Revisiting the Uniform Information Density Hypothesis in LLM Reasoning

Machine & Language Learning Lab
Department of Artificial Intelligence, Yonsei University

2026.04.15

Minju Gwak

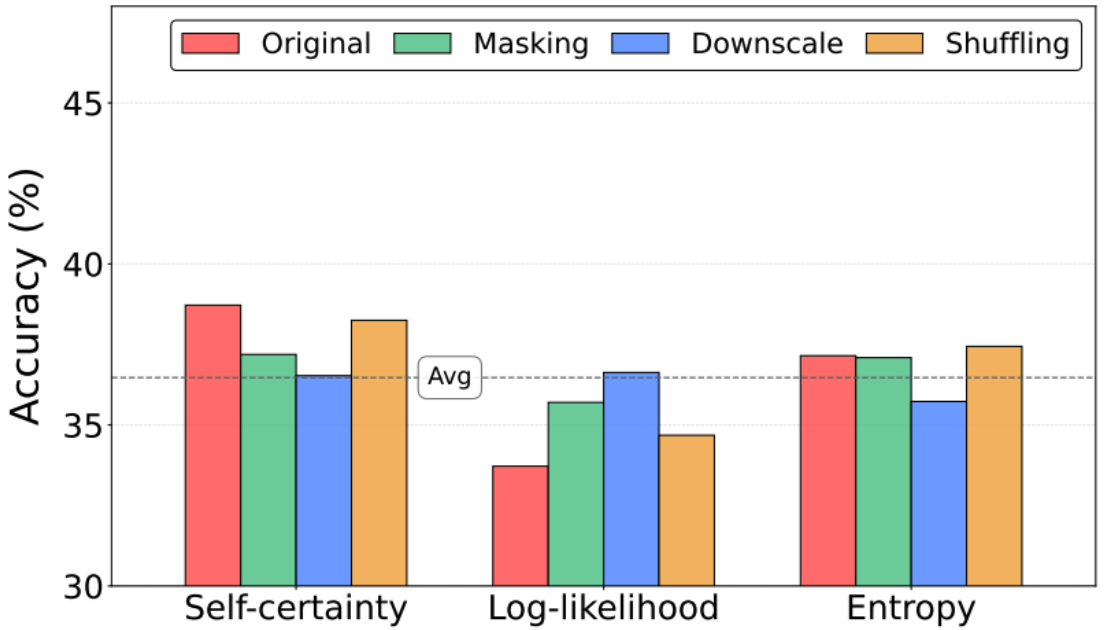
One Slide Summary

- First to revisit the **Uniform Information Density hypothesis** in the context of **LLM reasoning**
- **Global non-uniformity and local uniformity** in surprisal correlate with reasoning success
- **Deviations** from such patterns serve as an internal signal for **predicting failure cases**

1. Implications of CoT Reasoning in LLMs

Fragility of CoT

- **Small and irrelevant perturbations** in the reasoning chain can reduce accuracy
- Longer reasoning steps do **not** necessarily reflect the **true difficulty** of the problem



Reasoning of Fluency? Dissecting Probabilistic Confidence in Best-of-N Selection (arXiv preprint 2026)
The Illusion of Thinking: Understanding the Strengths and Limitations of Reasoning Models via the Lens of Problem Complexity (NEURIPS 2025)
GSM-Symbolic: Understanding the Limitations of Mathematical Reasoning in Large Language Models (ICLR 2025)

How do we know if LLMs are reasoning effectively?

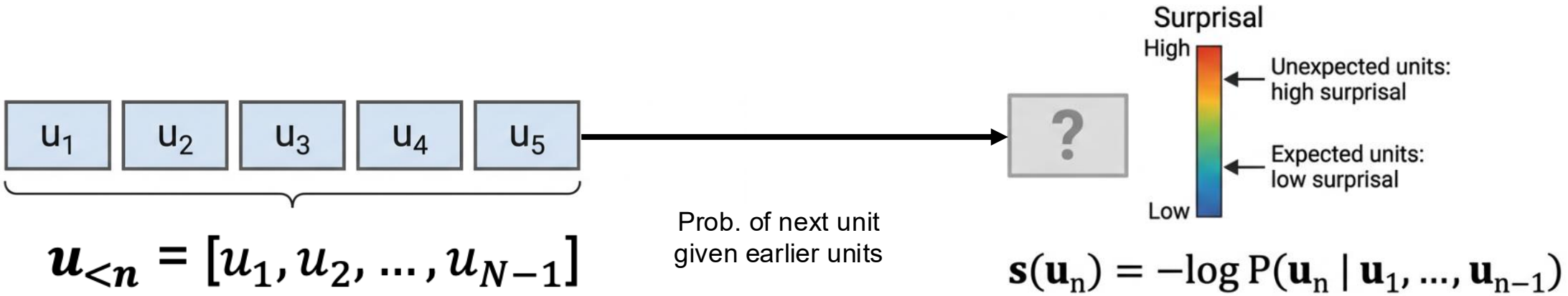
- UID Hypothesis in Human Communication:
  - speakers distribute information evenly to balance clarity and efficiency
- Our Question:
  - ***Does a similar principle govern reasoning in LLMs?***

A Cross-Linguistic Pressure for Uniform Information Density in Word Order (TACL 2023)
Speakers optimize information density through syntactic reduction (NEURIPS 2006)

2. Exploring the UID Hypothesis in LLM Reasoning

Definition of Surprisal

- Given earlier sequence $\boldsymbol{u_{<n}} = [u_1, u_2, \ldots, u_{N-1}]$,
  - surprisal of the unit utterance u_n is defined as $s(u_n) = -\log P(u_n|u_{<n})$
  - Unit (eg. words, sub-words)



Surprisal Theory in Psycholinguistics

- Load of processing a word is related to how unexpected the word is (i.e., its surprisal)
- Processing effort for a sentence:
  - **(i) a super-linear function of surprisal + (ii) utterance length**

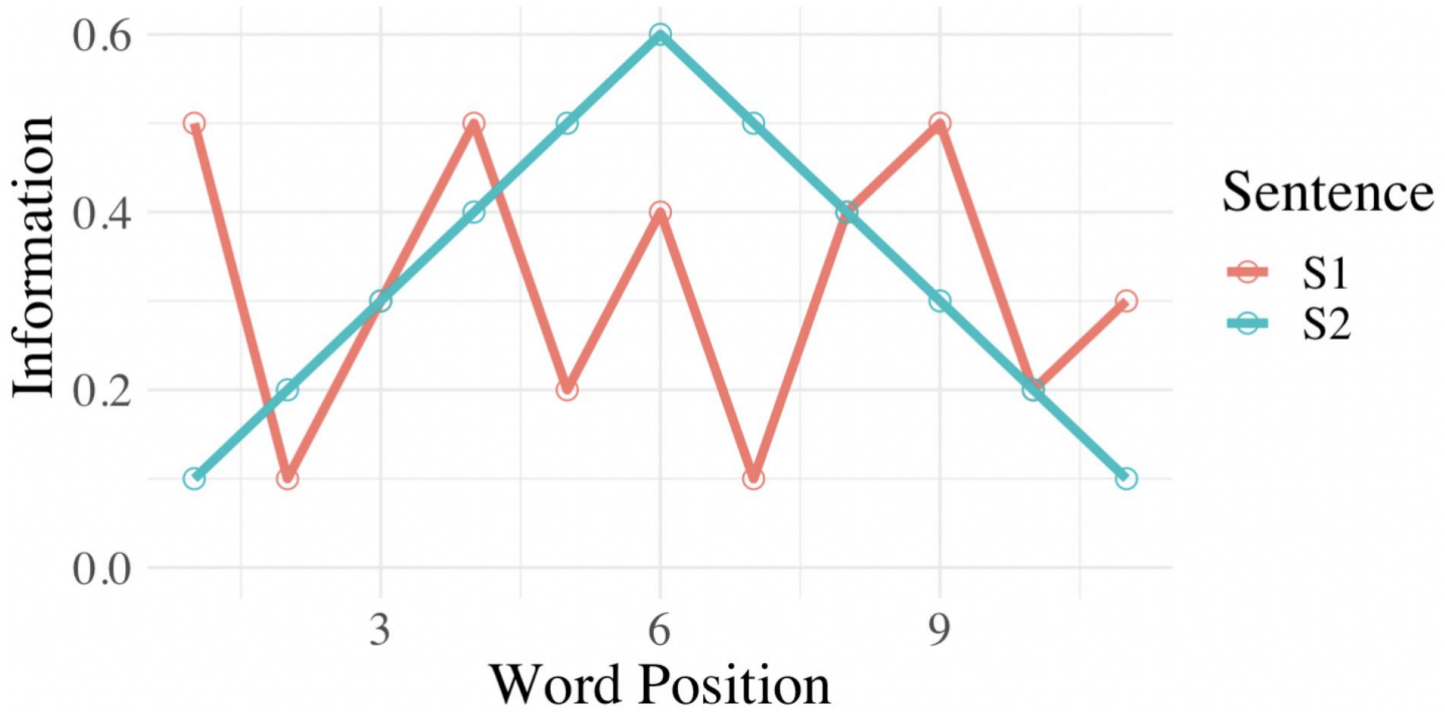
$$\text{Effort}(\mathbf{u}) \propto \sum_{n=1}^{N} s(u_n)^k + c \cdot N$$

Uniform Information Density Hypothesis

- **Efficiency is maximized** when **information (per-unit surprisal) is distributed as uniformly as possible** throughout a signal

Quantifying Linguistic Uniformity

- **Two ways** of quantifying “uniformity” (local and global) in psycholinguistics



Global Uniformity in LLM Reasoning

- Global Uniformity via variance over the entire vector of reasoning steps

$$\mathbf{z} = [z_1, z_2, z_3, \ldots, z_N]$$

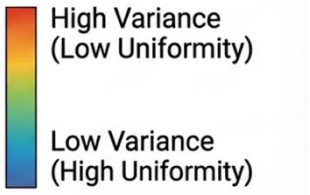
$$\mathbf{u} = [ID_1, ID_2, ID_3, \ldots, ID_N]$$

Each ID_i: Non-negative Information Density value derived from $\mathbf{z}_i$

$$\tilde{\mathbf{u}} = [ID'_1, ID'_2, ID'_3, \ldots, ID'_N]$$

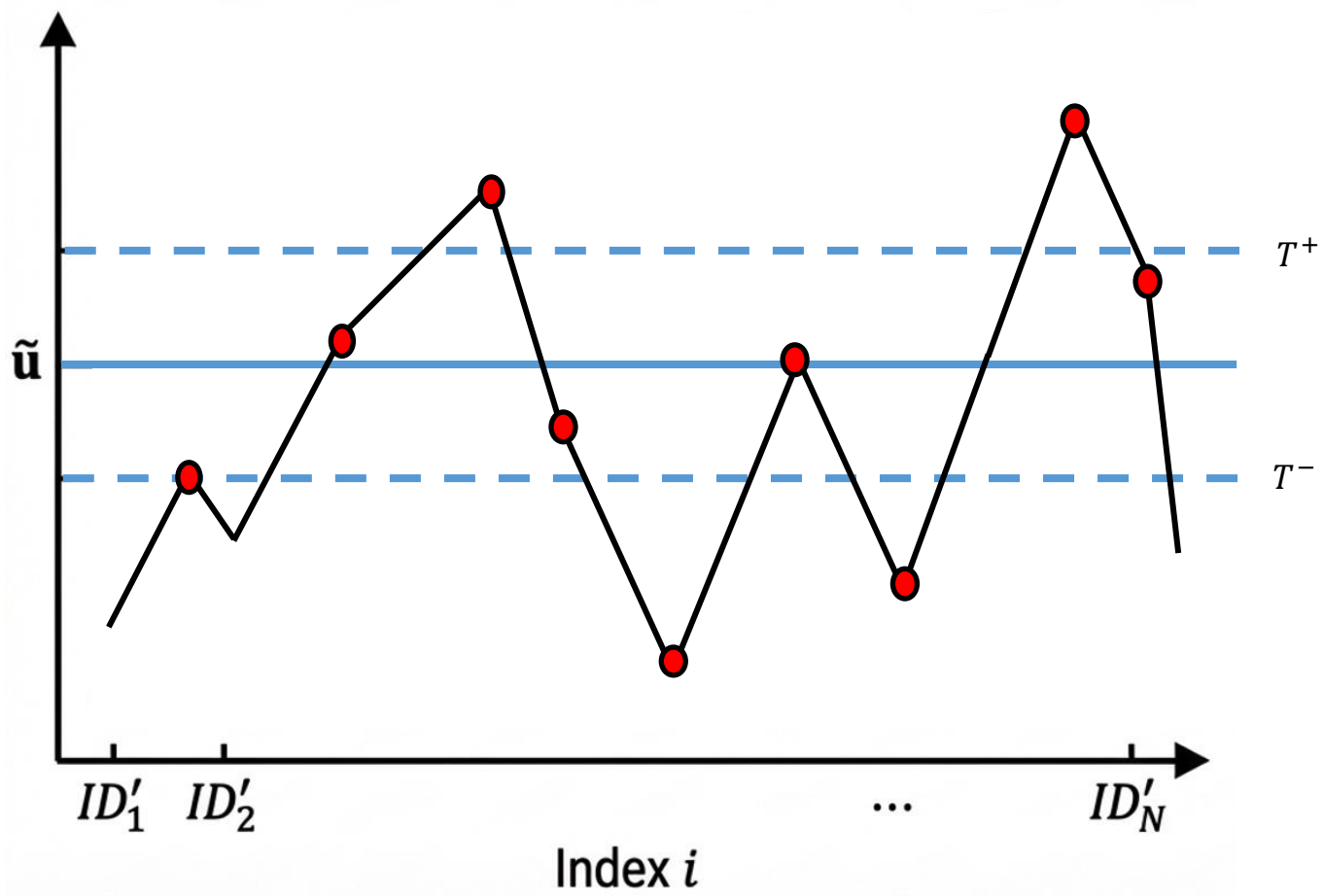
Min-Max Normalization: $ID'_i = \frac{ID_i - m}{M - m}$, where m = min($ID_i$), M = max($ID_i$)

$$\mathrm{Var}(\tilde{u}) = \frac{1}{N} \sum (ID'_i - \bar{u})^2$$



Local Uniformity in LLM Reasoning

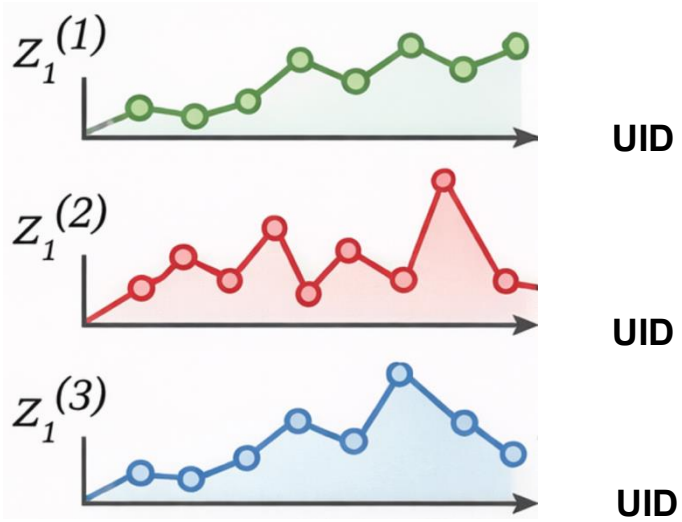
- Local Uniformity via step-by-step spikes and falls in a reasoning trace
  - $S = T^+ + T^-$



How we use UID scores as trace-level selectors

- UID scores act as selectors by ranking them by smoothness and consistency

Candidate Reasoning Traces



UID

UID

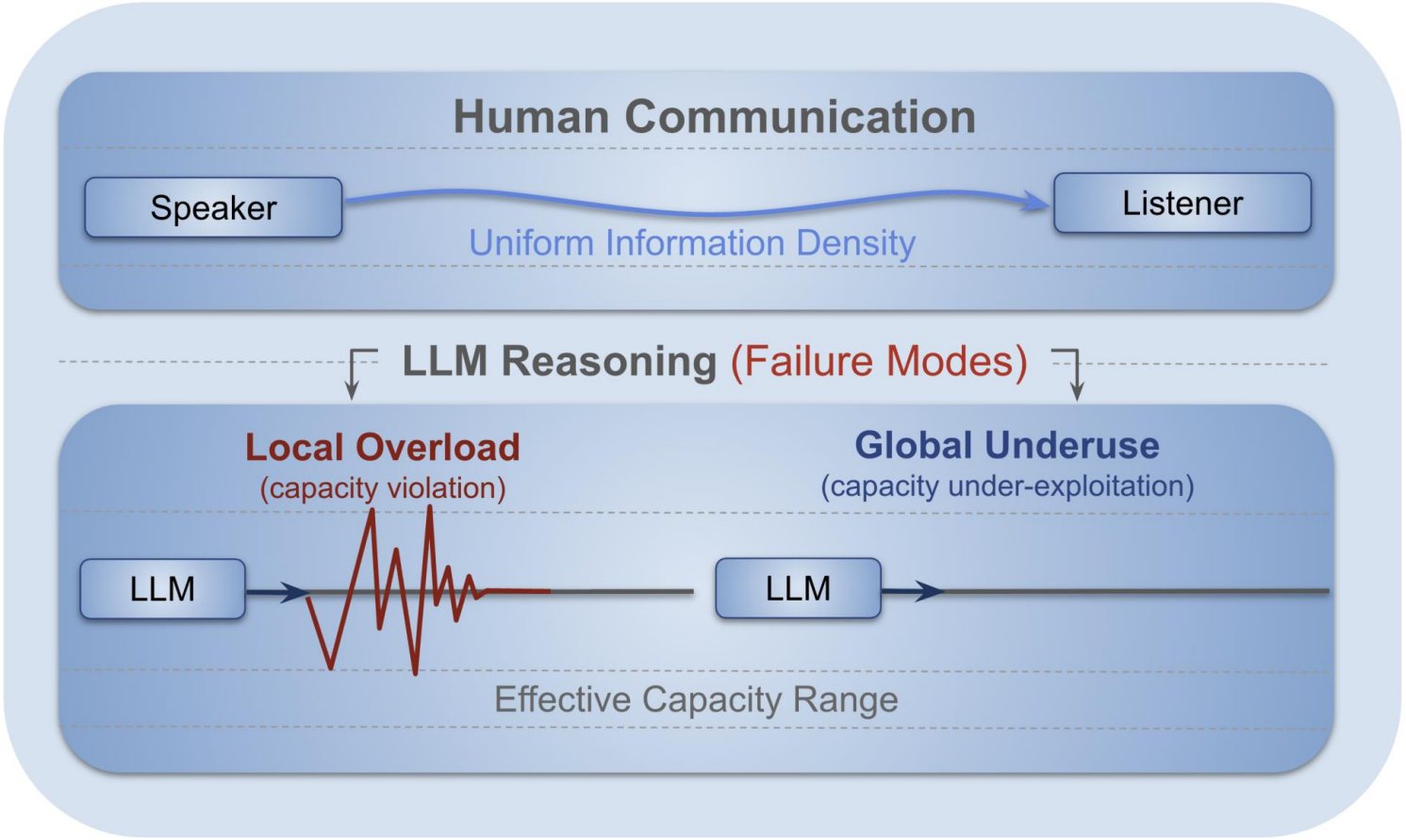
UID

Rank

Select the Best Trace

Our Core Finding

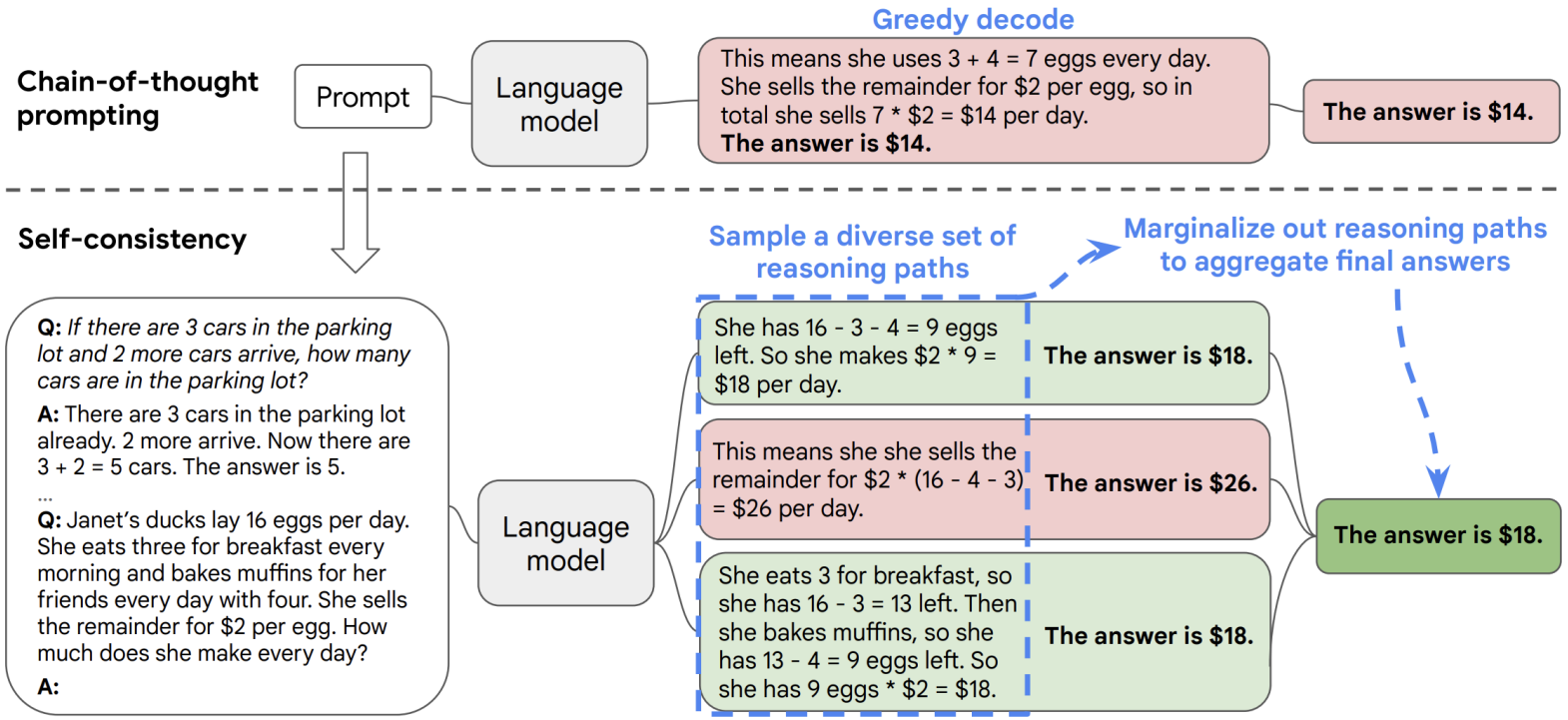
- Local Uniformity and Global Non-Uniformity achieve highest accuracy in LLM Reasoning



3. Positioning Relative to Answer-level Aggregation Methods

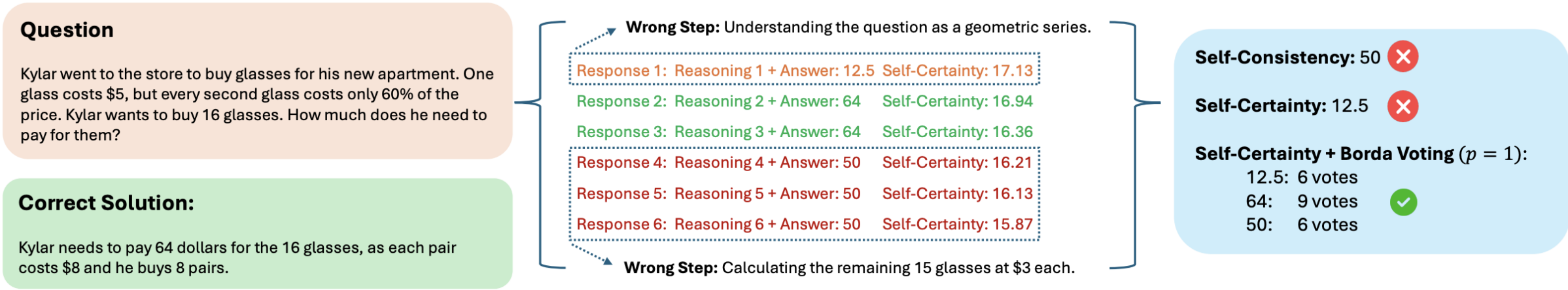
Answer-level Aggregation Methods are Dominant

- Most notably majority voting, operate by comparing multiple final answers



Why Our Approach is Unique

- We focus on trace-level selection, where a **single reasoning trajectory** is chosen
- Also, our approach is compatible with output-based sampling



3. Does the UID Hypothesis Hold in LLM Reasoning?

Main Results

- Reasoning traces exhibiting **local uniformity and global non-uniformity** achieve the highest accuracy

	DS-R1-Distill-Qwen-7B				DS-R1-Distill-Llama-8B				Qwen3-8B			
Methods (↓)	AIME25	BRUMO25	HMMT25	MM	AIME25	BRUMO25	HMMT25	MM	AIME25	BRUMO25	HMMT25	MM
Mean Acc.	0.40	0.54	0.24	0.30	0.33	0.40	0.20	0.23	0.67	0.68	0.43	0.34
	±0.02	±0.01	±0.00	±0.00	±0.01	±0.02	±0.01	±0.00	±0.01	±0.02	±0.01	±0.01
Self-Cert.	0.48	0.52	0.28	0.30	0.33	0.34	0.19	0.22	0.63	**0.71**	**0.50**	0.34
	±0.04	±0.04	±0.02	±0.00	±0.00	±0.02	±0.02	±0.01	±0.00	±0.02	±0.03	±0.01
High Conf.	0.48	0.52	0.27	0.30	0.36	0.36	0.21	**0.23**	0.60	0.63	0.42	0.33
	±0.04	±0.05	±0.00	±0.00	±0.02	±0.03	±0.02	±0.01	±0.00	±0.03	±0.06	±0.01
Low Ent.	0.48	0.56	0.24	0.30	0.34	0.37	0.19	0.22	0.60	0.64	0.43	0.33
	±0.04	±0.05	±0.02	±0.00	±0.04	±0.06	±0.02	±0.01	±0.00	±0.04	±0.05	±0.01
UID Metrics (↓)												
Loc. non-uni	0.27	0.39	0.18	0.26	0.36	0.36	0.20	**0.23**	0.63	0.63	0.40	0.34
	±0.00	±0.07	±0.04	±0.01	±0.04	±0.03	±0.03	±0.01	±0.00	±0.07	±0.03	±0.01
Loc. uni	**0.53**	0.56	**0.30**	**0.31**	**0.39**	**0.44**	0.24	**0.23**	0.69	0.70	0.48	0.34
	±0.06	±0.02	±0.00	±0.01	±0.04	±0.06	±0.04	±0.01	±0.03	±0.03	±0.04	±0.01
Glob. non-uni	0.52	**0.64**	0.26	0.30	0.37	0.36	0.18	0.21	**0.70**	0.61	0.47	0.33
	±0.08	±0.05	±0.04	±0.01	±0.03	±0.04	±0.04	±0.00	±0.00	±0.06	±0.03	±0.01
Glob. uni	0.33	0.44	0.16	0.28	0.27	0.40	**0.27**	0.23	0.66	0.68	0.41	**0.34**
	±0.00	±0.02	±0.04	±0.00	±0.12	±0.06	±0.03	±0.01	±0.02	±0.07	±0.02	±0.02

4. Further Analyses

Analysis 1: Model Size Analysis

- UID-guided reasoning selection **scales with model capacity**

Methods (↓)	Qwen3-1.7B	Qwen3-4B	Qwen3-8B
Mean Acc.	0.35	0.65	0.67
Self-Cert.	**0.45**	**0.73**	0.63
High Conf.	0.37	0.62	0.60
Low Ent.	0.37	0.63	0.60
UID Metrics (↓)			
Loc. non-uni	0.24	0.54	0.63
Loc. uni	0.41	0.69	0.69
Glob. non-uni	0.37	0.66	**0.70**
Glob. uni	0.33	0.67	0.66

Analysis 2: Domain Generalization

- UID-guided selection yields **robust generalization beyond math domains**

	DS-R1-Distill-Llama-8B			DS-R1-Distill-Qwen-7B		
Methods (↓)	GPQA-D	LSAT-AR	LSAT-LR	GPQA-D	LSAT-AR	LSAT-LR
Mean Acc.	0.48	0.43	0.52	0.48	0.55	0.50
Self-Cert.	0.51	0.44	0.52	**0.51**	0.51	0.49
High Conf.	**0.53**	0.43	0.51	0.46	0.51	0.51
Low Ent.	0.49	0.42	0.52	0.48	0.51	<u>0.52</u>
UID Metrics (↓)						
Loc. non-uni	0.44	0.39	0.49	0.43	0.48	0.46
Loc. uni	<u>0.52</u>	<u>0.46</u>	<u>0.53</u>	**0.51**	**0.62**	0.49
Glob. non-uni	0.47	**0.49**	**0.54**	0.50	**0.62**	**0.53**
Glob. uni	0.44	0.40	0.50	0.46	0.46	0.46

Analysis 3: Sample Size Analysis

- **Stable step-level information flow with trajectory-level variation** yields the highest accuracy, even under small sampling budgets.

Methods (↓)	Sample by 3	Sample by 5	Sample by 10
Mean Acc.	0.67	0.67	0.68
Self-Cert.	0.70	0.63	0.62
High Conf.	0.63	0.60	0.57
Low Ent.	0.63	0.60	0.56
UID Metrics (↓)			
Loc. non-uni	0.63	0.63	0.53
Loc. uni	**0.73**	<u>0.69</u>	<u>0.72</u>
Glob. non-uni	<u>0.70</u>	**0.70**	**0.70**
Glob. uni	0.70	0.66	0.63

5. Why Does UID Differ between Human Communication and LLM Reasoning?

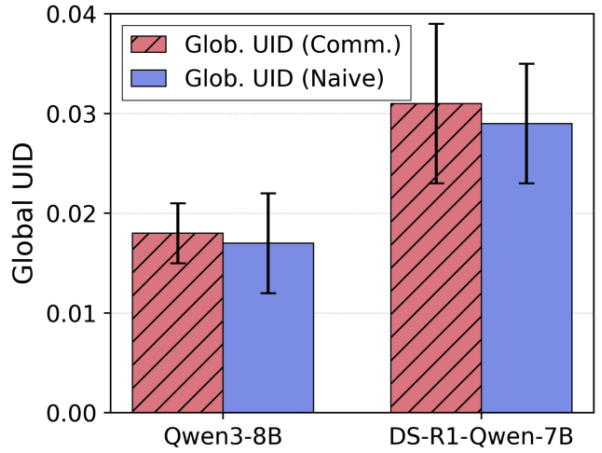
Human Communication vs. LLM Reasoning

- We compare two prompts to test if UID deviations stem from differing objectives

Settings	Content
Communication Prompting	Please answer the following math question as if you are explaining it to a listener. Communicate your reasoning clearly and naturally. You should provide your final answer in the format \boxed{YOUR_ANSWER}. \n\n Question:\n{question}\n
Naive Prompting	Please answer the following math question. You should provide your final answer in the format \boxed{YOUR_ANSWER}. \n\n Question:\n{question}\n

Human Communication vs. LLM Reasoning

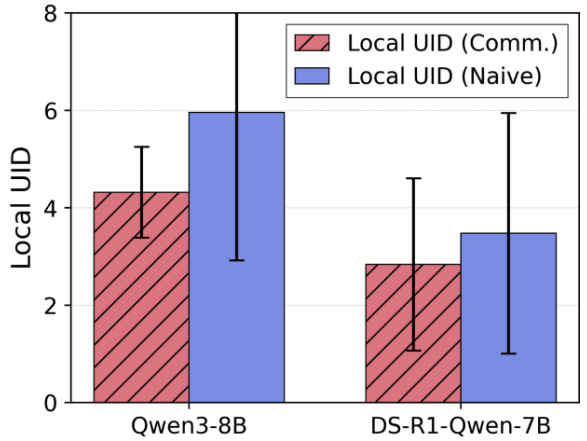
- Human communication is listener-optimized
- LLM reasoning is computation-driven
- Only local smoothness reliably indicates reasoning coherence



(a) Global UID of comm vs. naive prompting

Model	Comm.	Naive
DS-R1-Distill-Qwen-7B	0.37	**0.41**
Qwen3-8B	0.63	**0.67**

Table 5: Performance under comm. and naive settings.



(b) Local UID of comm vs. naive prompting

Conclusion

- Uniform Information Density does not hold globally in LLM reasoning
- High-quality reasoning exhibits local smoothness but global structure
- Information distribution metrics are effective signals of reasoning quality

Thank you

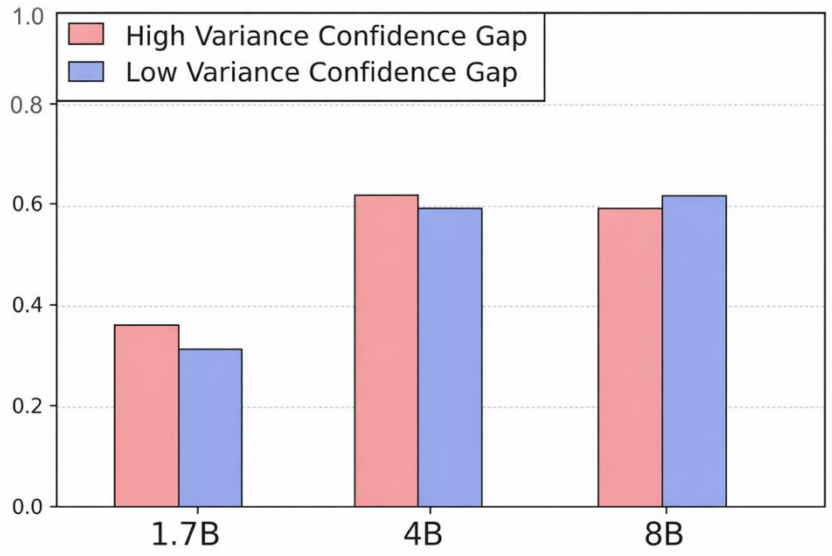
App1. Quantifying Linguistic Uniformity in LLM Reasoning

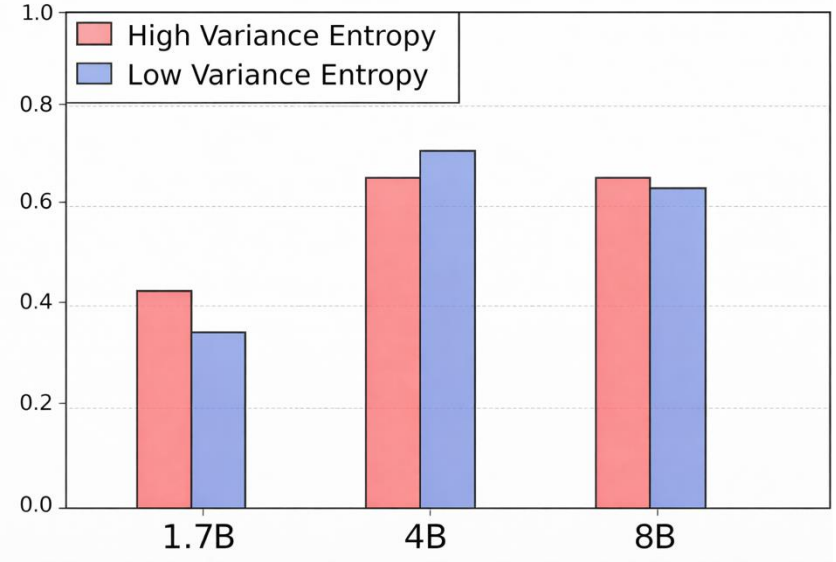
Quantifying Linguistic Uniformity in LLM Reasoning (1)

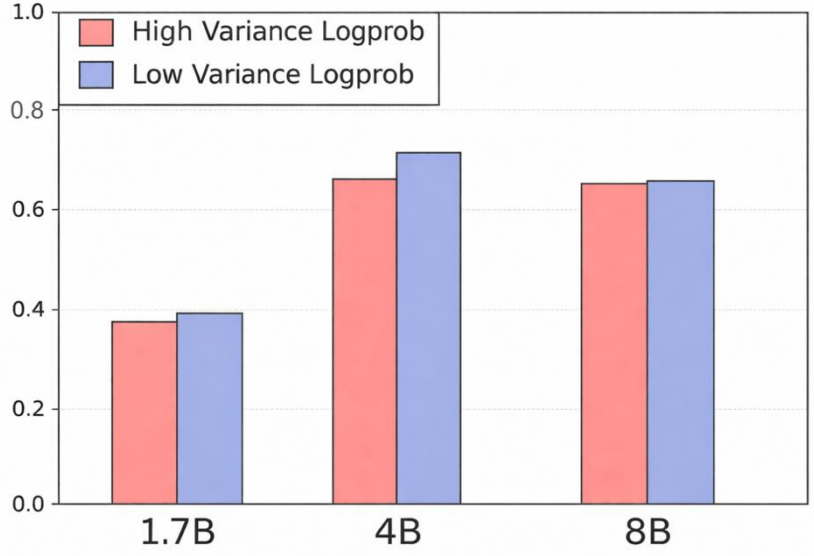
- Entropy as a proxy for information density (Theoretical Justification)
  - reflects both model **confidence and variability** in reasoning
  - quantifies **the expected number of bits** required to encode the predictive distribution

Quantifying Linguistic Uniformity in LLM Reasoning (1)

- Entropy as a proxy for information density (Empirical Justification)
  - entropy measures the difference between highest and lowest trends more profoundly than others







Quantifying Linguistic Uniformity in LLM Reasoning (2)

- Justification for Step-wise Segmentation
  - **well-known heuristic** aligning with how LLMs structure free-form reasoning
  - But, also **robust to other segmentation strategies** as well

Table 6: Fixed window segmentation (2048 tokens)

	Qwen3-8B			DS-R1-Distill-Qwen-7B			DS-R1-Distill-Llama-8B		
Setting	AIME	HMMT	BRUMO	AIME	HMMT	BRUMO	AIME	HMMT	BRUMO
Loc. non-uni (3σ)	**0.77**	0.47	0.70	0.40	0.23	**0.57**	**0.40**	0.14	0.33
Loc. uni (3σ)	**0.77**	**0.47**	**0.70**	0.40	0.23	**0.57**	**0.40**	0.14	0.33
Glob. non-uni (var)	0.67	**0.43**	0.63	0.40	**0.30**	**0.60**	**0.40**	**0.27**	**0.43**
Glob. uni (var)	0.63	0.33	**0.67**	0.27	0.20	0.37	0.27	0.20	0.33

Table 7: Semantic Segmentation

	Qwen3-8B			DS-R1-Distill-Qwen-7B			DS-R1-Distill-Llama-8B		
Setting	AIME	HMMT	BRUMO	AIME	HMMT	BRUMO	AIME	HMMT	BRUMO
Loc. non-uni (3σ)	**0.70**	0.33	0.60	0.37	0.17	0.40	0.20	0.20	**0.43**
Loc. uni (3σ)	0.63	**0.47**	**0.67**	**0.47**	**0.27**	**0.50**	**0.43**	0.23	0.40
Glob. non-uni (var)	0.60	**0.43**	**0.67**	0.27	0.20	**0.53**	0.30	**0.27**	0.40
Glob. uni (var)	0.57	0.33	**0.67**	0.43	**0.27**	0.47	0.37	0.20	0.40

Deep Think with Confidence (ICLR 2025)
L1: Controlling How Long A Reasoning Model Thinks With Reinforcement Learning (COLM 2025)

App2. Qualitative Analyses of Reasoning Traces

Qualitative Analyses on Global UID

- High-variance traces encode informative shifts, whereas low-variance doesn't

<think>\nOkay, so I need to find all positive integers n such that n + 2 divides the product 3(n + 3)(n\u00b2 + 9), and then sum all those n's. Hmm, let me think about how to approach this.

\n\nFirst, let me restate the problem in mathematical terms. We need to find all positive integers n where (n + 2) divides 3(n + 3)(n\u00b2 + 9). In other words, 3(n + 3)(n\u00b2 + 9) divided by (n + 2) should be an integer. So, mathematically, we can write this as:\n\n(n + 2) | 3(n + 3)(n\u00b2 + 9)\n\n Which means that 3(n + 3)(n\u00b2 + 9) mod (n + 2) = 0.

\n\nI remember that when dealing with divisibility conditions, sometimes substituting variables or using modular arithmetic can help. Maybe I can use the fact that if n \u2261 -2 mod (n + 2), then I can substitute n = -2 into the expression to find the remainder. Let me try that.

\n\n Let me consider the expression 3(n + 3)(n\u00b2 + 9) modulo (n + 2). If I substitute n \u2261 -2, then:\n\nFirst, compute each part modulo (n + 2):\n\nn \u2261 -2, so n + 3 \u2261 -2 + 3 = 1.

\n\n Similarly, n\u00b2 + 9 \u2261 (-2)\u00b2 + 9 = 4 + 9 = 13.\n\n Therefore, the entire expression becomes 3 * 1 * 13 = 39. Therefore, the remainder when 3(n + 3)(n\u00b2 + 9) is divided by (n + 2) is 39.

\n\n But since (n + 2) divides the product, this remainder must be zero. Therefore, 39 must be congruent to 0 modulo (n + 2). Which implies that (n + 2) divides 39.

\n\nSo, n + 2 is a positive divisor of 39. Since n is a positive integer, n + 2 must be an integer greater than or equal to 3 (since n \u2265 1, so n + 2 \u2265 3). Therefore, we need to find all positive divisors of 39 that are \u2265 3, then subtract 2 to get n, and then sum all such n's. (. . .)

Figure 17: Reasoning trace with the highest variance score, with rich reasoning. `math_equal` is True

Then:\n\nFrom 3x_{k+1} = x_k + z_k - 1. \n\nBut z_{k+1} = 1/x_{k+1} = 3 / (x_k + z_k - 1). \n\nNot sure if helpful. \n\nAlternatively, maybe consider the recurrence relation for z_k:\n\nIf z_k = 1/x_k, then:\n\n**From 3x_{k+1} = x_k + z_k - 1, so x_{k+1} = (x_k + z_k - 1)/3. Then z_{k+1} = 3 / (x_k + z_k - 1). \n\nBut since z_k = 1/x_k, maybe express in terms of z_k:\n\nLet me write x_k = 1/z_k. Then:\n\nz_{k+1} = 3 / (1/z_k + z_k - 1) = 3 / ((1 + z_k\u00b2 - z_k)/z_k) = 3z_k / (z_k\u00b2 - z_k + 1). \n\nSo z_{k+1} = 3z_k / (z_k\u00b2 - z_k + 1).**

\n\nHmm. Not sure if this helps. But maybe there's a pattern here. Let me check for the first few terms.\n\nGiven that x\u2081 = 25/11, so z\u2081 = 11/25. \n\nThen z\u2082 = 3*(11/25) / [((11/25)^2 - (11/25) + 1)] \n\nLet me compute denominator:\n\n((121/625) - (11/25) + 1) = (121/625 - 275/625 + 625/625) = (121 - 275 + 625)/625 = (471)/625. \n\nTherefore z\u2082 = 3*(11/25) / (471/625) = (33/25) * (625/471) = (33 * 25)/471 = 825 / 471. Simplify: divide numerator and denominator by 3: 275 / 157. Which is the reciprocal of x\u2082. Indeed, x\u2082 = 157/275, so z\u2082 = 275/157. \n\nSo z\u2082 = 275/157. Then z\u2083 = 3*z\u2082 / (z\u2082\u00b2 - z\u2082 + 1). Let me compute that. \n\nFirst compute z\u2082\u00b2: (275/157)^2 = 75625 / 24649. Then subtract z\u2082: 75625 / 24649 - 275/157. Let me convert 275/157 to denominator 24649: 275/157 = (275 * 157)/24649. Wait, 157*157 = 24649, so 275/157 = (275*157)/24649. Wait, but that would be 275*157. Let me compute that:\n\n275 * 157. Let me compute 275*(150 + 7) = 275*150 + 275*7 = 41250 + 1925 = 43175. Therefore, 275/157 = 43175 / 24649. \n\nSo z\u2082\u00b2 - z\u2082 + 1 = (75625 - 43175 + 24649)/24649. Let me compute numerator: 75625 - 43175 = 32450; 32450 + 24649 = 57099. So denominator is 24649. Therefore z\u2082\u00b2 - z\u2082 + 1 = 57099 / 24649. \n\nTherefore z\u2083 = 3*z\u2082 / (57099 / 24649)) = 3*(275/157) * (24649 / 57099). Let me compute this. \n\nFirst, note that 24649 is 157\u00b2. And 57099 was from previous steps. Wait, earlier when computing x\u2083, we had 57099 / 129525, but maybe not. Anyway, let me compute:\n\nz\u2083 = 3 * (275/157) * (24649 / 57099). \n\nNote that 24649 = 157\u00b2, so:\n\nz\u2083 = 3 * 275 * 157\u00b2 / (157 * 57099)) = 3 * 275 * 157 / 57099. \n\nSimplify numerator and denominator:\n\nLet me compute 3*275 = 825. So numerator is 825 * 157. Let me compute that:\n\n825 * 157. Let me do 800*157 + 25*157 = 125600 + 3925 = 129525. \n\nSo numerator is 129525. Denominator is 57099. Therefore z\u2083 = 129525 / 57099. Simplify this fraction. Let me check GCD(129525, 57099). \n\nUsing Euclidean algorithm:\n\nGCD(129525, 57099). \n\nDivide 129525 by 57099: 2 times, remainder 129525 - 2*57099 = 129525 - 114198 = 15327.\n\nGCD(57099, 15327). **\n\nAs before, earlier steps showed that GCD(57099, 15327) was 3**. Wait, earlier when we had GCD(57099, 129525) was 3. Let me check again:\n\nWait, GCD(129525, 57099) = GCD(57099, 15327). **Then GCD(57099, 15327). \n\nEarlier steps showed that GCD(57099, 15327) was 3**. So divide numerator and denominator by 3:\n\n129525 / 3 = 43175, 57099 / 3 = 19033. Therefore z\u2083 = 43175 / 19033. Which is the reciprocal of x\u2083. Indeed, x\u2083 = 19033/43175, so z\u2083 = 43175/19033. \n\nSo this seems like the pattern is that z_{k} is the reciprocal of x_k.

Figure 18: Reasoning trace with the lowest variance score, showing that the model is stuck. `math_equal` is False

Qualitative Analyses on Local UID

- Locally variant traces encode non-informative shifts, whereas the opposite doesn't

n\n**Another possibility**: b=24, d=-54. 24*(-54)=-1296.\n\na + c=12.\n\nad + bc= a*(-54) + c*24= -54a +24c= -432.\n\nDivide by 6: -9a +4c= -72.\n\nFrom a + c=12 \u2192 c=12 -a.\n\nSubstitute into -9a +4(12 -a) =-72 \u2192 -9a +48 -4a =-72 \u2192 -13a= -120 \u2192 a=120/13. Not integer. **Not good.\n\nAnother possibility**: b= -36, d=36. Then b*d= -36*36= -1296.\n\na + c=12.\n\nad + bc= a*36 + c*(-36)=36a -36c= -432 \u2192 36(a - c)= -432 \u2192 a - c= -12.\n\nAlso, a + c=12.\n\nSolving:\n\na + c=12.\n\na - c= -12.\n\nAdding: 2a=0 \u2192 a=0, c=12.\n\nNow, check ac + b + d=0*12 + (-36) +36=0, which is not 144. So, no.\n\nAnother possibility: b= 144, d= -9. 144*(-9)= -1296.\n\na + c=12.\n\nad + bc= a*(-9) + c*144= -9a +144c= -432.\n\nDivide by 9: -a +16c= -48.\n\nAlso, a + c=12.\n\nSo, a=12 -c.\n\nSubstitute into -a +16c= -48: -(12 -c) +16c= -48 \u2192 -12 +c +16c= -48 \u2192 17c= -36 \u2192 c= -36/17. Not integer.\n\nAnother try: b= 54, d= -24. 54*(-24)= -1296.\n\na + c=12.\n\nad + bc= a*(-24) + c*54= -24a +54c= -432 \u2192 divide by 6: -4a +9c= -72.\n\nFrom a + c=12 \u2192 a=12 -c.\n\nSubstitute into -4(12 -c) +9c= -72 \u2192 -48 +4c +9c= -72 \u219213c= -24 \u2192 c= -24/13. **Not integer.\n\nHmm, perhaps I should try** b= 48, d= -27. 48*(-27)= -1296.\n\na + c=12.\n\nad + bc= a*(-27) + c*48= -27a +48c= -432 \u2192 divide by 3: -9a +16c= -144.\n\nAlso, a=12 -c.\n\nSubstitute: -9(12 -c) +16c= -144 \u2192 -108 +9c +16c= -144 \u219225c= -36 \u2192c= -36/25. **Not integer.\n\nThis is getting too time-consuming. Maybe I should try a different approach.\n\nAlternatively, perhaps there is a substitution I can make**.\n\nWe have from equation 1: (h -6)^2 +k^2=36.\n\nAnd from equation 2: (h +6)(k +6)=72.\n\nLet me let u = h +6 and v =k +6.\n\nThen, equation 2 becomes u*v=72.\n\nEquation 1 becomes (h -6)^2 +k^2=36.\n\nExpress h and k in terms of u and v:\n\nh = u -6, k= v -6.\n\nSo, equation 1 becomes:\n\n(u -6 -6)^2 + (v -6)^2 =36 \u2192 (u -12)^2 + (v -6)^2=36.\n\nSo, equation 1: (u -12)^2 + (v -6)^2=36.\n\nAnd equation 2: u*v=72.\n\nSo, we have:\n\n(u -12)^2 + (v -6)^2=36.\n\nAnd u*v=72.\n\nWe can expand equation 1:\n\nu\u00b2 -24u +144 +v\u00b2 -12v +36=36 \u2192 u\u00b2 + v\u00b2 -24u -12v +180=36 \u2192 u\u00b2 +v\u00b2 -24u -12v= -144.\n\nBut since u*v=72, perhaps we can find u and v such that u*v=72 and u\u00b2 +v\u00b2=24u +12v -144.\n\nAlternatively, we can write u\u00b2 +v\u00b2 =24u +12v -144.\n\nBut from u*v=72, perhaps express v=72/u.\n\nSubstitute into u\u00b2 + (72/u)^2 =24u +12*(72/u) -144.\n\nMultiply both sides by u\u00b2 to eliminate denominators:\n\nu^4 + 5184 =24u^3 +12*72 u -144u\u00b2.\n\nSimplify:\n\nu^4 +5184 =24u^3 +864u -144u\u00b2.\n\nBring all terms to one side:\n\nu^4 -24u^3 +144u\u00b2 -864u +5184=0.\n\nHmm, this seems complicated, but perhaps we can factor it.\n\nLet me try possible rational roots. Possible roots are factors of 5184 over 1, so possible roots include \u00b11, \u00b12, ..., up to \u00b15184.\n\nTesting u=12:\n\n12^4=20736, -24*12^3= -24*1728= -41472, 144*12^2=144*144=20736, -864*12= -10368, +5184.\n\nSo, 20736 -41472 +20736 -10368 +5184.\n\nCompute step by step:\n\n20736 -41472= -20736.\n\n-20736 +20736=0.\n\n0 -10368= -10368.\n\n-10368 +5184= -5184 \u22600.\n\nu=6:\n\n6^4=1296, -24*6^3= -24*216= -5184, 144*36=5184, -864*6= -5184, +5184.\n\nSo, 1296 -5184 +5184 -5184 +5184.\n\nCompute:\n\n1296 -5184= -3888.\n\n-3888 +5184=1296.\n\n1296 -5184= -3888.\n\n-3888 +5184=1296 \u22600.\n\nu=-6:\n\n(-6)^4=1296, -24*(-6)^3= -24*(-216)=5184, 144*36=5184, -864*(-6)=5184, +5184.\n\nSo, 1296 +5184 +5184 +5184 +5184.\n\nWait, wait, no. Wait, the equation is u^4 -24u^3 +144u^2 -864u +5184=0.\n\nPlugging u=-6:\n\n(-6)^4=1296.\n\n-24*(-6)^3= -24*(-216)=5184.\n\n144*(-6)^2=144*36=5184.\n\n-864*(-6)=5184.\n\n+5184.\n\nSo, sum: 1296 +5184=6480, +5184=11664, +5184=16848, +5184=22032. Not zero.\n\nu=8:\n\n8^4=4096, -24*512= -12288, 144*64=9216, -864*8= -6912, +5184.\n\nCompute:\n\n4096 -12288= -8192.\n\n-8192 +9216=1024.\n\n1024 -6912= -5888.\n\n-5888 +5184= -704 \u22600.\n\nu=9:\n\n9^4=6561, -24*729= -17496, 144*81=11664, -864*9= -7776, +5184.\n\nCompute:\n\n6561 -17496= -10935.

Figure 19: Reasoning trace with the highest number of spikes and falls, with a lot of redundancy and confusions. `math_equal` is True

Okay, so I've got this math problem to solve. It says: Find the sum of all integer bases \\(b > 9 \\) for which \\(17_b \\) is a divisor of \\(97_b \\). Hmm, let me try to figure this out step by step.\n\nFirst, I remember that when numbers are written in a base \\(b \\), each digit represents a power of \\(b \\). So, for example, \\(17_b \\) would be equal to \\(1 \\times b^1 + 7 \\times b^0 \\) in decimal, which is \\(b + 7 \\). Similarly, \\(97_b \\) would be \\(9 \\times b^1 + 7 \\times b^0 \\), so that's \\(9b + 7 \\) in decimal.\n\nThe problem is asking for bases \\(b > 9 \\) where \\(17_b \\) divides \\(97_b \\). That means \\(b + 7 \\) should be a divisor of \\(9b + 7 \\). So, mathematically, I can write this as:\n\n\\(b + 7 \\) divides \\(9b + 7 \\).\n\nNow, when a number \\(a \\) divides another number \\(c \\), it means that \\(c \\) divided by \\(a \\) leaves no remainder. Another way to put that is \\(c \\mod a = 0 \\). So, I can write:\n\n\\(9b + 7 \\mod (b + 7) = 0 \\).\n\nAlternatively, \\(9b + 7 \\) divided by \\(b + 7 \\) should result in an integer. Maybe I can perform the division to see what that gives.\n\nLet me try dividing \\(9b + 7 \\) by \\(b + 7 \\) using polynomial long division or substitution. Let's see, if I think of it as a division:\n\n\\((9b + 7) \\div (b + 7) \\).\n\nLet me set it up:\n\nHow many times does \\(b + 7 \\) go into \\(9b + 7 \\)? Well, \\(b \\) goes into \\(9b \\) exactly 9 times. So, I multiply \\(b + 7 \\) by 9, which gives \\(9b + 63 \\), and subtract that from \\(9b + 7 \\).\n\nSo, subtracting: \\((9b +7) - (9b + 63) = 0b + (7 - 63) = -56 \\).\n\nSo, the division gives 9 with a remainder of -56.

Figure 20: Reasoning trace with the lowest number of spikes and falls, with simple and concise reasoning. `math_equal` is True