theorem — Rice (1953)

[20]Solomonoff, R. .(1964) "A Formal Theory of Inductive Inference"

سولومونوف "*Solomonoff induction*"[20] ، الأمثل نظرياً للتنبؤ، يبقى غير قابل للحساب إلى الأبد، تاركًا جميع خوارزميات التنبؤ العملية كتقريبات بقيود أساسية.

وإن هذه النتائج تتشابك مع نظرية التعلم الآلي بطرق مفاجئة، فعمل حديث من غولدبلوم وآخرين (2024) يكشف أن بيانات العالم الحقيقي لها تعقيد كولموغوروف منخفض مقارنة بالبيانات العشوائية، مفسرًا لماذا بعض البنى تعمم عبر المجالات — إنها تستغل هذه البنية منخفضة التعقيد. نظريات «لا وجبة مجانية»، التي تنص على أنه لا توجد خوارزمية تعلم تتفوق على الأخرى في المتوسط عبر جميع المشاكل، لا تنطبق على الممارسة لأن البيانات الحقيقية ليست موزعة بانتظام.

ومع ذلك، تظل الجدران النظرية غير قابلة للتجاوز أو التخطي. فالذكاء اصطناعي لا يمكنه تحديد ما إذا كان سيتوقف عند مدخلات معينة، أو إثبات اتساقه الخاص، أو التحقق من محاذاته من خلال التحليل الذاتي. بل حتى أن نظامَي ذكاء اصطناعي لا يمكنهما تحديد ما إذا كانا متكافئين وظيفيًا. ففي ورقة منشور حديثا على تعقيد التفسير تُظهر أن **أي تفسير أبسط بكثير من النموذج الأصلي يجب أن يختلف عند بعض المدخلات** "*any explanation significantly simpler than the original model must differ on some inputs*"[21] ، مع نمو التعقيد أسياً مع بُعد المدخل. وإذن فإن حلم الذكاء الاصطناعي الشفاف والقابل للتحقق يواجه استحالة رياضية عند كل منعطف.

وقد شهدت الفترة 2020-2025 تطورًا عاصفًا في كل من نقد الذكاء الاصطناعي والاستجابات الحاسوبية، فالمؤتمرات الكبرى تُظهر نموًا انفجاريًا — NeurIPS 2025 تلقى أكثر من 27,000 ورقة — مع تركيز متزايد على الأسس النظرية ودراسات الانبثاق، وقد تحول الإجماع الأكاديمي بشكل حاسم: **الاعتراضات الرياضية على الذكاء الاصطناعي القوي تُعتبر على نطاق واسع معيبة** "*mathematical objections to Strong AI are widely considered flawed*"[22]، بينما تبقى التحديات العملية هائلة.

وفي حين تقدم الاختراقات النظرية الحديثة منظورات جديدة، فإن اشتغال روساس وآخرين (2024) على الانبثاق الحاسوبي يوفر أطرًا صورية لفهم كيف يمكن للعمليات العيانية إظهار خصائص حاسوبية ذاتية الاحتواء — «برمجيات في العالم الطبيعي». فالأنظمة يمكنها إظهار إغلاق حاسوبي، فهذا بمثابة اقتراح بأن مسارات لقدرات الذكاء الاصطناعي قد تنبثق من تفاعلات معقدة دون تجاوز الحدود الصورية.

ولا يزال الحوسبيون يصوغون دفاعات متطورة ضد الاعتراضات الكلاسيكية، فحجة لوكاس-بنروز تواجه رفضًا شبه عالمي، مع دحضات صورية تُظهر عيوبًا منطقية ورياضية، فكما أظهرت تحليلات متعددة أن مشكلة الاتساق — البشر لا يمكنهم تأسيس اتساقهم الخاص — قُوِّضَت الحجة مقتولةً.

[21]Explanation complexity — recent work (2024)

[22]Academic consensus survey — NeurIPS (2025)

المراجع

1. XPPPPXOOOORRRSSSS
2. Turing, A. (1936). “On Computable Numbers, with an Application to the Entscheidungsproblem”
3. Church-Turing thesis formulation — Stanford Encyclopedia of Philosophy
4. Halting problem limitation — Turing (1936)
5. Melo et al. (2024). “On the Undecidability of Artificial Intelligence Alignment: Machines that Halt”
6. Rice, H. (1953). “Classes of Recursively Enumerable Sets and Their Decision Problems”
7. Gödel, K. (1931). “Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme”
8. Putnam, H. (1960). “Minds and Machines”
9. Hofstadter, D. (1979). “Gödel, Escher, Bach: An Eternal Golden Braid”
10. Complexity theory consensus — survey of complexity theorists (2024)
11. PSPACE-hardness of games — Stockmeyer & Meyer (1973)
12. Quantum computing limitations — Aaronson, S. (2013). “Quantum Computing Since Democritus”
13. Penrose, R. (1994). “Shadows of the Mind”
14. Lucas, J.R. (1961). “Minds, Machines and Gödel”
15. Chalmers, D. (1995). “Minds, Machines, and Mathematics”
16. Consensus on Lucas-Penrose — Internet Encyclopedia of Philosophy
17. Turing, A. (1938). “Systems of Logic Based on Ordinals”

18. Bekenstein, J. (1981). “Universal upper bound on the entropy-to-energy ratio”

19. Davis, M. (2004). “The Myth of Hypercomputation”

20. Physical Church-Turing thesis — Deutsch (1985)

21. Rice’s theorem — Rice (1953)

22. Solomonoff, R. (1964). “A Formal Theory of Inductive Inference”

23. Explanation complexity — recent work (2024)

24. Academic consensus survey — NeurIPS (2025)

25. 4E cognition framework — Clark & Chalmers (1998)