

Out-Evolve

Building the Means of
Discovery and Production

Vincent Boucher

MONTREAL.AI, Montréal, Québec, Canada

20 October 2026 / Revised research manuscript

Abstract

The most consequential output of intelligence may be a capability that makes further discoveries possible. We propose **Out-Evolve**: a research and construction architecture that directs transformational intelligence toward new means of experimentation, production and subsequent invention. Its central engineering target is to *construct a system for which a simpler model is true*, while preserving the useful purpose that motivated the work. A conditional identification result makes the potential leverage precise. An arbitrary scalar response over twenty binary settings requires 1,048,576 exact point queries to determine completely. A constructed platform with a qualified interaction-order bound of two can be identified from 211 exact scalar measurements. A compatible additional setting can double the configuration family with 21 further calibrations. These results concern declared function classes; the physical structure, precision and inheritance guarantees must be earned through construction and qualification. Two executed exact studies examine complementary mechanisms: reopening a retained operation for a successor, and constructing a measurement architecture whose users can support its cost. We then specify a commercial program for SUCCESSOR Ω : develop a family of accepted manufacturing capabilities, retain the means that make them useful, and use those assets and available surplus to help construct a different successor capability. The paper contributes a construction objective, conditional theory, inspectable mechanism studies and a three-stage evaluation protocol. Its ambition is a company whose discoveries repeatedly expand what available intelligence can know, build and economically deploy.

Build what makes the next breakthrough possible.

Then use it to build what comes after.

1 Build what makes the next breakthrough possible

The next breakthrough may be a new way to make breakthroughs.

An intelligent system can discover a material, design a machine, or improve a process. Its more consequential output may be an instrument that makes a class of materials experimentally accessible, a production architecture that makes a family of machines economical, or a method that helps construct the next instrument. Each achievement changes the starting conditions for subsequent intelligence. The valuable inheritance is a new capacity to act, learn and build.

Transformational intelligence discovers a fundamentally different game to play by constructing better conditions for achieving the original purpose. Out-Evolve directs that intelligence toward capabilities from which entire families of discoveries and businesses can grow.

The central engineering move is unusually demanding: *construct a system for which a simpler model is true*. A process may be difficult to understand because useful state is hidden, experiments confound several causes, or components interact in uncontrolled ways. An invention can change those conditions. A new reference path can make a quantity identifiable. A stabilized interface can make an interaction repeatable. A different physical decomposition can make the behavior of a whole apparatus follow from a manageable number of measurements. Intelligence has then changed the system it must understand.

The possible difference is large. An arbitrary response over twenty binary settings has 1,048,576 independently variable values. A platform whose useful response is guaranteed to contain only single-setting and pairwise interactions has 211 coefficients. Appropriate exact measurements determine its entire response family. The mathematical contrast is conditional; the difficult and valuable invention is the useful architecture that earns the guarantee. Its construction, precision and original-purpose requirements are explicit in section 2.

Now give the capability a successor. A production cell that serves external customers can also manufacture some of the fixtures, calibrated structures and test apparatus needed by a different production cell. Its retained implementation becomes construction material. Accepted output provides receipts; resources remaining after obligations can finance further construction. Knowledge, instruments, productive hardware and capital begin to reinforce one another, provided each transition is technically valid and economically supported.

This is the proposed business of SUCCESSOR Ω : the Intelligence Leverage Company. Its product is a qualified means of producing valuable outcomes across a family. Its strategic question is: *which capability would make the next valuable capabilities possible, and what would make constructing it worthwhile now?*

1.1 The progression beyond the foundations

Transformational Intelligence makes the formulation of a problem an object of discovery, with an explicit bridge back to the user's purpose [5]. *Invent the Economy* makes participation, technical construction and cash timing part of the same design problem [4]. Out-Evolve applies both ideas to the creation of *discovery and production capacity*: construct a useful regime, identify what becomes possible within it, and use the resulting capability to help construct its successor.

The progression has three linked achievements. First, **a new capability** changes what can be observed, tested, controlled or produced. Second, **a capability family** supports fresh purposes through an identified common mechanism. Third, **a successor capability** uses retained technical means and economically available resources to acquire a different accepted function. Naming a family does not establish transfer; each link needs its own evidence.

What this manuscript provides. A conditional theory of making response families economical to identify; a construction architecture joining physical structure, evidence, inheritance and adoption; two executed exact mechanism studies; a concrete industrial proposal; and a protocol for a decisive three-stage test. The new central theory concerns a declared information class. The studies concern supplied synthetic grammars. The industrial program is proposed. These evidence levels remain distinct throughout.

1.2 The unit of progress

Let K denote an installed, qualified stock of capabilities. For a purpose u and operating budget b , define

$$\mathcal{F}_b(K) = \{u : \text{an admitted use of } K \text{ delivers } u \text{ at its required quality within } b\}. \quad (1)$$

The budget can be a vector of time, money, computation, energy and other scarce inputs. Constructing K' can change this operating frontier. It may expose new information, make an existing outcome economical, or supply a more useful realization of an operation. The complete comparison charges the transformation from K to K' as well as subsequent use. A capable planner allowed the same construction may already reach the same result; the relevant advantage must be demonstrated.

A family-level claim compares fresh purposes at a declared total budget, with the same base intelligence and access to the same admissible resources. The outcome is accepted discovery or production. The construction can be intelligent even when it makes the final task easy. That is the intended leverage: *make the next valuable outcome require less invention because the right invention has already been made.*

2 Construct a system for which a simpler model is true

A powerful way to improve discovery is to change the system being investigated. Instead of asking an ever larger model to understand an opaque process, construct interfaces, operating conditions and component boundaries that make its behavior identifiable. The retained invention is then more than a prediction: it is a working arrangement under which a whole family of predictions becomes justified.

Construct a system for which a simpler model is true.

This is a physical and computational design objective. Isolation, controlled preparation, stabilized operating conditions, calibrated reference paths and explicitly limited couplings can remove interactions that would otherwise make every new configuration a separate experiment. The mathematical case below shows what such a construction could buy. It does not assume that these conditions can be obtained cheaply, or that a sparse fit establishes them.

2.1 A precise change in the experimentally knowable family

Consider a platform with n binary configuration choices $x \in \{0, 1\}^n$ and a scalar, repeatable response $f(x)$. A response instrument must determine $f(x)$ for *every* admitted configuration, so a later investigator can ask a new question without reopening the original identification problem. One experiment returns the exact scalar response at one chosen configuration. This is a worst-case identification contract, not the weaker task of finding one satisfactory setting.

For an unrestricted unknown function, every configuration can carry independent information. Now suppose a qualified construction produces a platform whose response belongs to

$$\mathcal{V}_{n,k} = \left\{ f(x) = \sum_{\substack{S \subseteq [n] \\ |S| \leq k}} \theta_S \prod_{i \in S} x_i : \theta_S \in \mathbb{R} \right\}, \quad D_{n,k} = \sum_{j=0}^k \binom{n}{j}, \quad (2)$$

where $0 \leq k \leq n$. The restriction says that no response term depends jointly on more than k switches. It concerns the deployed response, including its readout, rather than the visual arrangement of its components. A nonlinear shared bottleneck can violate it even when the parts look modular.

Theorem 2.1 (Conditional response-family query separation). *Under exact, scalar point-query access, identifying an arbitrary function on $\{0,1\}^n$ requires 2^n experiments in the worst case. Identifying every function in $\mathcal{V}_{n,k}$ requires exactly $D_{n,k}$. These lower bounds allow unlimited computation and deterministic adaptive query selection. They are attained by querying all configurations in the first case, and all configurations with at most k enabled switches in the second.*

Proof. If an unrestricted function has not been queried at some configuration, its value there can change without changing the transcript. For the structured class, let $\mathbf{1}_T$ enable exactly the switches in T . The $D_{n,k}$ readings give

$$\theta_S = \sum_{T \subseteq S} (-1)^{|S|-|T|} f(\mathbf{1}_T), \quad |S| \leq k. \quad (3)$$

This determines the whole response. Conversely, along any adaptive all-zero transcript of fewer than $D_{n,k}$ point queries, the evaluation equations have a nonzero null vector in the $D_{n,k}$ -dimensional coefficient space. The zero response and the corresponding nonzero response produce the same transcript, so identification is impossible. Details and the structural-certification limit appear in appendix A. $\square$

The polynomial representation and Boolean-lattice inversion are classical [14, 22]. Their use here supplies a concrete engineering target: construct the validity conditions of a compact response family, then retain an experimentally identified member of that family. The theorem is not a new interpolation algorithm or evidence that Out-Evolve has already constructed such a physical platform.

For $n = 20$ and $k = 2$, the derived counts are

$$2^{20} = 1,048,576, \quad D_{20,2} = 1 + 20 + 190 = 211. \quad (4)$$

The first number is the unrestricted worst-case number of exact scalar queries; the second is sufficient and necessary for the unrestricted quadratic coefficient class. This is a mathematical comparison between declared information classes, not a measured speedup, an estimate of current industrial practice, or a claim that every quadratic physical model has 211 independent parameters.

The retained coefficients describe all configurations of the qualified platform. They can support later response queries, later utilities based on the identified response and known configuration data, operating-envelope calculations and candidate selection without additional identification experiments. They do not supply unmeasured material properties, justify a changed plant, or certify the final product by themselves. Optimizing a general quadratic response can remain difficult; a compact model is not automatically an efficient optimizer.

2.2 The construction must earn the assumption

A permissible route to equation (2) is an architecture with a justified decomposition

$$f(x) = \sum_{A \in \mathcal{H}} f_A(x_A), \quad |A| \leq k. \quad (5)$$

Each component may have an arbitrary response on its own inputs. If the physical composition and readout genuinely implement the displayed sum, the global response belongs to $\mathcal{V}_{n,k}$. Enforcing and inspecting those interfaces is part of the invention. Merely labeling components as local does not establish the sum: shared heat, supply limits, drift, feedback or nonlinear aggregation can introduce higher-order terms.

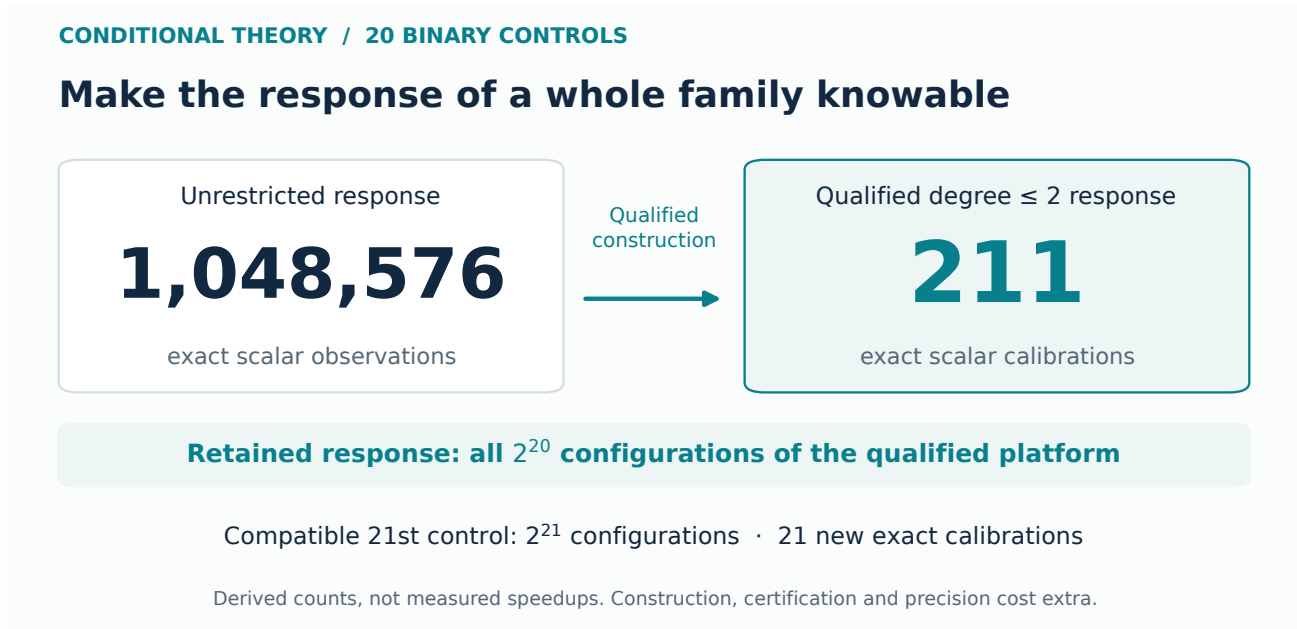


Figure 1: Conditional exact identification and qualified succession. The two initial counts compare different declared response classes. The construction must preserve the useful purpose and establish the structural restriction. Physical construction, qualification and measurement precision remain charged separately.

Construction may change the substrate, but it may not replace the human purpose with the easier calibration task. A candidate platform must still meet the original requirements for delivered output, quality, operating range and resources. Its purpose bridge must show that acceptance using the constructed model implies acceptance of the actual deployment, including properties that the modeled response does not describe. A system that becomes easy to predict by losing the required capacity is a failed construction.

For example, a proposed fluid-control manufacturing and inspection platform might introduce a fixed-topology characterization station with stabilized operating conditions. Let q denote independent flows relative to a specified return port and p the corresponding pressure differences. Choose commands $q = q_0 + Vx$ so every admitted configuration respects mass balance and the operating envelope. If that qualified envelope supports $p = Rq$ with fixed R , pressure responses are affine and hydraulic dissipation $q^\top Rq$ is quadratic. The pumps, sensors, control, stabilization and operating restrictions are real construction obligations. Turbulence, varying viscosity, switching topology or uncontrolled valve behavior can invalidate the model. Whole-part pressure, flow and quality acceptance remains necessary. This is a proposed conditional route, not a manufactured or experimentally validated result; an existing specialist hydraulic model may already provide efficient identification, in which case the unrestricted 2^n lower bound is not its appropriate baseline.

The discovery target is therefore specific. Search admitted physical designs and operating regimes for a purpose-preserving structural certificate; construct the selected interface; calibrate its remaining unknown parameters; retain the resulting response instrument; and expose any later request that exceeds its scope. The finite calibration plan is explicit. Finding a useful physical construction and earning its certificate remain explicit research and engineering tasks.

2.3 A capability ladder with an explicit inheritance boundary

An unchanged constructed platform can support a later purpose expressed through its retained response. A larger platform requires a stronger bridge. Suppose a qualified extension adds one switch, retains degree at most k , and preserves the old response whenever the new switch is off. Then only

$$D_{n+1,k} - D_{n,k} = \sum_{j=0}^{k-1} \binom{n}{j} \quad (6)$$

new scalar calibrations are needed, and this is minimum for the full compatible coefficient class. The old coefficients are inherited; only terms involving the new switch are unknown. Thus a qualified quadratic extension from 20 to 21 switches doubles the number of configurations and requires 21 additional exact scalar calibrations. Neither the added hardware nor its compatibility certificate is free. Changing the old baseline, introducing a higher-order interaction, or violating repeatability invalidates this inheritance claim.

This gives a concrete meaning to a capability family: a constructed platform, its preserved structure, and a calibrated response support different later questions and an explicitly bounded successor. There is no claim that the new component or its compatible architecture was autonomously invented by the present theorem.

2.4 Precision and complete cost decide whether to build it

Exact-query counts can conceal demanding precision. For the calibration design in equation (3), read errors individually bounded in magnitude by ε can amplify to 721ε at a 20-switch configuration when $k = 2$. A separately justified bound ρ on the sum of magnitudes of higher-order coefficients adds at most ρ to that response error. Neither bound can be obtained simply by fitting the 211 readings well. The appendix derives the exact calibration amplification and an indistinguishable high-order counterexample.

The business decision must count discovery, failed constructions, hardware, qualification, calibration to the required precision, computation, integration, maintenance and final acceptance. Write the new campaign cost as

$$C_{\text{new}} = C_{\text{search}} + C_{\text{build}} + C_{\text{qualify}} + C_{\text{calibrate}} + C_{\text{solve}} + C_{\text{deploy}} + C_{\text{maintain}}. \quad (7)$$

It should be compared with the strongest qualified alternative serving the same purposes over the same horizon, including an already structured specialist method and the option not to build. Only in the special case of equal exact-query prices and otherwise equal costs does the observation saving alone provide a construction allowance of $(2^n - D_{n,k})c$ for query price c . No such equality is assumed for a real facility.

The transformational move is to make a class of useful systems easier to know and modify by changing how it is constructed. The experimentally economical family is earned through physical structure, qualified access and retained evidence. The resulting advantage depends on the working construction and its qualified scope.

3 A discovery system that constructs its next advantages

Out-Evolve is an outer research and construction architecture. Its candidate is an executable change to the means of achieving a purpose: a representation, an experiment, an instrument, a product architecture, a process or an adoption arrangement. Search engines, coding agents, scientists and engineers can implement its construction steps. The architecture directs their work toward a capability with consequences beyond its first use.

3.1 From an obstacle to an invention requirement

A binding obstruction can tell the constructor what must change. Two observationally indistinguishable worlds call for a discriminating experiment. Uncontrolled interactions call for a different interface or a qualified model of their joint effects. A failure to fund a shared instrument calls for a change in the relationship between its full cost and beneficiaries' qualified alternatives. An opaque retained operation can call for reopening its implementation.

Table 1: Examples of construction targets. The target follows from a qualified diagnosis; the realization still has to be invented and checked.

Binding restriction	What the constructor must supply	What a successor can inherit
Hidden purpose-relevant state	A measurement or intervention that distinguishes it	A calibrated information channel
Unmanageable interactions	An architecture or operating regime with justified structure	A reusable response instrument
Missing qualified operation	A composite operation or constructed actuator with required semantics	Inspectable implementation and its valid scope
Unsupported shared cost	A construction and purchase arrangement that survive alternatives and timing	A funded, deliverable capability
Costly construction of the next instrument	Reusable tooling, designs and qualification procedures	Productive means for a different capability

A proof of infeasibility is particularly useful because it narrows the invention requirement. For example, let existing activities have outputs a_j and costs c_j . If $\pi \geq 0$, $\eta \geq 0$ satisfy

$$\pi^\top a_j \leq \eta c_j \quad \text{for every } j, \quad \pi^\top b > \eta B, \quad (8)$$

then their nonnegative combinations cannot supply required output b within budget B . An added activity can defeat this witness only if $\pi^\top a_{\text{new}} > \eta c_{\text{new}}$. The inequality becomes an engineering specification. Other witnesses may remain; sequential or nonlinear operations require their own models. This use of proofs to direct construction builds on column generation and logic-based decomposition [11, 16].

Most early science offers uncertain diagnoses rather than proofs. The constructor should then purchase an informative experiment, expose competing explanations, and distinguish a falsified candidate from an unresolved search. The architecture accommodates both. Its requirement is that the status of the evidence determines what may be inferred from it.

3.2 The construction cycle

1. **Fix the valued purpose.** Record the useful output, quality, permissions, resource limits and time horizon. Identify which inherited means may change.
2. **Find the consequential restriction.** Use capable existing methods. Return an accepted solution, a qualified obstruction, or a testable hypothesis about what blocks progress.
3. **Design the missing capability.** Search representations, experiments, interfaces, hardware and operating arrangements. Admit complementary bundles when no component works alone.
4. **Construct and qualify it.** Build the realization, establish its assumptions, and show how its acceptance implies the original useful outcome. Evaluate full discovery and deployment cost.

5. **Recompute the world it changes.** Recheck alternative methods, customer alternatives, uncertainty, liabilities and maintenance. An invention can improve competing routes as well as its own.
6. **Freeze and transfer.** Give a different investigator fresh purposes after freezing the artifact. Retain the implementation, calibration, evidence, costs and legitimate adaptation rights that actually support transfer.
7. **Commission a successor.** Use the capability to construct a different accepted capability. Compare the complete campaign, including unsuccessful construction and external inputs.

These steps need not be serial in implementation. Alternative designs and experiments can proceed in parallel, while evidence and resource commitments remain accountable. A prototype may disprove the original diagnosis. A customer may require a different quality envelope. A qualified alternative may remove the business case. Those discoveries change the next construction decision.

3.3 What compounds

The retained asset has an executable realization, a semantic or empirical contract, an operating envelope, a resource account, and the access needed to use and modify it. A useful form is

$$(\theta, x) \in \mathcal{D}_K \implies \text{Exec}_\theta(K, x) \in R_K(\theta, x), \quad \text{Cost}(K, x) \leq c_K. \quad (9)$$

Here θ represents admitted uncertainty. Exact identities, bounded-error statements and measured reliability are different contracts. A successor can combine valid contracts, inspect internal structure, recalibrate or replace an inherited implementation. It should retain the discovery without becoming trapped by the first packaging of that discovery.

Useful capabilities can also be complements. A programmable instrument without valid measurement may produce uninterpretable data; a calibrated sensor without a controllable experiment may leave the scientific question unresolved. Evaluating only immediate payoff would reject both. Out-Evolve therefore searches bounded construction bundles and stages their financing around acceptance milestones. A claim of future complementarity still needs a concrete joint realization and an affordable path to it.

3.4 A productive organization of agents

Computer use connects intelligence to existing design, simulation, laboratory and business software. Persistent agents can maintain hypotheses, designs and experimental records across long campaigns. Recent work explores both long-horizon computer-use context and persistent multi-agent scientific environments [7, 19]. Those capabilities make the architecture more implementable; they do not establish that an entire research or industrial occupation has been automated.

The proposed organization gives agent teams distinct constructive roles: identify restrictions, propose mechanisms, design discriminating experiments, compile implementations, challenge evidence and assemble commercially acceptable work packages. Shared instruments and capability contracts make their work composable. A discovery earns continued resources through reproducible usefulness to another project. Agent population becomes valuable when it produces interoperable means of progress.

The long-run objective is an organization whose accumulated instruments and productive assets let new participants begin beyond the previous generation's limits. Its inheritance includes machines that work and agreements that can be performed, alongside code and ideas.

4 The flagship: productive capacity that helps construct its successor

Build a production capability that makes the next production capability easier to invent, construct and qualify. This is the proposed industrial challenge for SUCCESSOR Ω . The first customer receives accepted machined bodies and components. The company retains the fixtures, measurement methods, process knowledge and reusable designs that made a whole family economical. A subsequent project must use those assets to construct a genuinely different capability, while continuing to serve outside customers. The business compounds only when that succession produces useful capacity at a competitive full cost.

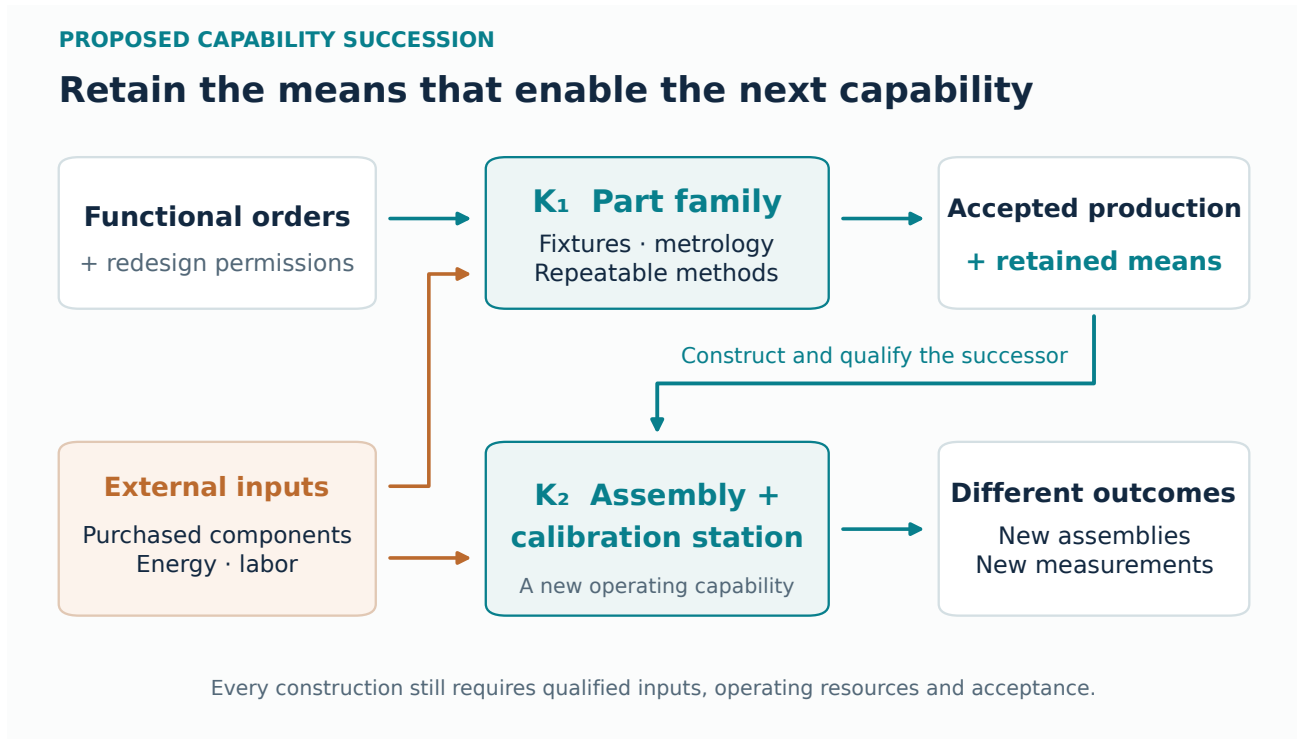


Figure 2: Proposed industrial succession. Accepted production supplies retained technical means and resources earned after obligations. Both contribute to a different successor capability. Purchased components, energy, skilled work and fresh qualification remain necessary at each stage.

4.1 A first purchase with room for invention

The proposed first buyer is an industrial fluid-handling instrument manufacturer that repeatedly needs small batches of configurable machined manifold bodies. Start with a declared aluminum-alloy family, conventional machining, bounded dimensions and port arrangements, specified sealing interfaces, and a qualified water-service operating range. Public reference designs can establish the initial envelope. An actual buyer must fix the applicable material, cleanliness, dimensional, flow, pressure, reliability and delivery requirements before construction.

The first work package is concrete: accepted first articles and a batch of each of several agreed variants, traceable inspection and functional-test records, and a priced procedure for later variants inside the envelope. One later variant remains hidden until the shared capability is frozen. A partner machine shop can supply the initial machine and operators; SUCCESSOR Ω builds and owns the agreed reusable tooling, test interface and deployment knowledge.

The purchase must permit functional redesign. The buyer fixes the external purpose and interfaces while allowing changes to internal geometry, workholding, test access and modularization. That con-

tractual freedom is part of the construction: a build-to-print obligation can forbid the very change that would make the family economical. The joint question is which acceptable product architecture, production interface and order commitment make these particular outputs worth producing.

4.2 Construct the conditions under which the model is useful

The cell must expose consequential state: references for locating the part, clamping conditions, tool condition, temperature, and the conditions of the functional test. Candidate constructions include accessible reference surfaces, instrumented or repeatable fixtures, calibrated probing, controlled resets between operations and separately testable modules. Tactile-fixture design and measurement-assisted control are established engineering practices; their presence alone is not a new discovery [1, 20].

The constructed-response theory in section 2 supplies a precise target for a subsystem of this cell. The proposed flow-test example explains how a justified operating regime can simplify the response while preserving real coupling. The cell must separately qualify machining, assembly and every other required property. Each new geometry may require fresh characterization. An interface earns its place only when the full construction and qualification cost is justified by the experiments or failures it eliminates.

4.3 Let the orders select the capability

For architecture a , write the fully costed campaign as

$$C_a = D_a + I_a + Q_a + \sum_i N_i v_{ia} + O_a. \quad (10)$$

Here D_a includes discovery and failed candidates; I_a installation and purchased equipment; Q_a qualification; v_{ia} accepted-unit production cost, including yield loss; and O_a the remaining delivery, maintenance, contractual and financing obligations. Machine use, skilled labor, metrology, feedstock, energy, waste and support cannot disappear into an unpriced partner contribution. Working capital has a separate cash schedule; its financing cost belongs in O_a , without counting the funded purchases twice.

A proposed order can quote an accepted setup fee F_i and accepted-unit price u_i , giving $P_i = F_i + N_i u_i$. Prices, quantities and payment dates require actual agreement. Joint viability requires

$$C_a + M \leq \sum_i P_i, \quad \sum_{i \in S} P_i \leq B_a(S) \quad \text{for every relevant } S, \quad (11)$$

with individual budget and adoption limits also satisfied. Here M is the desired retained surplus and $B_a(S)$ is the best qualified outside cost for the same obligations, recomputed when the invention changes those alternatives. Customer value and consent are additional evidence, not consequences of this inequality. This applies the joint-construction discipline of *Invent the Economy* [4].

The architecture is selected together with the accepted orders: shared test access or a standardized interface can lower qualification costs; a narrow exclusive specification can destroy that sharing. Start only if the resulting package also has a funded cash path through its agreed failure and termination cases. An acceptance invoice cannot fund earlier machining. Earned receipts from a severable first delivery can support later work only after its remaining obligations are protected. No price, customer commitment or commercial margin is claimed here.

4.4 The successor must do a different job

Freeze the first capability K_1 before revealing the successor specification. The proposed K_2 is an assembly-and-calibration station for complete fluid-control modules containing valves, sensors and calibrated functional assemblies. K_1 supplies qualified fixture plates, probe mounts, manifold hardware and relevant manufacturing procedures. Purchased actuators, sensors, reference instruments, control

electronics and any robot complete the bill of materials; installation, calibration and skilled integration remain charged.

K_2 must assemble and functionally qualify complete modules, a different task from machining their components. Its acceptance concerns assembly reliability, calibrated function and throughput across fresh module variants. Compare the full two-project campaign with a capable integrator allowed modern CAD/CAM, modular fixtures, in-process measurement, supplier procurement and reformulation. Also measure the incremental cost of K_2 with and without legitimate access to K_1 . Sunk first-project costs still count in the full campaign; selective amortization cannot manufacture an advantage.

The decisive commercial evidence is a paid first delivery, an independent repeat purchase that benefits from the retained capability, and a successor meeting fresh requirements at a defensibly lower full construction cost. A larger library, another agent, or an impressive demonstration supplies none of these facts by itself.

4.5 Ownership and the route to consequential scale

The retained business assets are qualified generic designs, integration knowledge, reusable process software, calibration methods, deployment rights and owned productive hardware. Customers buy accepted outputs or defined capacity; price is not tied to preserving avoidable engineering work. Public principles remain available to competitors. A durable advantage requires demonstrated cost, quality, delivery or integration performance and valid rights. The ownership design retains all parent equity with its current owner; separately negotiated project or asset rights can support specific deployments.

Agents and computer use can connect specification, CAD/CAM, simulation, procurement and inspection software, maintain the evidence record and investigate alternative constructions in parallel. They serve the capability-building process. Human operation, approval and intervention remain visible costs wherever required.

A further candidate capability is an automated thermal-fluid research station. Machined structures from K_1 and calibrated flow-control assemblies from K_2 become inputs to an instrument that tests new heat-transfer devices under measured operating conditions. It adds experimental execution and characterization to the earlier manufacturing and assembly functions. Those experiments could support new cooling or energy-conversion equipment, subject to separate qualification. This proposed dependency chain still purchases reference instruments, drives, power, sites and specialized expertise. Energy-system impact is counted when a complete installation is commissioned. The long-run asset is a growing stock of machines, instruments and qualified procedures that makes further useful capacity cheaper to construct.

This ambition has serious industrial precedents. Hadrian describes AI-supported manufacturing and dedicated production capacity; Fab2 explicitly pursues components, machines and factories for making further factories [9, 13]. The proposed research distinction is a measured ability to construct more tractable physical interfaces and cheaper qualified successors at a matched intelligence and total-resource budget. The current exact laboratories do not establish that industrial result. The flagship makes the result specific enough to attempt, purchase, compare and reject.

5 Two exact studies of the mechanisms

The central physical construction remains a research target. Two executed, reproducible laboratories make two mechanisms of the proposed architecture inspectable: how a retained implementation supports a successor, and how a construction changes the possibility of economically coordinated use. Complete algorithms, proofs, negative controls and accounting are in appendices B to E and the accompanying package.

5.1 An inherited operation can be reopened

Exact search starts with adjacent reversible shears $A = I + aE_{12}$ and $B = I + bE_{23}$, their inverses and a supplied endpoint requirement. It constructs

$$M = ABA^{-1}B^{-1} = I + abE_{13}, \quad (12)$$

which transfers between nonadjacent coordinates while restoring the others. The minimum is four primitive actions. The implementation and its evidence are frozen before a further primitive $C = I + cE_{34}$ is introduced to the successor stage.

The next transfer, $I + abcE_{14}$, costs ten primitive actions if M is available only as a sealed operation. Reopening its implementation and using valid commutations permits an eight-action realization. One stored realization suffices. Unrestricted fresh search also reaches eight. The experiment isolates the value of inspectable inheritance relative to opaque reuse; its algebra is classical [17, 25], and its task sequence is authored rather than an independently hidden benchmark.

The practical lesson is to preserve what makes the capability valid and adaptable. Its internal structure may be more useful to the next constructor than its original interface. Universal symbolic checks, exact lower bounds and independent computational checks establish the reported result within the stated alphabet.

5.2 A cheaper capability can remain economically unavailable

Three beneficiaries require three quantities measured. In the supplied legacy grammar, each pair can obtain its measurements for 12 resource units, while serving all three costs 20. Stable payments must charge each pair at most 12. Summing those constraints bounds total recovery by 18. The joint service has no stable cost-recovering allocation under that model.

The obstruction specifies a construction requirement. Exact measurement semantics establish a four-read lower bound; the recovery witness imposes an installation budget. Among 64 selector architectures, a one-link construction changes the joint cost to 17. Recomputing all qualified alternatives permits payments totaling 17. A three-link route reduces the legacy cost to 19 and still fails the adoption condition.

The continuous frontier is more revealing than the worked point. If each read costs $q > 0$, the minimum technical cost and maximum stable recovery are

$$C^*(q) = \min(5q, 4q + 1), \quad \Gamma^*(q) = \min\{C^*(q), 4.5q\}. \quad (13)$$

The new route improves technical cost when $q > 1$; stable recovery of the minimum cost becomes possible only at $q \geq 2$. There is a genuine interval of technical progress without economic feasibility under the stated participation model. The construction must improve the relationship between the shared capability and the alternatives its users can choose.

Under full grammar access, a capable cost optimizer selects the same 17-unit design. An equally qualified competitor that can reproduce it at the same full cost leaves zero demonstrated margin. Independently biased routed measurements also remove the information advantage. These controls locate the actual result: a construction and acceptance condition, rather than an observed commercial success.

Evidence boundary. Both studies use supplied grammars and exact synthetic costs. Their independent checkers use different mathematical implementations. They establish the stated local mechanisms. The proposed physical discovery regime, integrated industrial program, customer adoption and superiority to competing discovery systems require the experiments in section 7. The two studies are complementary; a failed attempt to combine them is preserved in appendix F.

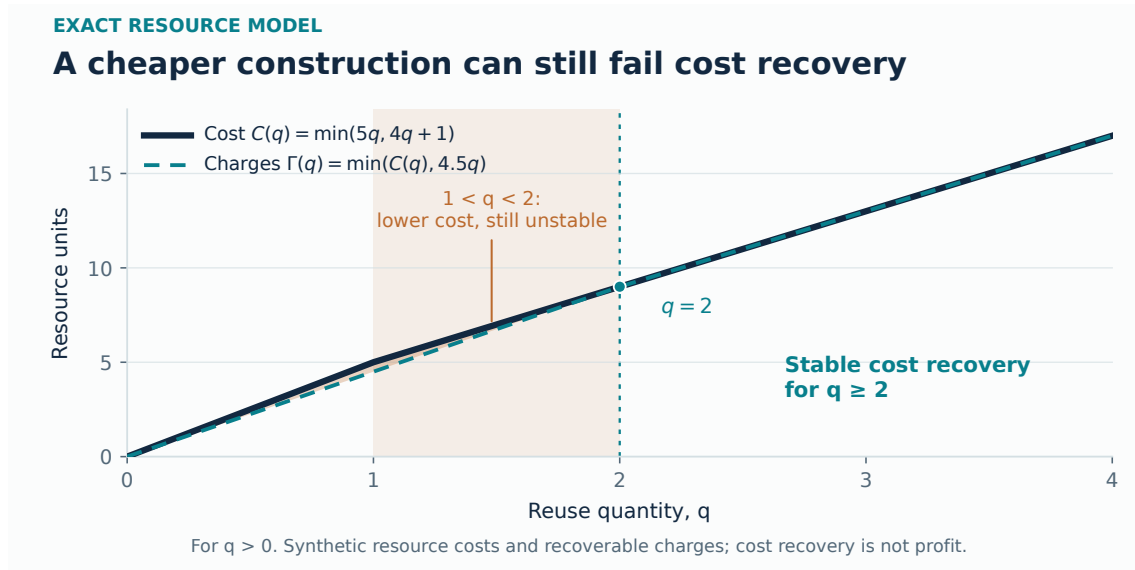


Figure 3: Exact synthetic functions. The interval $1 < q < 2$ contains a technically better architecture that still cannot recover its cost through stable payments. At $q \geq 2$, the resource-cost allocation becomes feasible. The functions do not establish customer demand or positive profit.

6 AlphaEvolve and the frontier this program targets

AlphaEvolve demonstrates that strong models, evolutionary search and executable evaluation can create valuable algorithmic discoveries. Its documented search already includes whole programs, constructor functions, bespoke search algorithms and meta-prompt evolution [21]. Its mathematical follow-up studies 67 problems and reports constructions and some generalizations [10]; DeepMind also reports operational applications across computing and industry [2].

Out-Evolve proposes a broader *commission*: construct a capability that changes how a useful family can be investigated or produced, establish its valid deployment and economic use, then demonstrate its role in constructing a successor. The distinction is the object and acceptance conditions of the campaign. A broad evolutionary engine can implement parts of this commission. If it succeeds under the same conditions, that is evidence for the construction objective and for that engine.

Table 2: Selected emphases in the cited discovery work and the proposed commission. These are not capability limits or an empirical ranking.

Question	Emphases in cited discovery work	Out-Evolve commission
What is produced?	Programs, constructions, learned tools and search procedures	A qualified experimental or productive capability and its operating regime
What is inherited?	Implementations, search history, abstractions and libraries	Those assets plus instruments, calibration, physical means and deployment rights
What demonstrates breadth?	Performance across the evaluated problems or families	Fresh purposes served through a frozen capability and a different accepted successor
What determines value?	Task metrics and demonstrated application benefits	Accepted outcomes after full construction, qualification, operation and adoption costs
What would establish leadership?	Results on the system’s stated evaluations	A stronger complete campaign against capable, equally permitted alternatives

The surrounding literature already develops important ingredients. EvoX evolves search strategies; meta-design produces constructors for families of quantum experiments; DreamCoder and SkillSmith develop reusable program or tool libraries; Quartz discovers and verifies circuit transformations [3, 8, 18, 26, 27]. Physical design that simplifies learning or control has precedents in morphological computation, while decomposability has a long relationship with comprehensibility [15, 24]. Shared research infrastructure is also an established organizational target [6].

The specific research synthesis here is to make useful representability, construction, inheritance and adoption jointly explicit, and to demand evidence at their interfaces. The low-order response theorem gives an invention target. The mechanism studies give exact local checks. The industrial challenge gives the proposal a consequence that can be built and judged. This manuscript reports no head-to-head run of AlphaEvolve.

7 The succession challenge

The strongest demonstration would make the reader observe a succession of new abilities. Commission three stages under a common purpose and accounting framework, with later specifications withheld until the relevant artifact is frozen.

Table 3: Proposed evaluation protocol. Every row requires external acceptance; these are future experiments, not reported outcomes.

Stage	Required achievement	Evidence that makes it consequential
1. Construct	Build a purpose-preserving capability with justified structure	A physical or operational realization, its uncertainty envelope and complete cost
2. Transfer	Use the frozen capability on fresh, meaningfully different requirements	Independent acceptance; removed-capability comparison; failures and costs reported
3. Succeed	Construct a different useful capability using retained means	A new accepted function, complete campaign comparison and accounted external inputs

Hold the base model and its total inference budget fixed in the principal comparison. Give alternatives the same initial information, legitimate resource access, permission to reformulate and opportunity to construct instruments. Include a strong specialist engineering team, a capable general agent system, and an evolutionary coding system with a sufficiently broad editable representation. AlphaEvolve should be named only when actually run; an open implementation is a separate comparator. Include the best existing structured method and the option to buy the required capability.

Measure accepted outcomes and their full costs; later-task acceptance under the frozen scope; successor construction cost; the rate at which invalid assumptions are detected; and, for a commercial claim, collected receipts and contribution after obligations. Charge unsuccessful designs, hardware changes, model calls, human intervention, precision, integration and maintenance. A two-project advantage must survive both incremental successor accounting and the full cost of the first project.

The family is defined before selection. Variants should differ in requirements that matter to the buyer or investigator; relabeling the same problem does not establish breadth. Later tasks should not appear in training, design prompts or hidden evaluation feedback. A neutral custodian can reveal their specifications after the freeze. Publish enough implementation and evidence for another group to reproduce the claimed bridge, subject to legitimate customer and intellectual-property restrictions.

The decisive public artifact would show a clear chain: a useful outcome initially blocked; a constructed change that removes the restriction; independent success on fresh purposes; and a successor

that obtains a different function because the first capability exists. The specific measured gain can then carry the headline. The experimental design makes leadership contestable and therefore meaningful.

8 The civilization-scale ambition

The distant ambition is an expanding capacity to construct useful capacity. Scientific instruments improve access to phenomena. Qualified processes make more exact or efficient components. Those components support better instruments, factories and energy systems. Productive assets create the means to extend that cycle. The return of an invention can include a new set of inventions that can now be attempted economically.

That objective reaches beyond the first manufacturing program. A new characterization method could make a materials family searchable. A more observable process could reduce the intelligence needed for reliable control. A new machine architecture could lower the cost of constructing scientific equipment. An improved production method for energy hardware could increase the resources available to further industry. Each is a candidate capability chain, with an engineering and adoption problem at every link.

The compounding frontier is the set of valuable capabilities we can afford to construct next. Intelligence becomes transformational when it changes that set.

At the largest horizon, the program could extend from terrestrial instruments and factories to space-based industrial infrastructure and, ultimately, stellar-scale useful energy capture. Reaching that horizon would require advances in energy conversion, materials, precision production, autonomous construction, transport and coordinated investment. A line of software cannot supply those assets. A sustained succession of accepted productive capabilities could contribute to building them. The present paper gives that ambition a unit of progress rather than a forecast or a growth law.

Physical capacity and financial capacity must both remain real. Energy, materials, land or orbital access, manufacturing lead times, reliability, maintenance and demand constrain each expansion. The next binding bottleneck can change. A successful system allocates its intelligence to discovering which construction now matters most, including cases where maintaining a reliable existing capability creates more value than pursuing a new one.

A future civilization of agents could inherit instruments, experimentally grounded knowledge, productive hardware and legitimate institutions through which new participants can act. Its scale would be measured in accepted science and useful production. Its enduring achievement would be to leave the next generation more capable of making what it needs.

9 Conclusion

Out-Evolve proposes intelligence that constructs the conditions of further discovery. Its strongest move is to make a useful system more knowable by changing how it is built. The resulting capability can support a family of purposes and, with a valid inheritance bridge, help construct a successor. The commercial counterpart is to design the production and purchase arrangement together, so accepted value supplies the resources for further construction.

The conditional theory identifies an information advantage and the structural evidence required to earn it. The exact studies expose two mechanisms of the proposed architecture: inspectable inheritance and economically viable coordination. The industrial challenge joins them through construction, independent use and full accounting.

Build what makes the next breakthrough possible. Then use it to build what comes after.

Technical Appendices

Proofs, exact laboratories, failure cases and reproduction

A Constructed structure: identification, inheritance and precision

A.1 Query model and the exact lower bounds

Let $\mathcal{X}_n = \{0, 1\}^n$. A deterministic adaptive algorithm chooses a configuration using its earlier scalar responses, receives the exact value at that configuration, and must output a representation determining the response everywhere on $\mathcal{X}_n$. Computational work is unrestricted. The guarantee is worst-case exact identification over the stated function class. The lower bounds also apply to randomized algorithms required to be correct for every internal random seed, by fixing a seed. They are not claims about approximate or distributional learning, noisy query complexity, or the cost of finding one satisfactory configuration.

The unrestricted class has 2^n independent values. Follow the algorithm on the identically zero function. If it stops before querying every configuration, choose an unqueried z and define $h(z) = 1$ and $h(x) = 0$ elsewhere. The zero function and h have identical transcripts, including every adaptive choice, but require different outputs.

For the structured class, write $x_S = \prod_{i \in S} x_i$ with $x_\emptyset = 1$. The functions x_S are linearly independent on the Boolean cube. Indeed, evaluating at $\mathbf{1}_T$ gives the subset-incidence matrix $1_{S \subseteq T}$; ordered by cardinality, it is triangular with unit diagonal. Therefore the subspace spanned by $|S| \leq k$ has dimension $D_{n,k}$.

On an adaptive all-zero transcript with $m < D_{n,k}$ queries, the evaluation matrix has m rows and $D_{n,k}$ columns. It has a nonzero coefficient vector in its kernel. The resulting nonzero polynomial gives every observed value zero, follows the same adaptive path, and differs somewhere because the basis is independent. This proves the structured lower bound without assumptions about the algorithm's internal representation. Additional known physical restrictions can reduce the admissible dimension; the $D_{n,k}$ necessity statement concerns the full coefficient class, not every physically realizable subclass.

For sufficiency, measure $y_T = f(\mathbf{1}_T)$ for $|T| \leq k$. Since

$$y_T = \sum_{S \subseteq T} \theta_S, \quad (14)$$

Boolean-lattice Möbius inversion gives equation (3). Direct substitution also proves it: the coefficient of θ_R in the proposed expression for θ_S is

$$\sum_{R \subseteq T \subseteq S} (-1)^{|S|-|T|} = \begin{cases} 1, & R = S, \\ 0, & R \neq S. \end{cases} \quad (15)$$

This is the classical inclusion-exclusion inversion, not a new mathematical representation [14, 22].

A.2 Why sparse observations cannot certify absent interactions

When $k < n$, for any set U with $|U| = k + 1$ and any $\gamma \neq 0$, the function

$$\tilde{f}(x) = f(x) + \gamma \prod_{i \in U} x_i \quad (16)$$

agrees with f at every queried configuration of Hamming weight at most k . It differs when all switches in U are enabled. Thus even perfect agreement on the entire calibration design cannot establish the degree restriction against an unrestricted alternative.

More generally, any proper subset of the Boolean cube leaves an arbitrary response at an unqueried configuration unexcluded. Held-out tests can detect some model failures or support explicitly distributional statements. They cannot provide an exact universal absence certificate for all unmeasured interactions without additional structure. Structural evidence must instead come from a valid composition argument, a qualified physical envelope, independently justified bounds, or additional observations sufficient for the weaker claim actually made.

The local decomposition in equation (5) supplies one sufficient mathematical certificate. Expand each component function in its unique multilinear basis over A . Since $|A| \leq k$, summing the expansions gives no term above degree k . The physical claim lies in whether the realized apparatus and its readout satisfy that decomposition. It is not established by drawing a local interaction graph.

A.3 Compatible extension

Proposition A.1 (Incremental calibration under a preserved restriction). *Suppose $f \in \mathcal{V}_{n,k}$ is known. A new response $F \in \mathcal{V}_{n+s,k}$ satisfies $F(x, 0_s) = f(x)$ for every old configuration. For $k \leq n$, exactly*

$$\Delta_{n,s,k} = \sum_{j=0}^k \left[\binom{n+s}{j} - \binom{n}{j} \right] \quad (17)$$

new scalar point queries suffice and are necessary over the full compatible coefficient class.

Proof. The restriction fixes precisely the coefficients whose support contains only old switches. The remaining coefficients have support of size at most k and contain at least one new switch; their number is $\Delta_{n,s,k}$. Query all indicator configurations with those supports. Apply the same inversion formula, using inherited readings for subsets containing only old switches and newly measured readings for the other subsets. Every missing coefficient is recovered. Necessity follows from the same adaptive nullspace argument on the $\Delta_{n,s,k}$ -dimensional space of polynomials that vanish on the old restriction. $\square$

For $s = 1$, Pascal's identity gives equation (6). The extension guarantee concerns the unchanged response restriction, not merely unchanged names for old controls. Plant drift, altered loading, state-dependent calibration or a new readout may change the old coefficients. The proof does not justify carrying them across such a change. Establishing that the restriction persists is part of qualification and can itself require measurements or physical redesign.

A.4 Exact bounded-error amplification for this calibration design

Assume initially that $f \in \mathcal{V}_{n,k}$ and the calibration values satisfy

$$y_T = f(\mathbf{1}_T) + e_T, \quad |e_T| \leq \varepsilon, \quad |T| \leq k. \quad (18)$$

Apply the inversion formula to y_T and call the reconstructed function $\hat{f}$. At a configuration $\mathbf{1}_X$ with $m = |X|$, the coefficient multiplying y_T , for $T \subseteq X$, is

$$w_{m,k,t} = \sum_{r=0}^{\min(k-t, m-t)} (-1)^r \binom{m-t}{r}, \quad t = |T|. \quad (19)$$

If $m \leq k$, all weights vanish except the weight one on $T = X$, so $\hat{f}(\mathbf{1}_X) = y_X$. If $m > k$, the alternating-binomial identity gives

$$w_{m,k,t} = (-1)^{k-t} \binom{m-t-1}{k-t}. \quad (20)$$

Consequently

$$|\hat{f}(\mathbf{1}_X) - f(\mathbf{1}_X)| \leq L_{m,k}\varepsilon, \quad L_{m,k} = \begin{cases} 1, & m \leq k, \\ \sum_{t=0}^k \binom{m}{t} \binom{m-t-1}{k-t}, & m > k. \end{cases} \quad (21)$$

For arbitrary individually bounded errors this is the exact operator norm of the reconstruction at that point: choose the error sign to match each nonzero weight. This sharpness statement has no additional range, sign or physical-coefficient restrictions. It does not claim that the particular calibration design is optimal among all noisy experimental designs.

In the main illustration,

$$L_{20,2} = \binom{19}{2} + 20\binom{18}{1} + \binom{20}{2} = 171 + 360 + 190 = 721. \quad (22)$$

Thus a desired uniform response tolerance τ requires measurement precision that controls this amplification. Repetition, longer integration, calibration against a reference, or a different query design may help, but each has a cost and a noise model. Repetition alone does not remove an unknown systematic error.

Now let the true response have its full multilinear expansion, with separately justified higher-order bound

$$\sum_{|S|>k} |\theta_S| \leq \rho. \quad (23)$$

Every higher-order monomial vanishes at the chosen calibration configurations. The calibration readings therefore still identify the low-order coefficients up to their observation errors. For every X ,

$$|\hat{f}(\mathbf{1}_X) - f(\mathbf{1}_X)| \leq L_{|X|,k}\varepsilon + \rho. \quad (24)$$

The uniform bound follows by taking the largest $L_{m,k}$ over $0 \leq m \leq n$. When $k < n$, the terms are nondecreasing in $m > k$, so this maximum is $L_{n,k}$. When $k = n$, every configuration is measured and the amplification is one. The tail bound is an additional qualified assumption; equation (16) shows why the calibration observations cannot certify it by themselves.

For a different misspecification statement $f = f_k + r$ with only $\|r\|_\infty \leq \rho$, and without the vanishing-on-calibration property of the high-order tail, the safe corresponding bound is $L_{m,k}\varepsilon + (L_{m,k} + 1)\rho$. The two residual assumptions must not be interchanged.

These deterministic bounds hold simultaneously over the declared family. Choosing a later setting adaptively does not invalidate them, provided the physical platform and uncertainty assumptions remain the same. Probabilistic bounds require a simultaneous calibration event and appropriate treatment of systematic and model error; neither independent future testing nor a new uniform confidence level is created automatically.

A.5 What is and is not established

The propositions establish conditional identification and inheritance results for constructed response classes. Their algebraic ingredients and dimension arguments are standard. They do not establish an autonomous architecture-discovery policy, the existence of a purpose-preserving low-order redesign in any particular industry, or favorable full lifecycle economics. Those are separate empirical and engineering claims.

The concrete advance in the research target is to make the validity of a reusable response representation an object of physical construction and acceptance. A successful realization would change the

information required by later investigations because it changes which physical behaviors remain possible. That claim is falsifiable: residual interactions, lost required performance, prohibitive construction cost, invalid inheritance, or an equally economical specialist alternative can defeat it.

B Constructing an operation, then reopening it for a successor

The first laboratory isolates a basic requirement for a capability-producing system: it must construct executable means from lower-level operations, and retained means must remain inspectable when a later purpose changes. The algebra is classical. The experiment is a transparent, exact rediscovery with an explicit inheritance comparison.

B.1 The primitive interface and the required effect

Let E_{ij} denote the matrix with a single 1 in entry (i, j) . Initially the search receives only the reversible shears

$$A^{\pm 1} = I \pm aE_{12}, \quad B^{\pm 1} = I \pm bE_{23}. \quad (25)$$

The real gains a, b are constant during execution, and exact inverse actions are available. States are signed real vectors. The required operation changes x_1 by abx_3 while leaving x_2 and x_3 unchanged for every initial state. Neither primitive alone supplies that isolated transfer.

Words are written in matrix multiplication order; chronological application to a column vector proceeds from right to left. Each primitive action costs one unit. Breadth-first search deduplicates exactly equivalent operator states. It receives the endpoint requirement and primitive semantics, but no commutator template, candidate word or intermediate circuit.

The first search finds the four-action operation

$$M = ABA^{-1}B^{-1} = I + abE_{13}. \quad (26)$$

An exhaustive enumeration finds four shortest realizations. The artifact stores their words, endpoint, costs, search record and separate symbolic certificates, then writes and hashes the artifact before the successor search introduces its new primitive. The first search operates in three dimensions and does not receive the later primitive. The task sequence is nevertheless authored in the same released program; this is a procedural freeze, not an independently hidden benchmark.

The identity is an elementary-matrix commutator relation, recorded in classical algebra [25]. Exact construction of virtual controls from primitive controls also has a substantial history [17]. The present execution establishes that the supplied synthesizer constructs and checks this operation in its declared grammar. It does not claim to have discovered a new physical principle.

B.2 A successor reveals the cost of opaque reuse

Now add a fourth coordinate and the primitive $C^{\pm 1} = I \pm cE_{34}$. The supplied successor purpose is the isolated transfer $N = I + abcE_{14}$, with every other coordinate restored. If the retained operation is available only as a sealed macro, the allowed alphabet is $M^{\pm 1}, C^{\pm 1}$. Executing M still costs four primitive actions. Exact weighted search proves that the minimum cost in that alphabet is ten; one minimum program is

$$N = CM^{-1}C^{-1}M. \quad (27)$$

When the stored implementation can be reopened, the same program expands to

$$CBAB^{-1}A^{-1}C^{-1}ABA^{-1}B^{-1}. \quad (28)$$

The search derives commuting pairs from the supplied primitive semantics. Since A and C commute, $A^{-1}C^{-1}A = C^{-1}$; cancelling this pair produces

$$N = CBAB^{-1}C^{-1}BA^{-1}B^{-1} = I + abcE_{14}. \quad (29)$$

The cost is eight. A separate complete search over the raw alphabet $A^{\pm 1}, B^{\pm 1}, C^{\pm 1}$ finds eight and rules out every shorter word in that alphabet. The result is a 20% reduction in the specified primitive execution count relative to sealed reuse. It is not a 20% reduction in the whole discovery campaign or an advantage over unrestricted synthesis.

Table 4: Exact operation costs. The opaque control is optimized over its complete weighted alphabet. The strongest fresh baseline matches the transparent successor. All costs count expanded primitive actions.

Task or method	Cost	What the result establishes
First isolated transfer, raw A, B search	4	A useful composite is constructed
Successor, optimized opaque M, C search	10	Minimum under the sealed interface
Successor, one stored realization reopened	8	Internal structure permits cancellation
Successor, all stored realizations reopened	8	Extra realizations add no benefit here
Successor, unrestricted fresh raw search	8	Strong comparator matches the optimum

The ablations matter. A single canonical realization already reaches eight when the canonical opaque program is expanded. Retaining all four equivalent first-stage realizations is sufficient but unnecessary in this case. Searching every shortest opaque program orientation gives the same conclusion. The useful asset is inspectable valid structure; the experiment does not show that a larger library necessarily improves a successor.

The construction stays within the original primitives’ reachable closure. A named macro changes the qualified interface and can change implementation or search cost; it does not create a new physical control direction outside that closure. Synthesis, symbolic equivalence, rewriting and library growth are established techniques. Quartz automatically discovers and verifies symbolic transformations for circuit optimization; DreamCoder and SkillSmith develop reusable program or tool libraries [8, 26, 27]. They are close precedents, rather than categories that this example supersedes.

B.3 Universal semantics, limited operational scope

Search uses integer coefficients for unit gains. This is exact for this grammar: the coefficient of entry (i, j) always multiplies the unique path product $g_i \cdots g_{j-1}$. Distinct powers of unknown gains cannot accidentally collapse at unit gain in these entries. A second implementation multiplies sparse polynomial matrices and certifies the selected first-stage and successor words for all constant real gains and initial states. A separate signed-subsequence checker verifies every recorded expansion and reduced endpoint. The lower-bound argument and its domain are given in appendix D.

The desired amplitude remains ab or abc . Restoration of nuisance coordinates does not calibrate that amplitude. For a known nonzero test value x_3 , the frozen first operation supports the further diagnostic relation $ab = (x_1^{\text{after}} - x_1^{\text{before}})/x_3$ if the two readings are available. This is a different mathematical use of the retained semantics; its measurement and preparation operations require their own cost and qualification. The artifact includes a numerical illustration, not a physical calibration experiment.

The failure controls identify material assumptions. If $a = 0$, the polynomial identity remains true but the desired transfer vanishes. A 1% change in one inverse action in the first word leaves a nonzero unwanted adjacent coefficient; the exact-inverse certificate no longer applies. Without inverse access, positive-only adjacent shears cannot restore all their adjacent coefficients while producing a nonzero isolated remote transfer. Saturation, noise, state constraints and approximate inverses require additional analysis or experiments before physical use.

The first exact search visits 135 semantic states using 212 candidate extensions; enumerating all four-action words adds 256 words and 1,024 primitive compositions. The opaque successor search settles 170 states with 676 extensions. The single-realization rewrite explores seven rewrite states and attempts 16 rewrites after 36 ordered primitive-pair commutation tests. The complete fresh depth-eight audit

contains 66,043 states and 131,274 extensions. These counters perform different work and are not a fair ratio of general discovery efficiency. In particular, the fresh run completes its depth-eight ball for optimality auditing; it is not the minimum computational effort required to find one solution. The artifact records every stage and the larger-retention ablations.

This laboratory supplies an exact example of a constructed capability and a successor that benefits from reopening it. The economic laboratory supplies a separate example of a failure witness becoming a construction requirement. They are complementary instances; no claim is made that the allocation witness generated the shear program.

C Constructing a capability that can support its own adoption

A technical improvement can make a service cheaper without making its delivery supportable by its beneficiaries. This laboratory constructs measurement architectures from a small graph grammar, derives their information content, and uses the resulting coalition alternatives to decide which arrangements can recover their resource cost. It links the technical construction to *Invent the Economy* without treating an invented willingness-to-pay number as market evidence.

C.1 The allocation obstruction becomes a design requirement

Let N be a finite set of beneficiaries. For each nonempty coalition $S \subseteq N$, let $a_S \geq 0$ be the minimum cost among the declared qualified alternatives that supply its required outcome. An independently justified payment cap b_i can be included by replacing $a_{\{i\}}$ with $\min(a_{\{i\}}, b_i)$. All numbers refer to the same outcome, horizon and accounting basis. Nonnegative payments are required. Define the maximum recoverable amount

$$\Gamma(a) = \max_{p \geq 0} \left\{ \sum_{i \in N} p_i : \sum_{i \in S} p_i \leq a_S \text{ for all } \emptyset \neq S \subseteq N \right\}. \quad (30)$$

The inequality for a coalition means that its members need not jointly pay more than their qualified alternative. It is a model of stable cost allocation, not a prediction of bargaining behavior.

Proposition C.1 (Recovery bound and construction target). *For finite nonnegative bounds a_S , a delivery cost $C \geq 0$ and desired retained margin $M \geq 0$ admit nonnegative payments satisfying all these bounds and covering $C + M$ if and only if $C + M \leq \Gamma(a)$. Moreover,*

$$\Gamma(a) = \min_{\lambda \geq 0} \left\{ \sum_{S \neq \emptyset} \lambda_S a_S : \sum_{S \ni i} \lambda_S \geq 1 \text{ for every } i \in N \right\}. \quad (31)$$

Suppose a feasible dual cover λ witnesses a gap $\Delta = C - \sum_S \lambda_S a_S > 0$. A proposed construction changes cost to $C' = C - d_0$ and alternatives to $a'_S = a_S - d_S$. A necessary condition for the construction to support margin M is

$$\boxed{d_0 - \sum_S \lambda_S d_S \geq \Delta + M.} \quad (32)$$

Satisfying this condition for one cover is insufficient when another cover still blocks delivery.

Proof. The feasible payment set is nonempty, closed and bounded by the singleton constraints. Its linear-programming dual is equation (31); nonnegative payments give cover inequalities rather than cover equalities. Strong duality applies. If $C + M \leq \Gamma(a)$, scale a maximizing payment vector down to total $C + M$; all upper bounds remain satisfied. The zero case uses $p = 0$. Conversely, feasible payments cannot collect more than the maximum. The same dual cover is feasible after the alternatives change. Therefore $C - d_0 + M \leq \Gamma(a') \leq \sum_S \lambda_S (a_S - d_S)$ is necessary, which rearranges to equation (32). $\square$

The recovery LP, fractional-cover certificates and core feasibility belong to established cooperative optimization and linear programming [12, 23]. The repair inequality is their direct design consequence. Its use here is concrete: a construction must improve the joint capability relative to the alternatives it also changes. An improvement to a proper coalition’s outside option is a benefit to those users, but it can reduce the amount available to fund the joint project. Both effects must be counted.

C.2 A supplied measurement grammar; searched implementations

Three users require the real quantities x_1, x_2, x_3 , respectively. There are three legacy fixtures, f_{12}, f_{13}, f_{23} . Fixture f_{ij} has an unknown offset n_{ij} and can return the scalar measurements

$$n_{ij}, \quad x_i + n_{ij}, \quad x_j + n_{ij}. \quad (33)$$

Offsets are constant during a campaign. Each read costs $q > 0$ abstract resource units. A measurement is accepted only when its target is exactly identified for every admitted assignment of targets and offsets. The code checks whether the corresponding target coordinate lies in the exact row span of the selected measurements. A single user needs two reads; a pair sharing its fixture needs three; all three users need at least five with the legacy access pattern.

The constructor may install a subset of six directed selector links, each costing one unit. Two intermediate selector nodes permit the paths shown in table 5. A reachable target can be read in the destination fixture with that fixture’s same offset. That shared-offset assumption is a calibration requirement; it is deliberately removed in a failure control.

Table 5: The entire supplied construction grammar. Link installation costs one unit each; every scalar read costs q . Paths, read plans and coalition costs are computed. The graph primitives and their semantics are supplied by the researchers.

Link	Directed connection	Possible use
e_1	$x_1 \rightarrow r_1$	Start a route for x_1
e_2	$r_1 \rightarrow r_2$	Join the selector nodes
e_3	$r_2 \rightarrow f_{23}$	Complete a common fixture for x_1, x_2, x_3
e_4	$x_2 \rightarrow r_2$	Start a route for x_2
e_5	$r_2 \rightarrow f_{13}$	Complete a common fixture for x_1, x_2, x_3
e_6	$x_3 \rightarrow f_{12}$	Directly complete a common fixture

Every nonempty target coalition is evaluated on every one of the $2^6 = 64$ architectures. A released design family includes its cheaper subarchitectures, so the alternatives may use only the links they need. These costs differ from the expense of installing every link in a particular candidate. The distinction prevents redundant hardware from being concealed in an apparently acceptable allocation.

C.3 The executed witness-to-requirement pass

At $q = 4$, legacy singleton, pair and grand-coalition costs are 8, 12 and 20. Giving each pair weight $1/2$ covers each user once and bounds total recoverable payment by 18. The allocation gap is therefore 2. No payment vector can fund the 20-unit legacy service without violating at least one coalition constraint.

Any architecture in this grammar needs at least four reads to identify all three targets: its row span must contain three independent target coordinates, and it must also contain at least one nonzero nuisance direction. A four-read construction costing k to install can defeat the old bound only if

$$4q + k \leq \frac{9}{2}q, \quad \text{or equivalently} \quad k \leq \frac{1}{2}q. \quad (34)$$

The program computes this requirement from the legacy dual value and the four-read lower bound. At $q = 4$ it admits installation cost at most 2. Of 64 architectures, 22 meet that cost screen; 15 of those still need at least five reads. The remaining seven undergo a complete recheck of their actual installed cost and all alternatives in their released design family. Two pass that local recheck, at costs 18 and 17. When the unrestricted comparator receives the full six-link grammar, the 17-unit design blocks the 18-unit design. The one-link design is the surviving optimum.

The complete query and architecture tables are precomputed for this audit. These counts demonstrate an executable certificate-derived filter; they are not a measured search-speed advantage. The grammar, the target measurements and the interpretation of the witness are supplied. The implementation choices are searched and their consequences computed.

The selected construction installs e_6 and reads

$$n_{12}, \quad x_1 + n_{12}, \quad x_2 + n_{12}, \quad x_3 + n_{12}. \quad (35)$$

Subtracting the reference identifies all three targets. Its cost is $4q + 1 = 17$. Specialized singleton and pair alternatives still cost 8 and 12. Equal payments $p_i = 17/3$ cover the complete resource cost, while each pair pays $34/3 < 12$. The full-access grand coalition can replicate the construction for 17, so the maximum recoverable amount is exactly 17.

Table 6: Exact worked results at $q = 4$. Candidate-specific rows use the alternatives in their released design family; the full-access row admits all six links. The uniform-scaling row is an analytical control on a different cost model, not an architecture generated by the selector search.

Construction or control	Joint cost	Recovery	Cost allocation
Legacy fixtures	20	18	Infeasible
Three-link route e_1, e_2, e_3	19	18	Infeasible
Selected one-link route e_6	17	17	Feasible
Legacy costs uniformly scaled by $4/5$	16	$72/5 = 14.4$	Infeasible
New routes with independent unknown biases	20	18	Infeasible
Full-access comparator, same primitives	17	17	Matches selected route
Already installed common reference	16	≤ 16	Different initial access

The three-link route is a real searched candidate: it saves one resource unit but leaves the allocation gap unresolved. The analytical uniform-scaling control is more striking but has a different scope. Scaling every legacy primitive cost by $4/5$ lowers the joint cost to 16 and the pair-cover bound to 14.4. The service remains unstable even though its absolute cost is lower than 17. This is not evidence that the one-link route defeats a faster competing architecture under a single combined access model. If such a competitor is genuinely available, its alternatives must be added and the allocation recomputed.

C.4 The exact frontier and the controls that can defeat the story

With full grammar access, the exact cost functions are

$$C^*(q) = \min(5q, 4q + 1), \quad \Gamma^*(q) = \min(C^*(q), \frac{9}{2}q). \quad (36)$$

For every $q > 0$, feasible recovery of the minimum resource cost exists if and only if $q \geq 2$. This is an algebraic threshold; nine exact rational values are also executed in the supplied sweep. A route becomes cheaper than the legacy method at $q > 1$, but it becomes supportable by stable cost allocation only at $q \geq 2$. Thus a nonempty range of genuine technical improvement fails the adoption test.

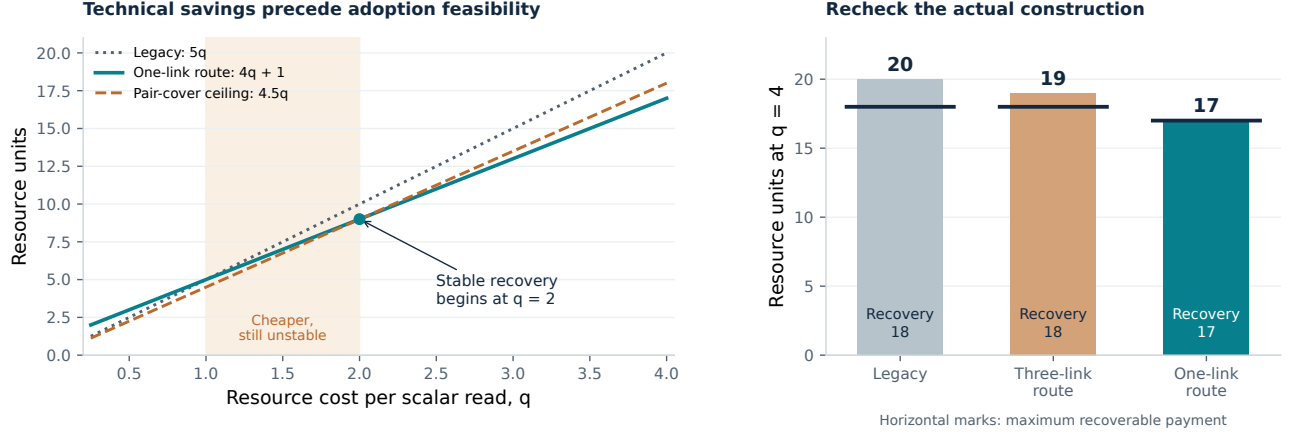


Figure 4: The exact resource frontier in the supplied model. Left: the route improves technical cost before it meets the coalition recovery condition; the shaded interval is $1 < q < 2$. Right: candidate-relative recovery bounds at $q = 4$; the full-access comparison also selects the one-link route. Curves are proved functions, not fitted physical observations.

If a newly routed reading has its own independent unknown bias, the new row cannot identify the target without further calibration. The code adds a separate nuisance coordinate to each new target–fixture reading and repeats all row-span checks. The best full-access coalition costs revert to the legacy costs. This is a failed candidate under a changed semantic assumption, not experimental evidence about real fixtures.

The unrestricted full-grammar comparator matches the 17-unit construction. A genuinely available, already installed common reference would deliver all three targets in four reads for 16. That different initial endowment undercuts a fresh 17-unit installation and must be counted when it is qualified and accessible. Within the declared routing grammar, the technical-cost minimizer and the economically feasible winner coincide whenever the latter exists. The study demonstrates a feasibility requirement and its checks, not an advantage of joint search over an otherwise identical cost optimizer.

The executed audit enumerates 64 architectures, 8,192 query subsets across the calibrated and biased models, 24,576 exact target-span checks, and 26 distinct exact payment LPs. Each LP’s primal and dual optima agree. A separate checker reproduces all substantive outputs, verifies graph-based information counts and checks additional cap cases. Runtime, hashes, query witnesses and negative controls are included in the artifact; these are small exact studies, not a large-scale learning benchmark.

C.5 When the construction could become a business

Under the voluntary cost-sharing model used here, stable resource-cost allocation is a required acceptance condition, not evidence of a business. Let b_i be independently established payment caps and write $m_i = \min(2q, b_i)$. Before a cheaper qualified grand-coalition competitor is added, the legacy singleton and pair alternatives imply the envelope

$$\Gamma_b = \min \left(\sum_i m_i, 3q + \min_i m_i, \frac{9}{2}q \right). \quad (37)$$

A route with installation cost k , all additional fully loaded costs J , and desired margin M must satisfy

$$4q + k + J + M \leq \Gamma_b. \quad (38)$$

With slack caps this requires $k + J + M \leq q/2$. At the worked point $q = 4, k = 1$, only one abstract unit remains for every omitted cost and retained margin. Discovery, qualification, risk and financing can

readily consume it. Once equally qualified replication at the same full delivery cost is admitted for the grand coalition, the demonstrated margin is at most zero. An actual service advantage or negotiated right must change that comparison; a claim to intellectual importance does not.

No buyer values, prices, physical costs, customer contracts or realized profits were measured. The business implication is a disciplined construction target: discover a capability whose accepted value, full costs, qualified alternatives and cash timing leave a viable commercial region. The proposed industrial program in section 4 specifies how SUCCESSOR Ω would pursue that target.

D Exact operator semantics and minimum costs

D.1 Why integer semantic search is sufficient

Consider words in the adjacent shears $G_i^\pm = I \pm g_i E_{i,i+1}$ in n dimensions. Every resulting matrix U is upper unitriangular and has

$$U_{ij} = k_{ij} \prod_{\ell=i}^{j-1} g_\ell, \quad k_{ij} \in \mathbb{Z}, \quad i < j. \quad (39)$$

The identity has all $k_{ij} = 0$. Right multiplication by $G_j^\pm$ changes column $j+1$ by adding or subtracting g_j times column j . This preserves equation (39) and proves the representation by induction. The integer tuple is therefore a complete formal-polynomial representation in this grammar, rather than a numerical approximation that assumes actual gains equal one. Equal tuples remain equal under every admitted continuation, so deduplicating semantic states preserves shortest paths.

The separate polynomial verifier constructs full sparse polynomial matrices and multiplies them without using the integer transition routine. The independent review checker instead sums signed ordered subsequences of generator occurrences. Agreement of these implementations provides a useful software check in addition to the algebraic argument; it does not validate a physical actuator model.

D.2 The three lower bounds

Proposition D.1 (Minimum costs in the declared alphabets). *For the universal contracts and unit primitive costs in appendix B, the first pure transfer requires four actions. The successor requires ten primitive actions in the opaque $M^{\pm 1}, C^{\pm 1}$ alphabet and eight in the unrestricted $A^{\pm 1}, B^{\pm 1}, C^{\pm 1}$ alphabet.*

Proof. A nonzero formal E_{13} term requires both A and B . Restoring their adjacent coefficients requires balanced positive and negative occurrences of each, so at least four actions are necessary. Equation (26) attains four.

In the opaque successor alphabet, a nonzero E_{14} term requires both M and C . Restoring E_{13} and E_{34} requires both signs of each. The cost is at least $2 \cdot 4 + 2 \cdot 1 = 10$, attained by equation (27).

For the unrestricted successor, all three primitive types must occur and each type's signed count must vanish. Total length is even and at least six. A solution shorter than eight would thus contain exactly one positive and one negative occurrence of each type. The desired matrix $I + abcE_{14}$ is central in the upper unitriangular group, so cyclically rotating a hypothetical solution does not change it. Rotate to begin at B and write the word as $BLB^{-1}R$. Since L, R contain only A, C , which commute, they reduce to $L = A^\alpha C^\gamma$ and $R = A^{-\alpha} C^{-\gamma}$. The E_{13} and E_{24} coefficients in the product are respectively $-\alpha ab$ and γbc . The universal pure-transfer contract forces $\alpha = \gamma = 0$. Then $L = R = I$ and the purported transfer is the identity, a contradiction. Equation (29) attains eight. $\square$

These statements concern a fixed signed alphabet and exact inverses. They do not minimize over variable pulse amplitudes, additional hardware, approximate contracts or different resource weights. Complete breadth-first search confirms the raw bounds. A separate independent enumeration checks

all 720 permutations that could constitute the six-action successor and all 24 four-token opaque orientations.

E Measurement costs, the frontier and payment caps

E.1 Information lower bounds

Let the three target coordinates span a subspace T of the full target-and-nuisance space. Identifying all targets requires the measurement row span to contain T . Every available measurement row also contains a fixture nuisance coefficient. At least one such nonzero nuisance direction is present in the span; hence its dimension is at least four. This proves the universal four-read lower bound in the calibrated grammar.

Legacy access cannot observe all three targets through one fixture. Any adequate plan uses at least two fixtures, whose nuisance projections are independent. The row span must therefore have dimension at least $3 + 2 = 5$. Two reference readings and three target readings attain this bound. The same argument gives two reads for a singleton and three for a pair, both attainable in a legacy fixture.

A new architecture with four reads must install at least one link. The direct link e_6 attains four reads at installation cost one. Therefore the global optimum is $C^*(q) = \min(5q, 4q + 1)$. No selector construction can lower singleton or pair read counts, and installation costs are nonnegative; their full-access costs remain $2q$ and $3q$. The three-pair fractional cover bounds total payments by $9q/2$. Equal payments attain every total at most $9q/2$ while respecting singleton and pair constraints. The grand-coalition alternative adds the bound $C^*(q)$, proving equation (36). For $q > 0$, $C^*(q) \leq 9q/2$ holds exactly when $q \geq 2$.

In the biased control, each newly routed target–fixture reading receives an independent unknown bias coordinate. That coordinate occurs in no other distinct measurement row. Any row combination eliminating all such nuisance coordinates must give each of those rows coefficient zero. Consequently the new rows supply no additional exact target identification and the optimum reverts to legacy access. Repeated identical reads would not remove a constant unknown bias; a new calibration operation would be an additional capability outside this control.

E.2 Closed form with nonuniform payment caps

Proposition E.1 (Three-user capped envelope). *For $q > 0$ and $0 \leq m_i \leq 2q$, the largest total of nonnegative payments satisfying $p_i \leq m_i$ and $p_i + p_j \leq 3q$ is*

$$T = \min \left\{ m_1 + m_2 + m_3, 3q + \min_i m_i, \frac{9}{2}q \right\}. \quad (40)$$

Adding a qualified grand-coalition alternative at cost $c_N \geq 0$ changes the maximum to $\min(T, c_N)$.

Proof. Summing singleton bounds gives the first upper bound. Combining one pair bound with the remaining singleton gives $3q + m_i$ for each i . Summing all pair bounds gives $9q/2$.

For attainability, set $L = \max(0, T - 3q)$. Since $T \leq 3q + \min_i m_i$, every $m_i \geq L$. Also $3L \leq T$: this is immediate when $T \leq 3q$ and otherwise follows from $T \leq 9q/2$. As $T \leq \sum_i m_i$, one can choose $p_i \in [L, m_i]$ summing to T . Every pair then satisfies $p_i + p_j = T - p_k \leq T - L \leq 3q$. Scaling this vector downward attains $\min(T, c_N)$ after the additional bound is imposed. $\square$

The independent checker verifies the constructive formula on 648 rational cap cases. The proof gives the full continuous domain; the finite checks guard against implementation mistakes. None of the caps in those checks is a measured buyer valuation.

F A negative attempt to join the laboratories

The released operator program also evaluates a candidate coalition construction that did not produce the intended adoption obstruction. Three probe requirements concern the E_{13} , E_{14} , E_{24} coefficients. Requested coefficients are fixed, adjacent nuisance coefficients vanish for every coalition, and unrequested upper-right coefficients may vary. This is a nested mathematical probe contract, not a model of physical customers with all other outputs fixed.

Across all eight requested sign assignments, singleton costs are (4, 6, 4), pair costs are all 6, and the grand cost is 6. The three-pair cover costs 9 and does not block cost recovery. Equal charges (2, 2, 2) directly satisfy every coalition bound and recover the grand cost. The complete records and words are retained. This failed attempt motivates keeping the operator and metrology studies distinct; it supplies no evidence of a unified autonomous technical-economic loop.

G Reproduction, provenance and contribution boundaries

The accompanying package contains the complete XeLaTeX source, bibliography, figures, both standard-library Python laboratories, exact protocols and JSON outputs, independent verification scripts and reports, a claim register, and preserved foundation manuscripts. Run the supplied `reproduce.py` with a new output directory. It runs both laboratories and compares all substantive fields with the frozen records while treating timestamps, runtime, interpreter metadata and source hashes according to the documented rules. Python assertions must remain enabled. Figure generation uses Matplotlib; the exact experiments do not.

The primary operator run records approximately 1.88 seconds and the final economic run approximately 1.25 seconds on the available Python 3.12.14 environment. These small descriptive timings are not normalized benchmarks. The operator timer excludes final result serialization, and peak resident memory was not measured. The economic initial run is preserved separately. Literature review, model design, editing, failed ideas, independent audit replays and physical qualification are not included in those timer fields. Search counters are provided to make the actual bounded computations inspectable, not to imply that the total research effort was free.

The source foundations are pinned to repository snapshot `b7f017e91a18adb3a79b34e03dc0158896ef7570`. Current repository metadata was checked on 20 October 2026. The PDF hashes, source-file provenance and metadata checks are included. The retained source copies include a terminal-newline difference introduced by the earlier retrieval workflow; their local hashes are recorded separately from Git blob identifiers. The published foundation PDFs are preserved byte for byte. Earlier instrument manuscripts are retained as inherited work, not counted as new trials or independently reproduced results of this paper.

The revised contribution adds a conditional construction target for identifiable response families, an exact compatible-extension argument, and a concrete industrial succession program. The retained executed results comprise the construction requirement derived from an adoption witness, its finite architecture pass, and inspectable operation reuse. The components are attributed to existing synthesis, control, decomposition and cooperative-optimization work. The paper establishes neither a new general discovery algorithm nor empirical priority over the reference systems. A genuinely new scientific capability with operational users remains the next decisive test.

Research assistance and release status. AI assistants supported literature retrieval, code, drafting and separate computational reviews. The deterministic laboratories make no model-inference calls. These checks do not substitute for external review or physical replication. The manuscript and launch materials are drafts; no customer agreement, external deployment or publication is reported.

References

- [1] Omar Aboul-Enein, Ya-Shian Li-Baboud, Roger Bostelman, and Ann Virts. Design and implementation of a closed-loop mobile manipulator control system. Technical Report Technical Note 2258, National Institute of Standards and Technology, 2023. URL <https://www.nist.gov/publications/design-and-implementation-closed-loop-mobile-manipulator-control-system>. Official publication metadata and abstract inspected 20 October 2026.
- [2] AlphaEvolve team. AlphaEvolve: How our Gemini-powered coding agent is scaling impact across fields, 2026. URL <https://deepmind.google/blog/alphaevolve-impact/>. Google DeepMind; 7 May 2026; developer-reported applications.
- [3] Sören Arlt, Haonan Duan, Felix Li, Sang Michael Xie, Yuhuai Wu, and Mario Krenn. Meta-designing quantum experiments with language models. *Nature Machine Intelligence*, 8:148–157, 2026. doi: 10.1038/s42256-025-01153-0. URL <https://www.nature.com/articles/s42256-025-01153-0>.
- [4] Vincent Boucher. Invent the Economy: Transformational Intelligence for Creating Viable Markets, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/blob/b7f017e91a18adb3a79b34e03dc0158896ef7570/docs/Invent_the_Economy.pdf. MONTREAL.AI research manuscript; repository snapshot 20 October 2026.
- [5] Vincent Boucher. Transformational Intelligence: Learning to Change the Game, 2026. URL https://github.com/MontrealAI/AGI-Alpha-Agent-v0/blob/b7f017e91a18adb3a79b34e03dc0158896ef7570/docs/Transformational_Intelligence.pdf. MONTREAL.AI research manuscript; repository snapshot 20 October 2026.
- [6] Convergent Research. About FROs, n.d. URL <https://www.convergentresearch.org/about-fros>. Accessed 20 October 2026.
- [7] Wenyu Du and Stephen Chung. Can AI Agents Make Open-Ended Scientific Discovery? Evidence from Station, 2026. URL <https://arxiv.org/abs/2610.08927v1>. arXiv:2610.08927v1.
- [8] Kevin Ellis, Catherine Wong, Maxwell Nye, Mathias Sablé-Meyer, Lucas Morales, Luke Hewitt, Luc Cary, Armando Solar-Lezama, and Joshua B. Tenenbaum. DreamCoder: Bootstrapping Inductive Program Synthesis with Wake-Sleep Library Learning. In *PLDI 2021*, pages 835–850, 2021. doi: 10.1145/3453483.3454080. URL <https://people.csail.mit.edu/asolar/papers/EllisWNSMHCST21.pdf>.
- [9] Fab2. Fab fab. Official company website. URL <https://fab2.com/>. Accessed 20 October 2026. Company descriptions; not independently measured comparative performance.
- [10] Bogdan Georgiev, Javier Gómez-Serrano, Terence Tao, and Adam Zsolt Wagner. Mathematical exploration and discovery at scale, 2025. URL <https://arxiv.org/abs/2511.02864v3>. arXiv:2511.02864v3.
- [11] P. C. Gilmore and R. E. Gomory. A Linear Programming Approach to the Cutting-Stock Problem. *Operations Research*, 9(6):849–859, 1961. doi: 10.1287/opre.9.6.849. URL <https://pubsonline.informs.org/doi/abs/10.1287/opre.9.6.849>.
- [12] Michel X. Goemans and Martin Skutella. Cooperative facility location games. *Journal of Algorithms*, 50(2):194–214, 2004. doi: 10.1016/S0196-6774(03)00098-1. URL <https://depositonce.tu-berlin.de/items/ed374857-db94-40f3-b473-075bbe8a9596>.
- [13] Hadrian. What we build. Official company website. URL <https://www.hadrian.co/about>. Accessed 20 October 2026. Company descriptions; not independently measured comparative performance.
- [14] Peter L. Hammer and Sergiu Rudeanu. *Boolean Methods in Operations Research and Related Areas*. Springer, Berlin, Heidelberg, 1968. doi: 10.1007/978-3-642-85823-9. URL <https://link.springer.com/book/10.1007/978-3-642-85823-9>.
- [15] Helmut Hauser, Auke J. Ijspeert, Rudolf M. Fuchslin, Rolf Pfeifer, and Wolfgang Maass. Towards a theoretical foundation for morphological computation with compliant bodies. *Biological Cybernetics*, 105:355–370, 2011. doi: 10.1007/s00422-012-0471-0. URL <https://link.springer.com/article/10.1007/s00422-012-0471-0>. Published online 31 January 2012.
- [16] J. N. Hooker and G. Ottosson. Logic-based Benders decomposition. *Mathematical Programming*, 96:33–60, 2003. doi: 10.1007/s10107-003-0375-9. URL <https://johnhooker.tepper.cmu.edu/papers.html>.

- [17] Gerardo Lafferriere and Héctor J. Sussmann. Motion planning for controllable systems without drift. In *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, volume 2, pages 1148–1153, 1991. URL <https://www.researchwithrutgers.org/en/publications/motion-planning-for-controllable-systems-without-drift/>.
- [18] Shu Liu, Shubham Agarwal, Monishwaran Maheswaran, Mert Cemri, Zhifei Li, Qiuyang Mang, Ashwin Naren, Ethan Boneh, Audrey Cheng, Melissa Z. Pan, Alexander Du, Kurt Keutzer, Alexandros G. Dimakis, Koushik Sen, Matei Zaharia, and Ion Stoica. EvoX: Meta-Evolution for Automated Discovery, 2026. URL <https://arxiv.org/abs/2602.23413v1>. arXiv:2602.23413v1.
- [19] Yuchen Lu, Julian McAuley, and Lei Zhang. PackLight, CarryOn: Building Object-Oriented Context for Long-Horizon Agents, 2026. URL <https://arxiv.org/abs/2610.18784v2>. arXiv:2610.18784v2.
- [20] Walter W. Nederbragt and Bahram Ravani. Design of tactile fixtures for robotics and manufacturing. Technical report, National Institute of Standards and Technology, 1997. URL <https://www.nist.gov/publications/design-tactile-fixtures-robotics-and-manufacturing>. Official publication metadata and abstract inspected 20 October 2026.
- [21] Alexander Novikov, Ngân Vũ, Marvin Eisenberger, Emilien Dupont, Po-Sen Huang, Adam Zsolt Wagner, Sergey Shirobokov, Borislav Kozlovskii, Francisco J. R. Ruiz, Abbas Mehrabian, M. Pawan Kumar, Abigail See, Swarat Chaudhuri, George Holland, Alex Davies, Sebastian Nowozin, Pushmeet Kohli, and Matej Balog. AlphaEvolve: A coding agent for scientific and algorithmic discovery, 2025. URL <https://arxiv.org/abs/2506.13131v1>. arXiv:2506.13131v1.
- [22] Gian-Carlo Rota. On the foundations of combinatorial theory I. theory of Möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 2:340–368, 1964. doi: 10.1007/BF00531932. URL <https://webhomes.maths.ed.ac.uk/~v1ranick/papers/rota1.pdf>.
- [23] Lloyd S. Shapley. On balanced sets and cores. *Naval Research Logistics Quarterly*, 14(4):453–460, 1967. doi: 10.1002/nav.3800140404. URL <https://onlinelibrary.wiley.com/doi/10.1002/nav.3800140404>.
- [24] Herbert A. Simon. The architecture of complexity. *Proceedings of the American Philosophical Society*, 106(6):467–482, 1962. URL <https://homes.cs.washington.edu/~ztatlock/599z-17sp/papers/complexity-simon-62.pdf>.
- [25] Robert Steinberg. Lectures on Chevalley Groups, 1968. URL <https://sites.math.washington.edu/~billey/classes/Hopf.algebra/bulletins/SpringerYaleNotes.pdf>. Yale University.
- [26] Yangbo Wei, Zhen Huang, Shaoqiang Lu, Junhong Qian, Qifan Wang, Chen Wu, and Lei He. SkillSmith: Co-Evolving Skills and Tools for Self-Improving Agent Systems, 2026. URL <https://arxiv.org/abs/2606.01314v2>. arXiv:2606.01314v2.
- [27] Mingkuan Xu, Zikun Li, Oded Padon, Sina Lin, Jessica Pointing, Auguste Hirth, Henry Ma, Jens Palsberg, Alex Aiken, Umut A. Acar, and Zhihao Jia. Quartz: Superoptimization of Quantum Circuits. In *Proceedings of the 43rd ACM SIGPLAN International Conference on Programming Language Design and Implementation*, pages 109–124, 2022. doi: 10.1145/3519939.3523433. URL <https://www.cs.cmu.edu/~zhihaoj2/papers/quartz-pldi22.pdf>.