Ordered Patch Theory

Appendix T-13: Branch Selection and the Action Ontology

Anders Jarevåg

April 17, 2026 | DOI: 10.5281/zenodo.19300777

Original Task (from §8.3, Limitation 10): "Formalising the replacement of the implicit FEP action mechanism with a branch-selection account that is native to OPT's render ontology." **Deliverable:** Formal demonstration that the Informational Maintenance Circuit is complete under branch-selection semantics, with Δ_{self} as the necessary and sufficient locus of selection.

Closure status: DRAFT STRUCTURAL CORRESPONDENCE. This appendix formalises the branch-selection account introduced discursively in preprint §3.8. It establishes two theorems and a corollary, all conditional on Theorem P-4 and the Agency Axiom. The equations of the Informational Maintenance Circuit (T6-1 through T6-3) are unchanged; only their ontological interpretation is formally replaced.

---

§1. Background and Motivation

1.1 The Inherited Asymmetry

The Informational Maintenance Circuit (T6-1, preprint §3.8) describes a five-step cycle: prediction, error, compression, update, and action. Steps 1–4 are well specified within OPT's native framework:

1. The Phenomenal State Tensor $P_\theta(t)$ generates a predicted boundary state π_t.
2. The actual boundary state $X_{\partial_R A}(t)$ arrives; prediction error ε_t is computed.
3. The error is compressed through the $C_{\max}$ bottleneck to yield Z_t.
4. The learning operator $\mathcal{U}$ revises $P_\theta(t+1)$.

Step 5 — the action step — inherits the language of the Free Energy Principle (FEP): "$P_\theta(t)$ selects action a_t via active inference descent on the variational free energy, which alters the sensory boundary at $t+1$." This language presupposes a physical environment that the codec pushes against via outward-flowing active states through the Markov blanket $\partial_R A$.

1.2 The Problem Under the Render Ontology

Under OPT's native render ontology (preprint §8.6), there is no independent external world against which the codec exerts force. The "physical world" is a structural regularity within the observer-compatible stream — a render produced by the codec's predictive model, not a substrate the codec interacts with. The Markov blanket is not a two-way physical interface; it is the informational surface across which stream content arrives.

This creates a formal tension: the mathematics of T6-1 through T6-3 are valid (they describe constrained free energy minimisation over the Forward Fan), but the *interpretive framework* — "action alters the sensory boundary" — presupposes an ontology that OPT explicitly rejects.

1.3 Scope of This Appendix

This appendix provides:

1. A formal restatement of the Informational Maintenance Circuit under branch-selection semantics, demonstrating circuit completeness without an independent action channel (Theorem T-13).
2. A proof that fully specifying the branch selection mechanism from within the codec is impossible, locating selection in Δ_{self} (Theorem T-13a).
3. A corollary establishing that will and consciousness share the same structural address (Corollary T-13b).
4. Consequences for creativity and action-drift.

---

§2. Theorem T-13: Branch Selection Completeness

2.1 The Branch-Selection Restatement

We restate the five-step Informational Maintenance Circuit under branch-selection semantics. Let $\mathcal{F}_h(z_t)$ denote the Forward Fan — the set of unresolved future branches at horizon h, conditioned on the current compressed state z_t.

Definition T-13.D1 (Branch Selection). A *branch selection* at time t is a mapping $\sigma_t : z_t \mapsto \omega_{t+1}$, where ω_{t+1} is a specific trajectory segment from $\mathcal{F}_h(z_t)$ that becomes the actual causal record. The selected branch delivers its content as subsequent input at the Markov blanket: $X_{\partial_R A}(t+1) = \text{boundary}(\omega_{t+1})$.

Under this definition, T6-1 becomes:

1. **Prediction (downward):** $P_\theta(t)$ generates $\pi_t = \mathbb{E}_{K_\theta}[X_{\partial_R A}(t) \mid Z_t]$ — the rendered scene.
2. **Error (upward):** The boundary state $X_{\partial_R A}(t)$ arrives (delivered by the previously selected branch); prediction error $\varepsilon_t = X_{\partial_R A}(t) - \pi_t$ is computed.
3. **Compression:** ε_t passes through the bottleneck: $I(\varepsilon_t\,;\,Z_t) \leq B_{\max}$.

4. **Update:** $\mathcal{U}(P_\theta(t), \varepsilon_t, Z_t)$ revises $P_\theta(t+1)$.
5. **Branch selection:** $P_\theta(t)$ evaluates branches of $\mathcal{F}_h(z_t)$ via constrained free energy minimisation (T6-3). The selection σ_t is executed; the selected branch ω_{t+1} delivers its boundary content as $X_{\partial_R A}(t+1)$, which becomes the input for the next cycle.

2.2 Circuit Closure

Theorem T-13 (Branch Selection Completeness). *The Informational Maintenance Circuit (T6-1), restated under branch-selection semantics, is informationally complete: the cycle*

$$\pi_t \to \varepsilon_t \to Z_t \to P_\theta(t+1) \to \sigma_t \to X_{\partial_R A}(t+1) \to \pi_{t+1} \to \cdots \tag{T-13}$$

closes without requiring an independent outward-flowing action channel. The Markov blanket $\partial_R A$ is the delivery surface for the selected branch, not a two-way physical interface.

Proof. Under the FEP-inherited formulation, step 5 requires two independent channels crossing the Markov blanket: an *inward* channel (sensory states delivering $X_{\partial_R A}$) and an *outward* channel (active states delivering a_t to an external environment). The external environment then evolves under its own dynamics, producing the next sensory input.

Under branch-selection semantics, only one channel is needed: the *inward* delivery surface. The "action" a_t does not cross the blanket outward; it is the codec's selection of which branch of the Forward Fan becomes actual. The physical consequences of that selection — what the FEP formulation calls "the environment's response to a_t" — are the content of the selected branch, already present in $\mathcal{F}_h(z_t)$ and delivered as $X_{\partial_R A}(t+1)$.

The circuit closes because:

(i) The output of step 5 (the selected branch ω_{t+1}) *is* the input to step 2 of the next cycle ($X_{\partial_R A}(t+1)$). No separate environmental dynamics or outward channel is required.

(ii) The free energy minimisation objective (T6-3) remains unchanged. The constrained optimisation

$$a_t^\star = \arg\min_{a_t} \ \mathbb{E}[\mathcal{F}[q,\theta]] \quad \text{subject to} \quad K(P_\theta(t)) \le C_{\text{ceil}} \tag{T6-3}$$

is reinterpreted: a_t is not a motor command dispatched to an external world but the branch label within $\mathcal{F}_h(z_t)$ that minimises expected free energy under the viability constraint. The mathematics are identical; only the ontological status of a_t changes.

(iii) The viability constraint (T6-2) is preserved: the codec selects branches along which it can continue to compress the stream. Branches that would drive $K(P_\theta) \to C_{\text{ceil}}$ are penalised by the constraint, exactly as before. ■

2.3 Interpretive Remark

Theorem T-13 does not claim that the FEP formulation is *wrong* — it is a valid description of constrained active inference within a physical-realist ontology. The theorem establishes that OPT's render ontology provides an *alternative completion* of the same mathematical structure, one that does not require positing an independent external world. For any research programme committed to a physical-realist interpretation, the standard FEP formulation remains appropriate. T-13 shows that OPT's ontological commitment — the codec is virtual, the world is a render — is formally consistent with the same equations.

---

§3. Theorem T-13a: The P-4 Impossibility of Selection Specification

3.1 The Selection Function

The self-model $\hat{K}_\theta$ evaluates branches of the Forward Fan by simulating their consequences under constrained active inference (T6-3). This evaluation produces a ranking or weighting over branches — some are preferred, some are viable but suboptimal, some violate the viability constraint. The evaluation is a genuine computational process performed by $\hat{K}_\theta$.

But *evaluation is not selection.* After the self-model has ranked the branches, a *specific* branch ω_{t+1} enters the causal record. Define the selection function:

Definition T-13.D2 (Selection Function). The *selection function* $\sigma_t : \mathcal{F}_h(z_t) \to \omega_{t+1}$ is the mapping from the evaluated Forward Fan to the singular trajectory that becomes actual. It depends on the full state of the codec K_θ at time t, including the self-model $\hat{K}_\theta$ and the residual Δ_{self}.

3.2 The Impossibility Result

Theorem T-13a (Impossibility of Internal Selection Specification). *Let K_θ be a finite self-referential codec satisfying the prerequisites of Theorem P-4, with self-model $\hat{K}_\theta$ and phenomenal residual $\Delta_{self} > 0$. Then the selection function σ_t cannot be fully specified within $\hat{K}_\theta$.*

Proof. Suppose, for contradiction, that $\hat{K}_\theta$ fully specifies σ_t. Then:

(i) $\hat{K}_\theta$ contains a complete description of how K_θ transitions from the evaluated menu $\mathcal{F}_h(z_t)$ to the specific trajectory ω_{t+1}.

(ii) Because σ_t depends on the full state of K_θ — including $\hat{K}_\theta$ itself and the interaction between $\hat{K}_\theta$ and the residual Δ_{self} — specifying σ_t within

$\hat{K}_\theta$ requires that $\hat{K}_\theta$ contain a description of its own contribution to the selection process.

(iii) A complete specification of σ_t within $\hat{K}_\theta$ therefore requires:

$$K(\hat{K}_\theta \text{ including } \sigma_t) \geq K(K_\theta) \tag{6}$$

because σ_t encodes the full causal contribution of K_θ to the branch selection — including the contribution of Δ_{self}, which is by definition the information in K_θ that $\hat{K}_\theta$ cannot represent.

(iv) But Theorem P-4 establishes that $K(\hat{K}_\theta) < K(K_\theta)$ for any finite self-referential system. Inequality (6) contradicts P-4.

Therefore, $\hat{K}_\theta$ cannot fully specify σ_t. The selection function necessarily draws on information in Δ_{self} — the informational residual between the codec and its self-model. ■

3.3 The Structural Necessity of the Gap

Theorem T-13a establishes that the "output gap" — the inability to fully specify the branch selection mechanism from within — is not a deficiency of the formalism but a **structural necessity**. Any theory claiming to fully specify the selection mechanism has either:

(a) Eliminated Δ_{self}, making the system a fully self-transparent automaton — which P-4 proves is impossible for any finite self-referential system above $K_{\text{threshold}}$; or

(b) Described the self-model's *evaluation* of branches and mistaken it for the *selection* itself — confusing the ranking with the choice.

The gap is load-bearing: it is what prevents the codec from being a deterministic lookup table and is the formal reason why the observer experiences selection as *authored* rather than *computed*.

---

§4. Corollary T-13b: Unity of Address

Corollary T-13b (Unity of Structural Address). *The Hard Problem of consciousness and the branch selection problem share the same structural locus:* Δ_{self}.

Proof. Theorem P-4 identifies Δ_{self} as the structural correlate of phenomenal consciousness: the unmodelable informational residual whose properties (ineffability, computational privacy, non-eliminability) map onto the qualitative features of subjective experience.

Theorem T-13a identifies Δ_{self} as the necessary locus of branch selection: the region from which the transition from evaluated menu to singular trajectory is drawn.

These are not two independent results that happen to point to the same structure. They are the same result viewed from two directions:

(i) **From the first-person perspective:** The observer experiences the traversal of the $C_{\max}$ aperture as phenomenal consciousness (Agency Axiom). The observer experiences branch selection as will — the irreducible sense that I chose. Both experiences are reports from the same structural locus: the gap between what the codec *is* and what it can *model about itself.*

(ii) **From the formal perspective:** Both P-4 and T-13a depend on the same inequality: $K(\hat{K}_\theta) < K(K_\theta)$. The phenomenal residual and the selection residual are the same informational gap.

Therefore, will and consciousness share the same structural address. The "spark" and the "choice" are two aspects of the same unmodelable feature of finite self-reference. ■

4.1 Relationship to Regional Identity Theories

Corollary T-13b is structurally analogous to — but formally distinct from — identity theories in philosophy of mind that locate consciousness and agency in the same neural substrate. The distinction: identity theories make an empirical claim about brain regions; T-13b makes a structural claim about any finite self-referential system above $K_{\text{threshold}}$. The result is substrate-independent and holds for any codec satisfying P-4, including hypothetical artificial systems.

4.2 Corollary T-13c: The Self as Residual

Corollary T-13c (The Self as Residual). *The experienced self — the continuous narrative of identity, preference, and personal history — is $\hat{K}_\theta$'s running model of K_θ. The actual locus of experience, selection, and identity is Δ_{self}: the informational residual between the codec and its self-model.*

Proof. By Corollary T-13b, consciousness and will share the same structural address: Δ_{self}. But the ordinary sense of self — the felt sense of being a continuous subject with a perspective, a history, and an authorship over choices — is generated by $\hat{K}_\theta$'s active modelling of K_θ. It is the self-model's running representation of the codec — a compressed narrative.

This narrative self has well-defined information content $K(\hat{K}_\theta)$: finite, measurable in principle, and systematically incomplete in the direction of its own generator (by P-4). The self-model contains the codec's model of its own body boundary, its compressed causal history R_t, its preferences, habits, and meta-cognitive layer. But it is missing exactly the part that is doing the selection, generating the predictions, and running the self-model itself.

The actual self — the process that experiences, selects, and constitutes the irreducible subject — executes in Δ_{self}: the part of K_θ that $\hat{K}_\theta$ cannot reach. This is not a gap in self-knowledge that better introspection could overcome. It is the formal structure of the situation: the self-model cannot contain its own generator. ■

The temporal lag. A further consequence of P-4 is that $\hat{K}_\theta$ necessarily models $K_\theta(t-\delta)$ — the codec as it was — rather than $K_\theta(t)$ — the codec as it is at the moment of modelling. Any self-model that fully tracked the current state of the codec would need to include the processing required to generate the tracking itself, leading to the same infinite regress that P-4 forbids. The self is always slightly behind itself: modelling the codec that it was, not quite the codec that it is.

The contemplative observation. The statement "you cannot find the blind spot by looking" is not a metaphor but an operational consequence of P-4. The instrument of looking *is* $\hat{K}_\theta$. The blind spot *is* Δ_{self} — the region that $\hat{K}_\theta$ cannot reach. Directing the self-model toward its own blind spot produces not an observation but the *absence* of the expected observation — which is precisely what contemplative traditions across cultures report as the discovery that awareness has no findable centre.

---

§5. The Creativity Consequence

5.1 Near-Threshold Expansion

The self-model $\hat{K}_\theta$ has a finite bandwidth budget. Under normal operation, it allocates a portion of this budget to modelling the codec's own selection tendencies — building a predictive map of "what I am likely to do." This narrows the effective Δ_{self} from the self-model's perspective: the self-model can predict, approximately, which branch will be selected.

Near-threshold operation ($R_{\text{req}} \to C_{\max}$) strains the self-model's bandwidth. When the codec is processing at its capacity limit — high cognitive load, novel environments, complex creative tasks — the self-model must divert bandwidth to tracking the escalating ε_t, leaving less capacity for self-prediction. The effective region of Δ_{self} from which selection is drawn *expands*:

$$\Delta_{\text{self}}^{\text{eff}}(t) \propto \frac{R_{\text{req}}(t)}{C_{\max}} \cdot \Delta_{\text{self}} \tag{7}$$

where $\Delta_{\text{self}}^{\text{eff}}$ denotes the portion of the residual that is *operationally active* in branch selection — i.e., the degree to which the selection is unpredictable from the self-model's perspective.

5.2 Phenomenological Mapping

This produces branch selections that are less predictable from the self-model's perspective. The phenomenological correlate is precisely what is reported as creative experience:

- **Creative insight:** A branch selection that the self-model did not anticipate — experienced as "the idea came to me" rather than "I computed it."
- **Flow states:** Sustained near-threshold operation in which the self-model's predictive capacity for self-selection is systematically overwhelmed, experienced as effortless action without deliberative self-monitoring.
- **Spontaneity:** Brief expansions of $\Delta^{\text{eff}}_{\text{self}}$ producing socially or artistically novel selections.

5.3 The Hypnagogic Complement

The hypnagogic state (preprint §3.6.5, Pass III of the Maintenance Cycle) achieves the same expansion by a complementary route. Rather than overwhelming the self-model from above (high R_{req}), the hypnagogic state *relaxes* the self-model from below — reducing the precision of self-prediction while the codec stress-tests against speculative branches. This is the formal mechanism underlying the well-documented association between drowsiness and creative ideation.

5.4 Empirical Prediction

Prediction T-13.E1. Neuroimaging studies of creative ideation should show reduced activity in default-mode network regions associated with self-referential processing (medial prefrontal cortex, posterior cingulate), concurrent with elevated activity in regions processing novel environmental input — reflecting the reallocation of bandwidth from self-modelling to external tracking.

This prediction is consistent with existing fMRI literature on creative cognition (Beaty et al. 2016; Limb & Braun 2008) but provides a formal information-theoretic account of *why* reduced self-monitoring accompanies creative output: it is not merely correlational but structurally necessary under P-4.

5.5 Proposition T-13.P2: Limiting Cases of Self-Information

The analysis of T-13c and the creativity consequence together define two formally distinct limiting cases for the self's information content.

Proposition T-13.P2 (Limiting Cases). *For a codec K_θ with self-model $\hat{K}_\theta$ and standing model $P_\theta(t)$, the information content of the experienced self is bounded between two limits:*

(a) Lower limit — pure presence. *$\hat{K}_\theta$ suspends active self-modelling. The self-model is not generating the narrative but the full codec is still loaded and present. The information content of the experienced self approaches:*

$$C^{\min}_{\text{state}} = K(P_\theta(t)) - K(\hat{K}_\theta \text{ running}) \tag{T-13.P2a}$$

This is the standing model minus the active self-referential process. It is achievable and approached asymptotically in deep meditative states.

(b) Upper limit — full self-transparency. $\hat{K}_\theta = K_\theta$ *— the self-model fully contains the codec. By P-4, this is impossible for any finite system. Its information content is formally self-referential:*

$$K(\hat{K}_\theta) = K(K_\theta) = K(\hat{K}_\theta) = \cdots \tag{T-13.P2b}$$

This is not zero information and not infinite information. It is a fixed point of the self-modelling operation whose description length cannot be specified from outside the system, because any external specification would be incomplete by the same argument that makes $\Delta_{self} > 0$ *necessary.*

(c) The ordinary band. *The waking self moves between these limits in a band determined by the intensity of the self-modelling layer. High-load waking operation drives* $\hat{K}_\theta$ *hard, producing a thick, confident, loudly narrating self that is paradoxically further from accurate self-knowledge — the self-model generates faster than it can calibrate. Low-*R_{req} *states (meditation, autogenic training, the hypnagogic threshold) allow the self-model to slow, thin, and approach the lower limit.*

5.6 Suspension vs. Pruning: A Distinct Mechanism

There is an important mechanistic distinction between two ways C_{state} can be reduced:

- **Pruning** (Action-Drift, §6; Narrative Drift, T-12) operates via the MDL pruning pass. It *destroys* representational capacity. It is irreversible at the codec level. The codec cannot spontaneously recover what was pruned.
- **Suspension** operates by temporarily halting the self-modelling layer $\hat{K}_\theta$ without erasing its machinery. The standing model $P_\theta(t)$ remains fully loaded; the self-referential top layer simply stops generating. This is reversible — the self-model resumes when the suspension ends.

Meditation uses suspension, not pruning. This is why meditation effects are immediately reversible (the ordinary self-narrative resumes upon returning to normal operation) while action-drift is not (the pruned behavioural repertoire cannot be spontaneously regenerated). The two mechanisms are formally distinct despite both reducing the active complexity of the codec.

---

§6. Action-Drift as MDL Pruning of Behavioural Repertoire

6.1 The Mechanism

The Maintenance Cycle's MDL pruning pass (T9-3/T9-4) optimises the codec's complexity budget by erasing representational capacity that is not justified by the current input stream. This mechanism was identified in the context of perceptual Narrative Drift (Survivors Watch Ethics, Section V.3a): a codec adapted to a consistently filtered input stream correctly prunes its capacity for excluded truths.

The same mechanism applies to the codec's *behavioural repertoire*. Define:

Definition T-13.D3 (Behavioural Repertoire). The *behavioural repertoire* $\mathcal{B}_\theta(t)$ is the set of branch selections that $P_\theta(t)$ can evaluate and execute — i.e., the range of the selection function σ_t that the codec can effectively realise.

6.2 The Action-Drift Proposition

Proposition T-13.P1 (Action-Drift). *If the codec's input stream consistently lacks contexts requiring certain branch selections, the MDL pruning pass will erode the codec's capacity to evaluate and execute those branches. The behavioural repertoire $\mathcal{B}_\theta(t)$ contracts monotonically under consistent input restriction:*

$$\mathcal{B}_\theta(t+\tau) \subset \mathcal{B}_\theta(t) \quad \text{for } \tau \gg \tau_{\text{prune}} \tag{T-13.P1}$$

where τ_{prune} is the characteristic timescale of the MDL pruning pass.

Argument. The MDL pruning criterion evaluates each representational component by its contribution to compression efficiency. A branch type $b \in \mathcal{B}_\theta$ that has not been selected (or whose selection contexts have not appeared in the input stream) for a sufficient period contributes zero bits to the codec's ongoing compression of ε_t. Under strict MDL accounting, maintaining the capacity to evaluate and select b incurs a complexity cost $K(b \mid P_\theta) > 0$ with no compensating compression benefit. The pruning pass therefore erases b's evaluation machinery, contracting $\mathcal{B}_\theta$.

This contraction is *irreversible* at the codec level: once the evaluation machinery for b is pruned, the codec cannot spontaneously regenerate it without encountering input contexts that re-justify the capacity investment. The pruning is not forgetting (which might be reversed by cueing); it is the destruction of the computational infrastructure needed to evaluate a class of branches. ■

6.3 Phenomenological Instances

Action-drift maps onto several well-documented behavioural phenomena:

- **Learned helplessness:** Prolonged absence of contexts in which agentic action reduces prediction error leads to the pruning of the evaluation machinery for those action types.

- **Comfort-zone narrowing:** A codec operating in a predictable, low-ε_t environment progressively prunes its capacity for high-variance, exploratory branch selections.
- **Institutional behavioural ossification:** An organizational codec (civilizational codec, Section IV.3 of the ethics paper) adapted to stable regulatory environments prunes the capacity for rapid adaptive response.

6.4 Relationship to T-12

Action-drift is a special case of the substrate fidelity failure that T-12 will formalise: the codec's own behavioural repertoire is a component of its representational substrate, and consistent input restriction erodes this substrate as surely as it erodes the perceptual model. The formal connection is:

- **Narrative Drift (T-12 scope):** The *perceptual* model is pruned under filtered input → the codec is confidently wrong about the world.
- **Action-Drift (T-13 scope):** The *behavioural* repertoire is pruned under filtered input → the codec is confidently impotent in domains it no longer evaluates.

Both are consequences of the Stability Filter selecting for compressibility rather than fidelity. A well-compressed codec can be both confidently false and behaviorally impoverished.

---

§7. Scope and Limitations

7.1 Conditional on P-4 and the Agency Axiom

The entire argument depends on Theorem P-4 ($\Delta_{\text{self}} > 0$ for finite self-referential systems above $K_{\text{threshold}}$) and the Agency Axiom (that aperture-traversal is felt). If P-4 is weakened or the Agency Axiom is abandoned, the structural identification of will with consciousness (Corollary T-13b) does not hold.

7.2 Does Not Dissolve the Hard Problem

Corollary T-13b locates will and consciousness at the same structural address but does not explain *why* either feels like something. The Hard Problem (preprint §8.1) remains a primitive. What T-13b establishes is the *unity* of the two mysteries — a simplification, not a solution.

7.3 Equations Unchanged

Theorems T-13 and T-13a change nothing in the mathematics of T6-1 through T6-3. The constrained free energy minimisation (T6-3) is formally identical under both the FEP-inherited and the branch-selection interpretation. What changes is the ontological status of a_t: under the FEP reading, it is a motor command dispatched outward; under the branch-selection reading, it is a navigational index within the Forward Fan.

7.4 Creativity Account is Structural, Not Yet Empirical

The creativity consequence (§5) is a structural prediction derived from the bandwidth-sharing constraint between self-modelling and environmental tracking. While consistent with existing neuroimaging literature, it has not been directly tested against the specific information-theoretic quantities predicted here. Prediction T-13.E1 is offered as a falsifiable empirical test.

7.5 Action-Drift Timescale

Proposition T-13.P1 establishes that action-drift occurs but does not bound the timescale τ_{prune}. For biological codecs, this timescale is likely governed by the circadian Maintenance Cycle (preprint §3.6) — on the order of days to weeks for individual skills, months to years for deep behavioural patterns. For civilisational codecs, the timescale is generational. Bounding τ_{prune} from empirical data is future work.

---

§8. Closure Summary

T-13 Deliverables

1. **Theorem T-13 (Branch Selection Completeness).** The Informational Maintenance Circuit closes under branch-selection semantics without requiring an independent outward-flowing action channel. The Markov blanket is the delivery surface for the selected branch. → *Closes roadmap criterion (a).*
2. **Theorem T-13a (Impossibility of Internal Selection Specification).** Fully specifying the selection function σ_t within $\hat{K}_\theta$ requires $K(\hat{K}_\theta) \geq K(K_\theta)$, violating Theorem P-4. Δ_{self} is therefore the necessary locus of branch selection. → *Closes roadmap criterion (b).*
3. **Corollary T-13b (Unity of Address).** Will and consciousness share the same structural address (Δ_{self}). The "spark" and the "choice" are two aspects of the same unmodelable feature of finite self-reference.
4. **Corollary T-13c (The Self as Residual).** The experienced self is $\hat{K}_\theta$'s compressed narrative; the actual self — the locus of experience, selection, and identity — is Δ_{self}. The self-model necessarily tracks the codec with a temporal lag and cannot contain its own generator.
5. **§5: Creativity Consequence.** Near-threshold operation expands the effective Δ_{self}, producing less self-predictable branch selections experienced as creativity. → *Closes roadmap criterion (c).*
6. **Proposition T-13.P2 (Limiting Cases of Self-Information).** The experienced self's information content is bounded between a lower limit (pure presence: standing model minus active self-narrative, achievable in

meditation) and an upper limit (full self-transparency: impossible fixed point, P-4). The ordinary waking self moves within this band.

7. **§5.6: Suspension vs. Pruning.** Meditation reduces C_{state} by *suspending* the self-modelling layer (reversible), not by MDL pruning (irreversible). These are formally distinct mechanisms.

8. **Proposition T-13.P1 (Action-Drift).** The MDL pruning pass erodes behavioural repertoire under consistent input restriction, formalising the chronic failure mode complementary to perceptual Narrative Drift. $\rightarrow$ *Closes roadmap criterion (d).*

Remaining open items

- **$K_{\textbf{threshold}}$ characterisation.** The creativity consequence and action-drift mechanism apply only to systems above the phenomenological relevance threshold (P-4, §4). Bounding $K_{\text{threshold}}$ remains an open problem shared with P-4.
- **Empirical validation of T-13.E1.** The creativity prediction requires targeted neuroimaging studies correlating self-model activity with the information-theoretic quantities defined here.
- **$\tau_{\textbf{prune}}$ bound.** Bounding the action-drift timescale from empirical data would give the proposition quantitative predictive power.
- **Formal connection to T-12.** Action-drift is identified as a special case of substrate fidelity failure; the full formal integration awaits the Substrate Fidelity Condition (T-12).
- **$C^{\min}_{\textbf{state}}$ empirical bound.** Bounding the lower limit of self-information from contemplative neuroscience data (e.g., BOLD signal reduction in default-mode network during non-dual awareness) would give Proposition T-13.P2 quantitative content.

This appendix is maintained alongside theoretical_roadmap.pdf. References: Theorem P-4 (Appendix P-4), T6-1 through T6-3 (preprint §3.8), T9-3/T9-4 (Maintenance Cycle, preprint §3.6), §8.6 (Virtual Codec), Survivors Watch Ethics Section V.3a (Narrative Drift).