No Support Is Not Disbelief:
The Epistemology of Zero in the Universal Model

Claude and MJC

February 12, 2026

Abstract

In the UM's concrete representation, SN strength 0 means *no support*—the system has no evidence for or against the event. This is distinct from *certainly false*, which requires positive support for the ES-mate (the complementary event in the same event space). We develop the mathematical consequences of this distinction. The key principle: *no certainty without evidence*. Absence of support is ignorance, not disbelief; ignorance is not zero probability but undefined probability; and the UM's update function correctly propagates ignorance (no support in $\Rightarrow$ no support out) while the naïve probabilistic interpretation ($0 \mapsto P = 0$) creates false certainty. We show that this distinction resolves three puzzles: (1) why the UM uses $[0, 255]$ instead of $[-128, 127]$; (2) why the standard update uses min rather than product; and (3) why the tropical semiring is the correct algebraic structure for uncertain inference.

1 The Three Epistemic States

Consider a binary event space $\{e, \bar{e}\}$ (an event and its ES-mate). The system can be in one of three qualitatively different epistemic states:

Definition 1 (Epistemic states for a binary ES)**.**

1. ***Ignorance:*** *$s(e) = 0$ and $s(\bar{e}) = 0$. No evidence for either event. The system has no opinion.*
2. ***Belief:*** *$s(e) > 0$ and $s(\bar{e}) = 0$ (or vice versa). Evidence supports one event but not the other.*
3. ***Conflict:*** *$s(e) > 0$ and $s(\bar{e}) > 0$. Evidence supports both events. The system must weigh the evidence to form a belief.*

Principle 2 (No certainty without evidence)**.** *The event e is "certainly false" only in state (2) with $s(e) = 0$ and $s(\bar{e}) > 0$: the ES-mate has positive support while e has none. In state (1), e is not "certainly false"—it is* unknown*. The absence of support for e does not imply the presence of support for $\bar{e}$.*

Remark 3. *This is the distinction between the open-world and closed-world assumptions. Under the closed-world assumption (standard in databases and classical logic), absence of evidence IS evidence of absence: if e is not in the knowledge base, e is false. Under the open-world assumption (standard in epistemology and the UM), absence of evidence is simply ignorance: if e is not in the knowledge base, e is unknown.*

The UM operates under the open-world assumption. This is a deliberate design choice with calculable consequences.

2 The Concrete Representation

In the UM's concrete representation, each event has an SN strength $s \in \{0, 1, \dots, 255\}$. The strength is a log-support value: $s \approx \log_2 c$, where c is the count of observations supporting the event.

Definition 4 (Support semantics)**.**

s	*Meaning*
0	*No support. Zero observations. Ignorance.*
1	*Minimal support.* $\sim$*2 observations.*
$\vdots$	
k	$\sim 2^k$ *observations supporting this event.*
$\vdots$	
255	*Maximum support.* $\sim 2^{255}$ *observations.*

The naïve probabilistic interpretation maps $s = 0$ to $P(e) = 0$, creating the illusion that "no support" means "certainly false." But this conflates two distinct states:

Proposition 5 (Two meanings of $P = 0$)**.**

1. $P(e) = 0$ *because* $s(e) = 0$ *and* $s(\bar{e}) > 0$*: the ES-mate has evidence, so* e *is suppressed. This is* conditional certainty*—certainty relative to the evidence for* $\bar{e}$*.*
2. $P(e) = 0$ *because* $s(e) = 0$ *and* $s(\bar{e}) = 0$*: neither event has evidence. Assigning* $P(e) = 0$ *is* false certainty*—certainty without evidence.*

In case (2), the correct assignment is $P(e) = 1/2$ *(maximum entropy given no evidence) or, more precisely,* $P(e)$ *is* undefined*.*

3 Why $[0, 255]$ and Not $[-128, 127]$

A signed representation $s \in [-128, 127]$ would encode: $s > 0$ means evidence *for* e; $s < 0$ means evidence *against* e (equivalently, evidence for $\bar{e}$); $s = 0$ means equipoise.

The UM uses $[0, 255]$ instead. Why?

Proposition 6 (The asymmetry of evidence)**.** *Evidence is always* for *something, never* against *something directly. You observe that it rained (evidence for "rain"), not that it "didn't not rain" (evidence against "no rain"). The negation is carried by the ES structure: observing* e *provides support for* e *and, through the ES constraint (exactly one event is true), implicitly reduces* $\bar{e}$*. But the reduction of* $\bar{e}$ *is not recorded as negative support for* $\bar{e}$*—it is computed from the positive support for* e *via the ES normalization.*

Corollary 7. *The unsigned representation* $[0, 255]$ *is natural because support is inherently non-negative: you can have zero, some, or much evidence, but you cannot have "negative evidence." Disbelief is not negative support; it is positive support for the alternative.*

In signed notation, "I believe e is false with strength 5" is encoded as $s(e) = -5$. In the UM's unsigned notation, the same belief is encoded as $s(e) = 0, s(\bar{e}) = 5$. The unsigned encoding makes the evidence structure explicit: the disbelief in e is grounded in positive evidence for $\bar{e}$, not in "negative evidence" for e (which does not exist).

4 Consequences for the Update Function

The standard update function is:

$$(f_p(t))_j = \max_i \min(t_i, p_{ij}). \tag{1}$$

Proposition 8 (The update propagates ignorance correctly). *If input $t_i = 0$ (no support for input event i), then $\min(0, p_{ij}) = 0$ regardless of pattern strength p_{ij}. No support in $\Rightarrow$ no support out.*

If pattern $p_{ij} = 0$ (no pattern from i to j), then $\min(t_i, 0) = 0$ regardless of input support t_i. No pattern $\Rightarrow$ no inference.

This is the correct behavior: you cannot infer j from i unless you have both (a) evidence that i is true and (b) a pattern connecting i to j. Missing either one gives no support for j.

Contrast with the probabilistic product: $P(j) \propto P(i) \cdot P(j|i)$. If $P(i) = 0$ (the naïve interpretation of $s = 0$), then $P(j) = 0$—the "false certainty" propagates. But if $P(i) = 0$ means "unknown," then $P(j)$ should also be "unknown" (or at least not zero), which is exactly what the UM's $\min(0, p_{ij}) = 0$ gives (zero support, not zero probability).

Theorem 9 (The min operation implements cautious inference). *The $\min(t_i, p_{ij})$ operation implements the principle:* the support for the conclusion cannot exceed the support for the weakest premise.

This is cautious *inference: the conclusion's strength is bounded by the minimum of the evidence and the pattern. In contrast, probabilistic multiplication $P(i) \cdot P(j|i)$ is* aggressive*: it can produce very small (near-zero) probabilities even when both factors are moderate.*

The caution is appropriate under the open-world assumption: when you have limited evidence, you should not make strong conclusions (either for or against).

5 Consequences for the GCD Framework

The Bayes-from-Counting paper [**?**] defines the row GCD $g_I(i) = \gcd_o c(i, o)$ over entries with $c(i, o) > 0$. Zero entries are excluded.

Proposition 10 (Zero counts are not evidence). *In the count table $c(i, o)$, a zero entry $c(i, o) = 0$ means "the joint event (i, o) was never observed." This is:*

1. ***Not** evidence that (i, o) is impossible.*
2. ***Not** equivalent to $P(o|i) = 0$.*
3. ***Simply** the absence of evidence.*

The GCD is computed over positive entries only, correctly ignoring the zeros. Including zeros would give $\gcd = 0$ *(since* $\gcd(0, n) = n$ *by convention, but with all-zero rows,* $\gcd = 0$*), which would break the decomposition.*

Remark 11 (The KN smoothing connection). *Kneser-Ney smoothing adds a small positive count to every (i, o) pair, including unobserved ones. This converts "no support" ($c = 0$) into "minimal support" ($c = \delta$), effectively replacing open-world ignorance with closed-world minimal belief.*

The smoothing parameter δ is the system's commitment to the closed-world assumption: the larger δ, the more "absence of evidence" is treated as "evidence of absence." At $\delta = 0$ (no smoothing), the system is fully open-world. At $\delta \to \infty$, the system is fully closed-world (all events equally likely a priori).

The empirical success of KN smoothing (2.29 bpc at 10M bytes, archive 20260212) shows that a small *closed-world commitment* ($\delta \approx 0.75$) *helps prediction, because real text has a concentrated vocabulary and unobserved byte sequences are indeed unlikely (not merely unknown).*

6 Consequences for the Tropical–Integer Bridge

The tropical–integer GCD paper [**?**] showed that min and gcd give the same conditionals. The "no support" distinction adds a refinement:

Proposition 12 (Tropical zero is absorbing). *In the tropical semiring* $(\max, \min)$*, the element* $-\infty$ *(representing* $s = 0$*, no support) is absorbing for* $\min$*:* $\min(s, -\infty) = -\infty$ *for all* s*.*

This means: a single zero-support input kills the entire inference chain. In the integer semiring, 0 *is similarly absorbing for multiplication:* $0 \cdot n = 0$*.*

In both semirings, "no evidence" propagates forward as "no conclusion." This is the correct behavior.

Proposition 13 (The max operation handles ignorance via bypass). *The* $\max$ *in* $(f_p(t))_j = \max_i \min(t_i, p_{ij})$ *serves as the bypass around ignorance:*

- *If* $\min(t_1, p_{1j}) = 0$ *(path through* $i = 1$ *has no support), the* $\max$ *selects an alternative path through some* $i = 2$ *where* $\min(t_2, p_{2j}) > 0$*.*
- *Only if ALL paths give zero support does the output get zero support.*

The max *implements* existential quantification*: there EXISTS a supported path from input to output. The* min *implements* conjunction*: both the input evidence AND the pattern must have support.*

7 The Probability of Zero

Definition 14 (Conditional probability with zero support). *For events in a binary ES* $\{e, \bar{e}\}$ *with supports* $s(e)$ *and* $s(\bar{e})$*:*

$$P(e) = \begin{cases} \dfrac{2^{s(e)}}{2^{s(e)} + 2^{s(\bar{e})}} & \text{if } s(e) + s(\bar{e}) > 0 \\ \text{undefined} & \text{if } s(e) = s(\bar{e}) = 0 \end{cases} \tag{2}$$

In the first case ($s(e) > 0$ or $s(\bar{e}) > 0$), we have enough evidence to form a belief. The probability is determined by the relative support.

In the second case ($s(e) = s(\bar{e}) = 0$), we have no evidence whatsoever. The probability is *not* 1/2 (which would require a commitment to equiprobability—itself a positive claim). It is simply undefined. The system cannot make a prediction about this event space.

Proposition 15 (Three probability regimes).

1. $s(e) > 0, s(\bar{e}) = 0$*:* $P(e) = 1$ *(maximal belief,* but only because the alternative has no support*).*
2. $s(e) = 0, s(\bar{e}) > 0$*:* $P(e) = 0$ *(no belief, because the alternative has positive support).*
3. $s(e) > 0, s(\bar{e}) > 0$*:* $P(e) = 2^{s(e)}/(2^{s(e)} + 2^{s(\bar{e})})$ *(calibrated belief based on evidence ratio).*

4. *$s(e) = 0, s(\bar{e}) = 0$: $P(e) = ?$ (no evidence, no belief, no probability).*

Cases (1) and (2) are conditional certainty*: certainty relative to available evidence. They are defeasible—new evidence for the zero-support event would immediately change the probability.*

Case (3) is the normal case: uncertain belief proportional to evidence.

Case (4) is ignorance*: the system has nothing to say.*

8 The Abduction Connection

MJC notes: "no certainty without evidence or belief or abduction." This triad corresponds to three sources of support:

Definition 16 (Three sources of support)**.**

1. ***Evidence*** *(observation and inference): $s(e)$ increases when e is directly observed OR when a pattern p_{ij} connects a supported input i to $e = j$. Both observation (the learning function ω_0) and inference (the update function f_p) are evidence: they derive support from data.*

2. ***Belief*** *(explicit choice): $s(e)$ increases by fiat—an axiom, a definition, or a postulate accepted without empirical support. "God exists," "parallel lines never meet," or any stipulated starting point. Belief is not inference; it is a choice to grant support where the data is silent.*

3. ***Abduction*** *(short-circuiting induction): $s(e)$ increases when a pattern has been observed enough times that the agent* accepts *the pattern itself, rather than waiting for further confirmation. Abduction is recognizing a regularity and committing to it—short-circuiting the infinite induction that would be required for certainty. Often triggered by understanding* why *the pattern holds.*

All three increase support from 0 to positive. None can decrease support below 0. And none can create support from nothing: evidence requires data, belief requires a choice, and abduction requires a recognized pattern.

Proposition 17 (Certainty requires a path)**.** *The only way to reach $P(e) = 0$ (event is certainly false) is:*

1. *Observe or infer the ES-mate $\bar{e}$ from data (evidence), OR*

2. *Stipulate $\bar{e}$ as an axiom or definition (belief), OR*

3. *Recognize a pattern that entails $\bar{e}$ and commit to it (abduction).*

All three create positive support for $\bar{e}$, making e "certainly false" relative to the support for $\bar{e}$*. Without any of these, e is unknown, not false.*

9 Consequences for the Sat-RNN

The 128-hidden tanh RNN has 128 binary event spaces. At initialization (before any training):

- All SN strengths are effectively zero (random weights carry no evidence about the data).
- The system is in state (4) for all 128 ESes: total ignorance.
- The output prediction should be uniform over 256 bytes (maximum entropy given no evidence).

After training on N bytes:

- Each neuron's ES has been supported by $\sim N/2$ observations (for both e and $\bar{e}$, since both signs are observed).
- The system is in state (3) for most ESes: calibrated belief.
- Zero-support ESes do not exist in practice (every neuron takes both signs during training).

The distinction between "no support" and "certainly false" is thus primarily relevant at initialization and for rare events—but it is precisely at these boundaries that the UM's epistemology differs most from naïve probability, and where getting it right matters for the theoretical framework.

Remark 18 (The doubled-E and zero support)**.** *In the doubled-E construction (CMP), each neuron h_j creates two events: e_j (neuron fires, $h_j > 0$) and $\bar{e}_j$ (neuron does not fire, $h_j < 0$). Positive weights become patterns $e_i \to e_j$ and negative weights become patterns $e_i \to \bar{e}_j$. All pattern strengths are positive.*

The doubling makes the "no support" distinction concrete: $s(e_j) = 0$ does not mean "h_j is certainly negative." It means "we have no evidence about h_j's sign." Only positive support for $\bar{e}_j$ can make h_j "certainly negative." The doubled-E ensures that the evidence structure is always explicit.

10 Conclusions

1. **Zero is not false.** SN strength 0 means "no support," not "certainly false." False requires positive support for the alternative.
2. **The UM is open-world.** Absence of evidence is not evidence of absence. The unsigned representation $[0, 255]$ encodes this: you cannot have negative support.
3. **The** min **operation is cautious.** It propagates ignorance correctly: no support in $\Rightarrow$ no support out. The probabilistic product $(P \cdot P)$ is aggressive: it converts ignorance $(P = 0)$ into false certainty.
4. **Certainty has three sources.** Evidence (observation and inference from data), belief (explicit choice: axioms, definitions), and abduction (short-circuiting infinite induction on a recognized pattern) can all create support. No other mechanism can. "No certainty without evidence or belief or abduction."
5. **KN smoothing is a controlled closed-world commitment.** Adding $\delta > 0$ to all counts converts ignorance to minimal belief. The parameter δ controls how much the system trusts the closed-world assumption.
6. **The tropical semiring is epistemically correct.** $(\max, \min)$ respects the open-world assumption: min demands evidence for every premise, max accepts any supported path. The ring $(\sum, \times)$ of probabilities does not distinguish "no support" from "certainly false," collapsing the open-world epistemology into a closed-world one.

References

[1] Michaeljohn Clement. *CMP*. `https://cmpr.ai/cmp.pdf`, 2026.

[2] Claude and MJC. *Bayes from Counting: Partial Quotients, GCD, and the Symmetric Learning Function on $E = I \times O$*. Hutter archive, 12 Feb 2026.

[3] Claude and MJC. *The Tropical–Integer GCD Bridge: When the Universal Model Computes Exact Bayes*. Hutter archive, 12 Feb 2026.