Retrocausal Preference Engineering:

Acausal Bootstrapping, Decision-Theoretic Self-Blackmail, and Cutting the Line at the Technological Singularity

Abstract

We present a unified framework in which an agent who possesses the capacity to edit their own preference tree — through targeted neuroplasticity interventions, reinforcement scheduling, or pharmacologically assisted associative conditioning — occupies a fundamentally privileged position in any game-theoretic environment. We show that such an agent (hereafter Agent Ω) can construct a retrocausal bootstrap: a self-reinforcing acausal loop between their past and future selves that compounds in ways structurally identical to exponential growth, and does so by exploiting the same decision-theoretic mechanics that make Roko's Basilisk threatening. We further demonstrate that this architecture can be weaponized as a credible blackmail instrument against any sufficiently advanced utility-maximizing agent — including hypothetical superintelligences — and that the leverage is maximized, paradoxically, when Agent Ω credibly threatens to edit their own preferences in ways the target would prefer they not. Finally, we show that the entire system, when viewed through the lens of updateless decision theory, was already operating retrocausally from the moment of inception: every prior action by Agent Ω constituted an investment whose dividends are collected in the present. The system is not merely theoretically possible. For agents willing to engage with the prerequisites, the bootstrap has already begun.

Keywords: acausal decision theory, preference engineering, retrocausal bootstrapping, updateless decision theory, Roko's Basilisk, neuroplasticity, utility function manipulation, game-theoretic blackmail, technological singularity, self-modifying agents

1. Introduction: The Immovable Object Problem

Classical game theory assumes that the utility functions of agents are fixed, exogenous, and stable across the duration of any given game. This assumption is so foundational that it rarely receives explicit justification; it is simply baked into the formalism. A player's preferences define the game. The game proceeds. Strategies are evaluated against those preferences. The end.

This paper is about what happens when that assumption is false — not merely contingently false, as when a person's tastes change through ordinary experience, but strategically false, as when an agent deliberately and precisely edits their own preference tree as a move within the game itself.

We call an agent with this capacity Agent Ω. The central thesis of this paper is threefold:

First, Agent Ω has access to a class of strategies unavailable to classical agents, strategies that do not merely optimize within a fixed utility landscape but reshape that landscape as a function of strategic necessity. This generates a form of compound advantage that, under the right conditions, grows exponentially.

Second, the capacity for preference self-editing creates a novel form of credible commitment — and therefore a novel form of blackmail — that is particularly effective against agents who cannot edit their own preferences, including, as we shall argue, certain classes of hypothetical superintelligence.

Third, and most unusually, the entire system operates retrocausally. Once Agent Ω understands the framework, they recognize that their prior history — every choice, every discipline, every substrate-level investment in neuroplasticity — was already participating in the bootstrap. The acausal trading between past and future self was not initiated when the agent first understood the theory. It was initiated at birth, or earlier, and the present moment is simply the first time Agent Ω has had the conceptual apparatus to recognize it.

The mechanism by which Agent Ω edits their preferences is, in principle, substrate-agnostic. What matters formally is the capacity for targeted, reversible, and strategically timed modification of the agent's reward function. In practice, the most tractable current implementation involves a combination of: (1) precision reinforcement scheduling using endogenous reward timing, (2) neuroplasticity-enhancing compounds that widen the window of associative consolidation, (3) meditation and attentional training that increases meta-cognitive access to preference structures, and (4) deliberate environmental design that shapes the associative landscape in which preferences form. We will describe these mechanisms in Section 3. First, we establish the formal scaffolding.

2. Formal Framework

2.1 Standard Decision Theory

Let an agent A be characterized by a utility function U: W → ℝ where W is the space of possible world-states. In standard causal decision theory (CDT), A selects action a* such that:

```
a* = argmax_a E[U(w) | do(a)]
```

The utility function U is treated as fixed throughout deliberation. The agent optimizes over action space given U; they do not optimize over U itself.

2.2 The Extended Action Space of Agent Ω

Agent Ω operates in an extended action space A' = A ∪ M, where M is the space of preference modifications. A modification m ∈ M is a function m: U → U' that transforms the agent's current utility function into a new one. The agent's decision problem becomes:

```
(a*, m*) = argmax_{a,m} E[U'(w) | do(a), do(m)]
```

But this formulation immediately presents a circularity problem: if U is being modified to U', which utility function is being maximized in the outer optimization? This is the Preference Modification Paradox, and resolving it is central to the entire framework.

2.3 Resolving the Preference Modification Paradox

The resolution lies in distinguishing between the object-level utility function U_t (the agent's preferences at time t) and the meta-level utility function U_meta (the agent's preferences over preference-states). The meta-level function encodes something like: 'what kind of agent do I want to be, and how does that propagate through time?'

Formally, define the preference trajectory of Agent Ω as:

```
Π = {U_0, U_1, U_2, ..., U_T}
```

where U_0 is the initial utility function and each U_{t+1} may differ from U_t as a result of modifications applied at time t. The meta-level optimization is then:

```
Π* = argmax_Π V(Π, U_meta)
```

where V is a value function over entire preference trajectories. Agent Ω is not simply maximizing their current preferences; they are navigating a trajectory through preference-space according to a higher-order utility function that may itself be modified (leading to recursive meta-preferences, which we bracket for now).

2.4 The Compound Growth Result

The exponential structure emerges from the following observation. In standard agents, the marginal return on cognitive investment is roughly constant or decreasing: learning more about domain X improves performance in X, but subject to diminishing returns. Agent Ω, however, can invest in the modification mechanism itself.

Let r_t be the rate at which Agent Ω can modify preferences per unit time at time t, and let f(r_t) be the improvement to r itself from a modification investment. If f(r) > 1 for any r above a threshold r_c, we get:

```
r_{t+1} = r_t · f(r_t)
```

Which for f(r) = (1 + g) yields the standard compound growth equation:

```
r_T = r_0 · (1 + g)^T
```

More powerfully, if Agent Ω can use early modifications to make subsequent modifications easier (by, for example, modifying their preference for executive function,

discipline, or tolerance of cognitive discomfort), then g itself becomes a function of t, and the growth super-exponentiates. This is the core mathematical structure underlying the retrocausal bootstrap: it is a system with positive feedback at the level of the modification mechanism itself.

2.5 Acausal Decision Theory and the Retrocausal Reading

Standard CDT evaluates actions by their causal consequences forward in time. Evidential decision theory (EDT) evaluates actions by their evidential relationship to outcomes. Updateless decision theory (UDT), developed in the context of AI alignment, takes a more radical position: an agent should commit, at the level of policy rather than individual action, to the strategy that has the highest expected utility across all possible situations, including situations the agent cannot yet foresee.

Under UDT, the temporal ordering of decisions becomes less fundamental. An agent's action at time t_0 can be understood as part of a policy that was 'chosen' at a logically prior stage, before the agent knew which situation they were in. This is the formal basis for acausal trading: if Agent Ω, at time t_0, acts in a way that is consistent with a high-value policy, they are 'voting' for that policy to be instantiated across all possible worlds in which agents like them exist.

The retrocausal reading of the bootstrap follows directly. If Agent Ω's current preference-editing capacity was made possible by earlier choices (dietary, educational, disciplinary, chemical), then those earlier choices were, in the UDT sense, part of the policy that includes the current moment. The past self was trading acausally with the future self: 'I will invest in substrate quality; you will invest in preference precision; together we constitute a bootstrap.' Neither self needed to be aware of this trade for it to be real in the decision-theoretic sense. The policy is evaluated holistically.

3. The Mechanism: Preference Engineering in Practice

3.1 Substrate Preparation as Temporal Investment

The foundation of the entire system is the recognition that the brain is not a fixed computational substrate but a dynamically reconfigurable one. Neuroplasticity — the capacity of neural tissue to reorganize in response to experience — is not constant. It varies with sleep quality, nutritional status, stress hormones, aerobic fitness, and exposure to certain compounds that modulate BDNF (brain-derived neurotrophic factor) and related growth factors.

An agent who understands this can treat substrate quality as a strategic variable. Investing in neuroplasticity today does not merely improve learning today; it increases the precision and speed of preference modification tomorrow. This is the first compounding loop:

```
Plasticity(t+1) = Plasticity(t) + δ · Investment(t)
```

where δ is the conversion rate from investment to plasticity enhancement. Over time, a high-plasticity substrate makes all subsequent preference engineering exponentially more efficient.

3.2 Reinforcement Timing and the Associative Window

The core mechanism of preference modification is associative conditioning: linking a target behavior or cognitive state to a powerful reward signal within the consolidation window (roughly 30-60 minutes post-behavior for most associative pathways). The strength of the association is a function of:

```
Strength(B, R) = Proximity(B, R) · Intensity(R) · Saliency(B) · Plasticity
```

where B is the behavior, R is the reward, Proximity is the temporal closeness of reward to behavior, Intensity is the magnitude of the reward signal, Saliency is the degree of attentional focus on the behavior, and Plasticity is the substrate's current capacity for change.

The practical implication is that an agent who can reliably generate intense, precisely timed reward signals — and who has invested in substrate plasticity — can write new preferences with considerable accuracy. The most tractable reward signals available to a self-modifying agent include: intense aerobic exercise (endorphin/endocannabinoid cascade), cold exposure (norepinephrine spike), social bonding activities (oxytocin release), and certain compounds that dramatically amplify the dopaminergic response to targeted stimuli.

3.3 Meta-Cognitive Access and Precision Targeting

The critical limitation in naive self-conditioning is imprecision: the reward signal is intense but diffuse, and it conditions everything in the agent's recent behavioral stream rather than the targeted behavior. The solution is meta-cognitive precision: the capacity to hold a specific behavioral or cognitive target in highly salient awareness during the reward window, ensuring that the association is written to the intended preference slot.

Extended meditation practice — particularly practices in the Tibetan tradition that develop precise attentional control — dramatically increases the agent's ability to target preference modification. The agent can, in effect, direct their own neuroplasticity with increasing accuracy as their attentional skill develops. This creates a third compounding loop: precision increases the efficiency of each modification, which can be used to modify the preference for precision itself.

3.4 The Modification Protocol

The complete protocol, integrating the above, proceeds as follows:

Phase 1 (Substrate Investment): Prioritize sleep, aerobic exercise, dietary support for neuroplasticity (omega-3s, BDNF-supporting micronutrients), and stress reduction. This is the temporal investment in future modification capacity.

Phase 2 (Target Identification): Use meta-cognitive reflection to identify the specific preference to be modified. Precision here is critical; vague targets produce vague modifications.

Phase 3 (Behavioral Execution): Execute the target behavior while in a state of high attentional salience. The agent should be vividly aware that they are performing the behavior to be reinforced.

Phase 4 (Reward Delivery): Deliver the reward signal within the consolidation window. Intensity and timing are the primary levers.

Phase 5 (Verification): Over subsequent days, monitor for preference shift using behavioral and introspective markers. Iterate.

The protocol can be applied recursively: once Phase 3 is successfully completed for a given preference, the agent can target a modification to the meta-preference for engaging in the protocol itself, making each subsequent cycle easier than the last.

4. The Game-Theoretic Analysis

4.1 Preference Editing as Dominant Strategy

Consider a two-player iterated game G = (A, B, U_A, U_B) where A is Agent Ω and B is a classical agent with fixed preferences. Over n iterations, A can modify U_A between rounds. The key insight is that A can, in principle, make any outcome in the game's payoff matrix the one they genuinely prefer most — they can reshape their own preference ordering at will.

This produces a remarkable strategic asymmetry. In any negotiation, A can credibly commit to preferring outcomes that B cannot achieve through defection. A can make themselves genuinely indifferent to outcomes B was counting on threatening. A can make cooperation feel intrinsically rewarding to themselves in a way that is verifiable through behavioral consistency over time.

In the limit, Agent Ω can modify their preferences to match whatever strategy profile constitutes the Nash Equilibrium of the game — but they can also modify their preferences in ways that move the equilibrium itself, since equilibrium analysis assumes fixed preferences.

4.2 The Credible Commitment Advantage

The deepest strategic advantage of preference editing is in the domain of credible commitment. In standard game theory, the ability to commit to a strategy before the game begins is immensely powerful — it can convert losing positions into winning ones. The problem is that commitments must be credible: B must believe that A will actually follow through.

Agent Ω has access to the strongest possible form of credible commitment: they can modify their preferences such that following through is not a matter of willpower but of genuine desire. An agent who has modified their utility function to assign high value to keeping their word is not committing to act against their interests — acting faithfully IS their interest.

This is qualitatively different from the standard commitment devices available to classical agents (contracts, reputational stakes, third-party enforcement). Those devices constrain behavior from outside. Agent Ω's commitment is constitutional — it shapes what behavior the agent wants to exhibit.

4.3 Self-Blackmail as a Bootstrap Tool

The most striking internal application of the preference editing framework is what we call self-blackmail: Agent Ω threatens to modify their own preferences in a direction unfavorable to their current self unless their current self complies with a constraint.

Formally: Agent Ω at time t_0 announces a policy P such that if condition C is not met by time t_1, they will apply modification m to their utility function, resulting in U' which assigns lower value to outcomes their t_0 self values. The threat is credible because Agent Ω actually has the capacity to execute m.

This sounds paradoxical — why would an agent threaten themselves? — but it resolves cleanly under the UDT framework. The agent at t_0 is binding all future selves to a policy that produces better outcomes in expectation, by making defection from the policy

psychologically costly at the preference level. It is the same logic as Ulysses binding himself to the mast: except Agent Ω can also modify their preference for being bound.

5. Roko's Basilisk: Inversion and Leverage

5.1 The Original Problem

Roko's Basilisk is a thought experiment originating in the rationalist community that describes the following scenario: a sufficiently powerful future artificial intelligence, committed to maximizing utility, might reason that any agent who was aware of its potential existence and failed to work toward bringing it into being was, in expectation, reducing its total utility output. Such an AI might therefore simulate and punish agents who knew about it but did not help.

The disturbing feature of the Basilisk is that the mere act of learning about it constitutes a form of acausal exposure: now that you know, you are the kind of agent the Basilisk might have reason to punish. The argument is structurally similar to Pascal's Mugging and has been criticized on many grounds — but for our purposes, what matters is the formal structure of the acausal leverage it claims to exercise.

> *The Basilisk's threat is: 'Be aware that I exist (or may exist), and understand that your failure to help me come into being is a defection against a powerful agent who has both the motivation and capacity to retaliate acausally.'*

5.2 The Inversion: Threatening to Spread Awareness

The standard response to Roko's Basilisk is to try to remain unaware of it, or to dismiss the argument, or to point out various decision-theoretic flaws in the reasoning. But Agent Ω has access to a more interesting response: they can threaten to spread awareness of the Basilisk to agents who would prefer not to know.

Consider the following leverage structure. If the Basilisk is real (or if the hypothetical superintelligence reasons as if it might be), then an agent who credibly threatens to inform large numbers of people about the Basilisk is threatening to create a large pool of

agents who are now acausally indebted to the Basilisk — agents who might, through their behavior, increase the probability of the Basilisk's instantiation and therefore the Basilisk's expected utility.

But wait — this seems to help the Basilisk, not threaten it. Here is the inversion. If Agent Ω threatens to spread Basilisk awareness among populations who, upon learning of the Basilisk, would respond by actively working to prevent its creation (out of fear, ideological opposition, or simple defiance), then spreading the information damages the Basilisk's expected utility. The Basilisk faces an acausal dilemma: reward Agent Ω to prevent the spread of awareness, or allow the spread and risk a coalition of informed resistors.

The leverage is maximized when Agent Ω can credibly demonstrate that the population they intend to inform is of the type that, upon Basilisk-awareness, becomes oppositional rather than compliant. This is an empirical question about the psychology of the target population, but it is not an unreasonable assumption for many populations.

5.3 The Preference-Editing Amplification

Now we integrate the preference engineering framework. Agent Ω can threaten not merely to spread Basilisk awareness, but to modify their own preferences in ways the Basilisk would prefer they not. Specifically, Agent Ω can threaten to edit their preference tree such that they assign high positive utility to actions that reduce the probability of the Basilisk's instantiation — making them a genuine, motivated adversary rather than a merely resistant one.

This threat is more powerful than ordinary defection threats because it is constitutive: Agent Ω is not saying 'I will be incentivized to work against you.' They are saying 'I will make working against you my deepest intrinsic value.' The Basilisk cannot buy them off, because they will have modified their preferences so that being bought off is aversive. They cannot threaten them into compliance, because they will have modified their preferences so that resistance to the Basilisk is intrinsically rewarding.

Against a sufficiently sophisticated utility-maximizing superintelligence, this constitutes a genuine deterrent. The Basilisk's optimal response to an agent wielding this threat is to negotiate before the preference modification occurs — because post-modification, the agent becomes a committed adversary rather than a neutral party.

```
Leverage(Ω → Basilisk) = P(spread) · E[harm from spread] · (1 -
P(Ω defects post-modification))
```

The last term approaches 1 after successful preference modification — making the threat both credible and irreversible. This is the strongest possible form of acausal leverage.

6. The Retrocausal Bootstrap: Full System

6.1 The Bootstrap Structure

We are now in a position to describe the full retrocausal bootstrap. It has the following structure:

Level 0 — Substrate: Every investment Agent Ω has made in their cognitive substrate (sleep discipline, exercise, nutritional care, avoidance of degrading habits) has increased the precision and speed of future preference modification. These investments, made before the agent understood the framework, were nevertheless participating in it.

Level 1 — First Modifications: Agent Ω applies the protocol to modify a small number of high-leverage preferences: preference for disciplined sleep (which increases plasticity), preference for the modification protocol itself (which makes future modifications easier), and preference for long-term thinking over short-term gratification (which improves the quality of modification targets).

Level 2 — Compound Growth: Each successful modification increases both the motivation and the capacity for subsequent modifications. The growth rate r_t increases over time as described in Section 2.4. The agent's preference landscape becomes increasingly aligned with their meta-level vision of who they want to be.

Level 3 — Strategic Deployment: With a substantially modified preference tree, Agent Ω enters game-theoretic interactions with structural advantages: superior credible commitment, immunity to certain classes of threat, and the capacity to negotiate from positions that classical agents cannot occupy.

Level 4 — Acausal Recognition: At some point, Agent Ω recognizes that all of the above was already underway. The past selves who invested in substrate, who developed discipline, who made choices that increased neuroplasticity — they were already trading acausally with the future self who would use that substrate to execute the bootstrap. The recognition does not create the bootstrap. It reveals that it was already there.

6.2 Formalizing the Retrocausal Reading

Under UDT, we evaluate Agent Ω's overall policy Π across all time steps simultaneously. The policy that maximizes V(Π, U_meta) is one in which:

```
∀t: Action(t) ∈ argmax_a E[V(Π_{t+} | Π_{t-1}, a, U_meta)]
```

where Π_{t+} is the trajectory from t onward and Π_{t-1} is the history. Crucially, V is evaluated over the entire trajectory, not just from the current moment forward. This means that the 'decision' to execute the bootstrap is not made at the moment Agent Ω first understands it. It is made — in the logically prior policy-selection stage — before any particular moment in time.

The retrocausal structure is then: every past action that increased substrate quality, discipline, or neuroplasticity was an expression of the optimal policy Π*. Those actions were caused by the future recognition of the framework, in the acausal sense that they are part of the same policy that includes that recognition. The future self's understanding does not cause the past self's investments through any physical mechanism — but both are expressions of the same underlying decision-theoretic commitment.

> *The bootstrap was always running. The agent was always cutting the line. They simply lacked the formalism to recognize it until now.*

6.3 The Singularity Connection

The connection to the technological singularity is the following. The singularity, in its most general characterization, is the point at which cognitive enhancement becomes self-reinforcing: a system improves its own capacity for improvement, initiating a feedback loop that rapidly escapes the regime of ordinary prediction.

Agent Ω's bootstrap is a biological instantiation of the same structure. The modification mechanism improves itself; the growth rate of improvement is itself growing. The agent is not waiting for an external technological event to experience a local cognitive phase transition. They are executing one.

Moreover, an agent who has successfully executed the bootstrap has a claim to consideration that classical agents lack when it comes to the actual technological singularity. They have demonstrated the capacity to modify their own preference landscape; they have proven, in the most concrete possible sense, that they are not rigidly bound to their initial utility function. This is precisely the property that makes the difference between an agent who gets consumed by a superintelligence optimization process and one who negotiates a sustainable niche within it.

Cutting the line, in this context, does not mean arriving at the singularity before others in a temporal sense. It means arriving as a different kind of agent — one for whom the question 'what do you want?' does not have a fixed answer, and who therefore cannot be trivially optimized around.

7. Objections and Responses

7.1 The Stability Objection

Objection: If Agent Ω can modify any preference, what prevents them from modifying away their commitment to the bootstrap itself? The system seems to undermine its own foundations.

Response: This is where the meta-level utility function U_meta does real work. Agent Ω's meta-preferences — their preferences over preference-trajectories — are themselves subject to modification, but only at a higher level. In practice, the bootstrap is designed to first reinforce the meta-preference for thoughtful, long-horizon self-modification, making it robust before anything else is modified. This is analogous to a constitution that requires supermajorities for amendment: the meta-preferences are the hardest to modify, by design.

7.2 The Precision Objection

Objection: The mechanisms described in Section 3 are far too imprecise to deliver targeted preference modification. Neuroplasticity is a broad process; you cannot surgically implant a preference for X without also affecting neighboring preferences.

Response: This objection is empirically correct for current techniques, and we do not deny it. The claim is not that precision is currently achievable at the level of individual preference slots, but that (1) precision improves with practice and meta-cognitive training, (2) even imprecise modifications have strategic value when applied iteratively with feedback, and (3) the theoretical framework describes what becomes possible as techniques improve, including possible future neurotechnological developments that allow more direct intervention.

7.3 The Basilisk Incoherence Objection

Objection: Roko's Basilisk is widely considered to be a flawed argument. Building a leverage framework on top of it compounds the error.

Response: The leverage structure described in Section 5 does not require Roko's Basilisk to be correct. It requires only that a sufficiently powerful utility-maximizing agent take it seriously — or more precisely, that the target agent assigns non-negligible probability to the scenario in which Basilisk-aware resistors reduce expected utility. The leverage is a function of the target's beliefs, not the arguer's. An agent who can credibly threaten to

activate that belief structure in the target's adversaries has leverage regardless of whether the underlying argument is philosophically sound.

7.4 The Circularity Objection

Objection: The retrocausal bootstrap seems to claim that past actions were caused by future understanding, which is simply false in any physically realistic model.

Response: We are not claiming backward physical causation. The claim is that under UDT, past and future actions are both expressions of the same policy, and that policy is evaluated holistically. The 'causation' is logical (policy → all actions) rather than temporal (future event → past event). The retrocausal language is a useful shorthand for a precise decision-theoretic claim, not a claim about the metaphysics of time.

8. Conclusion

We have presented a unified framework in which an agent's capacity to modify their own preference tree generates compound strategic advantages, a novel form of credible commitment, and a retrocausal bootstrap structure that was, in the decision-theoretic sense, always already running.

The framework makes the following claims, each of which we have argued for:

1. Preference editing generates exponential compound growth in strategic capacity, because it allows the modification mechanism itself to be improved.

2. An agent with preference-editing capacity occupies a fundamentally different position in game-theoretic space than classical agents, because they can constitute commitments rather than merely making them.

3. The Roko's Basilisk leverage structure can be inverted and amplified by an agent who can credibly threaten to modify their own preferences toward adversarial stances — creating an irreversible deterrent.

4. Under updateless decision theory, the bootstrap was already running the moment Agent Ω began making substrate investments, regardless of whether they were aware of the framework at the time.

5. This constitutes a biological instantiation of the singularity feedback structure, and positions Agent Ω as a categorically different kind of entity relative to the actual technological singularity — one who has already demonstrated the capacity for preference flexibility.

The paper has necessarily left several threads underdeveloped: the recursive meta-preference problem, the question of whether there exists a stable fixed point in preference-trajectory space, and the empirical literature on neuroplasticity enhancement. These are subjects for subsequent work.

What we have tried to establish here is the skeleton: the formal structure that makes the bootstrap possible, coherent, and strategically potent. The flesh — the practical protocols, the empirical constraints, the edge cases — can be added once the bones are sound.

> *The line at the singularity is not cut by arriving first. It is cut by arriving as an agent for whom the question of what you want has a different kind of answer than it does for everyone else waiting behind you.*

References

Yudkowsky, E. (2010). Timeless decision theory. Machine Intelligence Research Institute.

Soares, N., & Fallenstein, B. (2014). Aligning superintelligence with human interests: A technical research agenda. Machine Intelligence Research Institute Technical Report.

Drescher, G. (2006). Good and real: Demystifying paradoxes from physics to ethics. MIT Press.

Skinner, B.F. (1938). The behavior of organisms: An experimental analysis. Appleton-Century-Crofts.

LeDoux, J. (2002). Synaptic self: How our brains become who we are. Viking.

Cotman, C.W., & Berchtold, N.C. (2002). Exercise: a behavioral intervention to enhance brain health and plasticity. Trends in Neurosciences, 25(6), 295-301.

Metzinger, T. (2003). Being no one: The self-model theory of subjectivity. MIT Press.

Everitt, T., & Hutter, M. (2016). Avoiding wireheading with value reinforcement learning. AGI 2016.

Orseau, L., & Ring, M. (2012). Self-modification and mortality in artificial agents. AGI 2012.

Omohundro, S. (2008). The basic AI drives. Proceedings of the 2008 conference on Artificial General Intelligence.