NEURAL INFORMATION PROCESSING SYSTEMS

NeurIPS 2026

FUJITSU

The Butterfly Effect in Reasoning

Branch-Structured Distributional Memory for Stochastic Agents

Ishan Jindal*, Vivek Dhamale*, Mahesh Chandran

Fujitsu Research of India

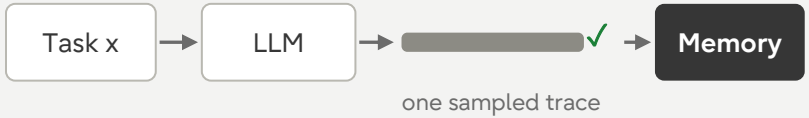
*Equal contribution

Problem: memory from a single trajectory

FUJITSU

Under stochastic decoding, one chain of thought is just one sample of the reasoning.

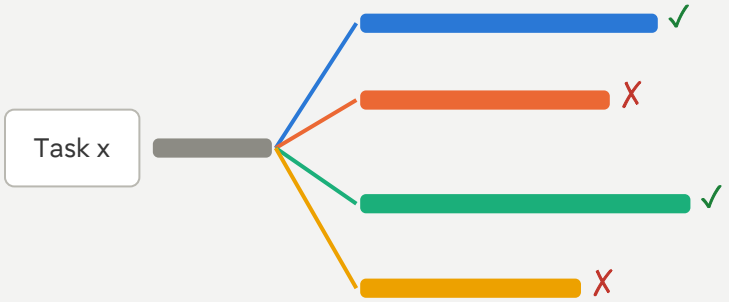
Today: trajectory as experience



Reasoning banks, experience libraries and metacognitive handbooks store past reasoning so agents improve over time.

The unit of experience is **one realized trajectory**.

Reality: reasoning branches early



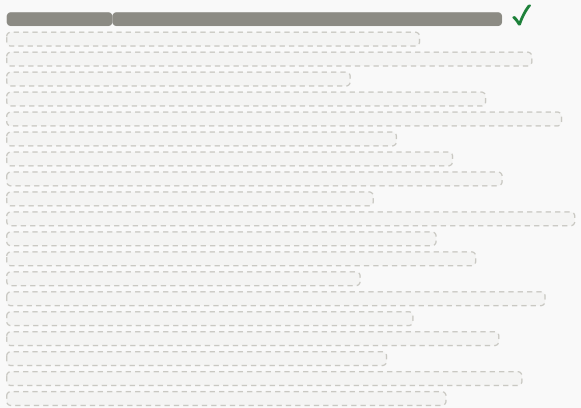
Small differences in the first tokens steer generation to different decompositions, assumptions, answers and failure modes: the **butterfly effect**.

Memory from one trace overfits to that path (its phrasing, exploration order and local mistakes) and misses alternative strategies and recurring failure modes. Sampling more full traces is redundant and saturates the memory generator's context.

Idea: trajectory distribution as experience

Single-trajectory memory

One sampled trace becomes the stored experience.

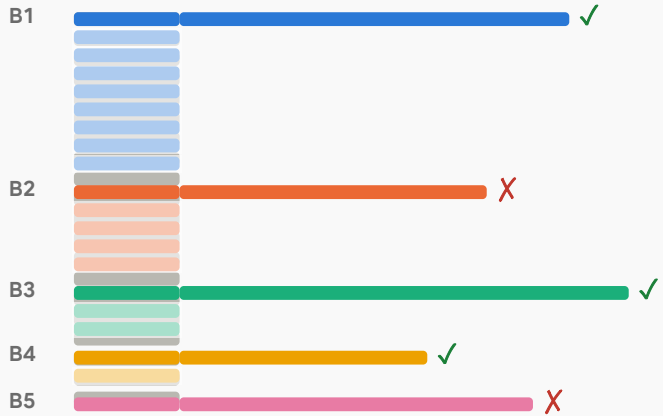


Branches covered 1 / 5

✓ Positive strategies | ✗ Anti-patterns | ⇄ Contrastive rules

BranchDIME

Prefixes → branches → one representative each → memory



Branches covered 5 / 5

✓ Positive strategies | ✗ Anti-patterns | ⇄ Contrastive rules

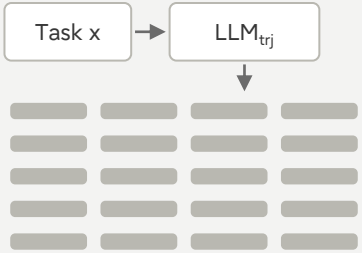
1 Sample prefixes | **2 Discover branches** | **3 Expand representatives** | **4 Distill memory**

Each bar is one sampled reasoning trace; its length is the tokens generated. Dashed bars are samples the single-trajectory method never observes. Branch sizes follow the average DSL-8B branch mass (9 / 5 / 3 / 2 / 1 of 20 prefixes). Schematic, not to scale.

BranchDIME: discover branches, then expand

FUJITSU

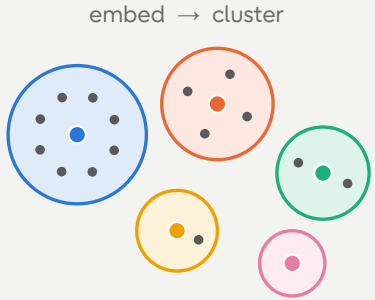
1 Sample short prefixes



Draw n short reasoning prefixes from LLM$_{trj}$. No full trajectories are generated yet.

n = 20 · k = 256 tokens

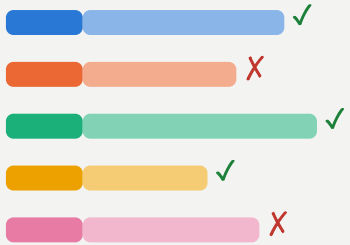
2 Discover branches



Embed prefixes as unit-norm TF-IDF vectors; cluster with spherical K-means (or agglomerative for adaptive B).

B = 5 branches

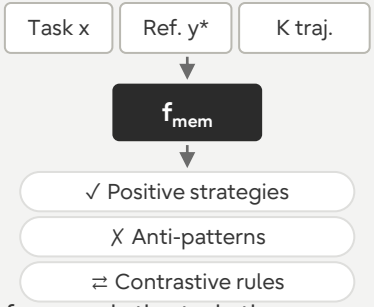
3 Expand representatives



Continue the prefix closest to each centroid to a full trajectory by prefix conditioning.

K = 5 trajectories

4 Distill memory



f$_{mem}$ reads the task, the reference answer and all K trajectories, and writes reusable memories.

f$_{mem}$ = GPT-5-Mini

Every multi-trajectory method gets the same K = 5 full trajectories for the memory generator; they differ only in **which** trajectories are selected. Prefix sampling explores more of the distribution before that budget is spent.

Three kinds of memory, used at inference

FUJITSU

✓ Positive strategies

Reasoning patterns associated with correct outcomes.

e.g. Given advancing and receding contact angles, estimate the static angle between them.

✗ Anti-patterns

Recurring failure modes in incorrect reasoning.

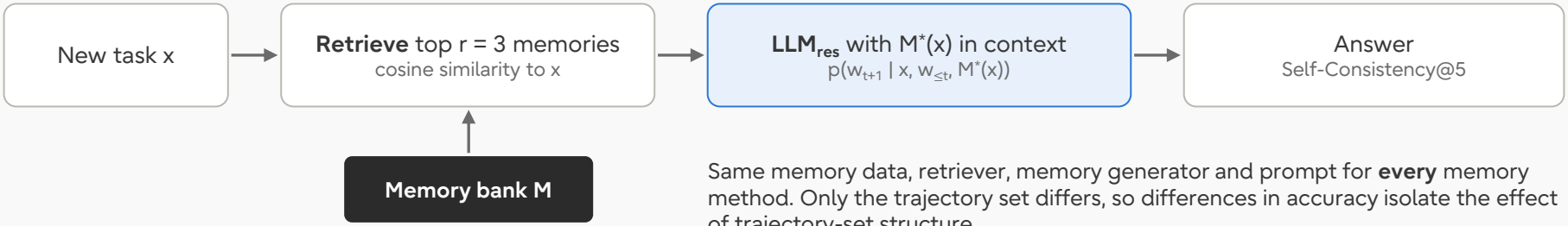
e.g. Don't transfer the wetting state inferred for water directly to oil.

⇄ Contrastive rules

Structural differences between successful and failing trajectories.

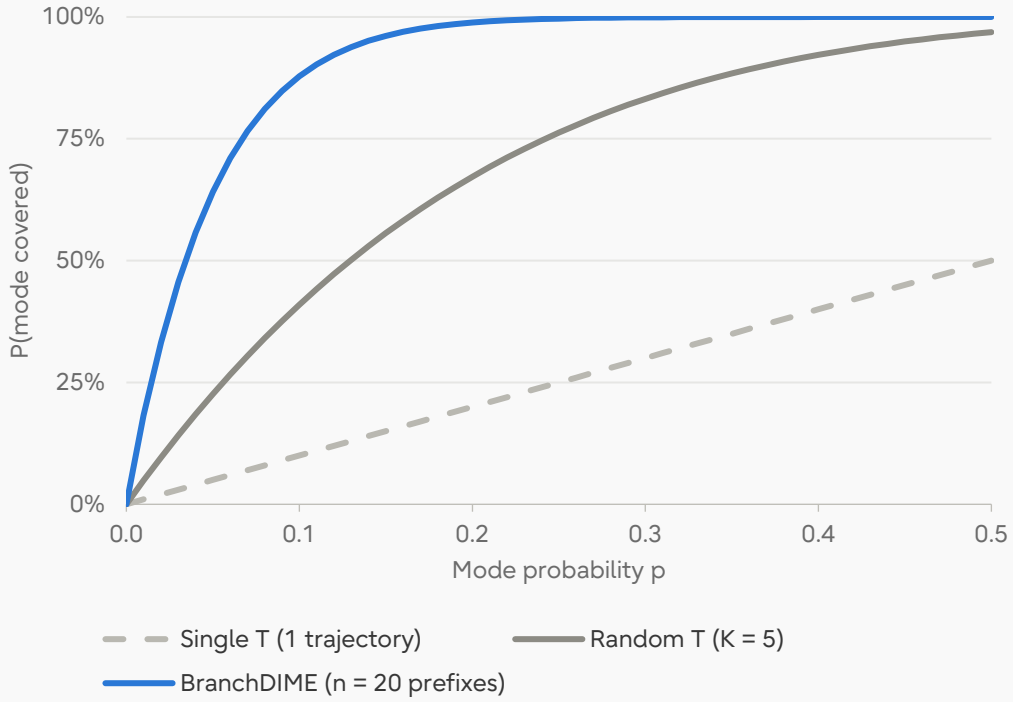
e.g. Check the sign of cos θ: roughness raises angles above 90° and lowers those below.

At inference



Same memory data, retriever, memory generator and prompt for **every** memory method. Only the trajectory set differs, so differences in accuracy isolate the effect of trajectory-set structure.

Why branch first? Rare modes get covered



Proposition 1 (mode coverage)

A mode followed by a fraction p of samples is covered with probability

Random K: $1 - (1 - p)^K$

BranchDIME: $1 - (1 - p)^n$

At p = 5% (a minority mode):

Single T	**5.0%**
Random T (K = 5)	**22.6%**
BranchDIME (n = 20)	**64.2%**

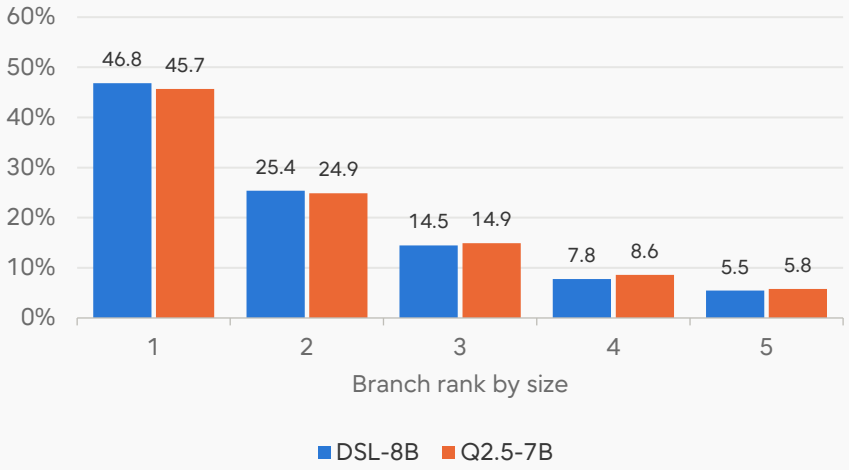
Rare modes: $1 - (1 - p)^n \approx np$
$\rightarrow \approx n / K =$ **4×** coverage advantage at n = 20, K = 5

Idealized: assumes prefix clustering recovers every mode present in the prefix sample. Extra cost over Random T ≈ (n − K)·k prefix tokens.

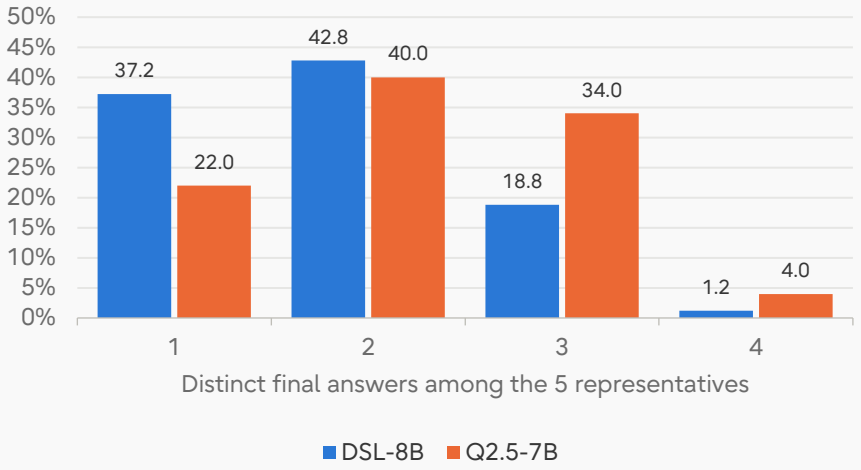
Branches carry mass and change answers

On GPQA-M, branches are skewed but not collapsed, and representatives often disagree.

Branch mass: % of the 20 prefixes in each branch



Outcome diversity: % of questions by # distinct answers



The largest branch holds **< 50%** of prefixes (46.8% DSL-8B, 45.7% Q2.5-7B).
Keeping only the consensus branch discards more than half the distribution.

Representatives reach **1.84 / 2.20** distinct answers on average.
62.8% / 78.0% of questions get at least two; 20.0% / 38.0% get three or more.

Main results: better memory, same latency

	GPQA-M → GPQA-Diamond					MATH-M → MATH500				
Memory	DSQ-7B	Q2.5-7B	DSL-8B	Avg	η (s)	DSQ-7B	Q2.5-7B	DSL-8B	Avg	η (s)
No memory (CoT)	0.491	0.333	0.406	0.410	8.68	0.808	0.728	0.831	0.789	6.55
No memory (SC@5)	0.525	0.343	0.470	0.446	21.66	0.865	0.756	0.866	0.829	11.22
Single T	0.533	0.349	0.505	0.462	23.20	0.882	0.764	0.890	0.845	17.34
Random T	0.535	0.342	0.512	0.463	23.60	0.890	0.767	0.886	0.848	17.52
Consensus T	0.530	0.347	0.485	0.454	95.60	0.891	0.754	0.889	0.845	71.04
BranchDIME	**0.581**	**0.392**	**0.534**	**0.502**	**23.84**	**0.896**	**0.795**	**0.902**	**0.864**	**17.91**

All memory methods: same memory data, retriever, generator (GPT-5-Mini) and prompt; Self-Consistency@5 inference; mean of 3 seeds. η = retrieval + response time, averaged over models.

GPQA-Diamond

+8.7%

relative over Single T
0.462 → 0.502 (avg)

MATH500

+2.2%

relative over Single T
0.845 → 0.864 (avg)

Latency

≈ same

23.8 s vs 23.2 s (GPQA)
17.9 s vs 17.3 s (MATH)

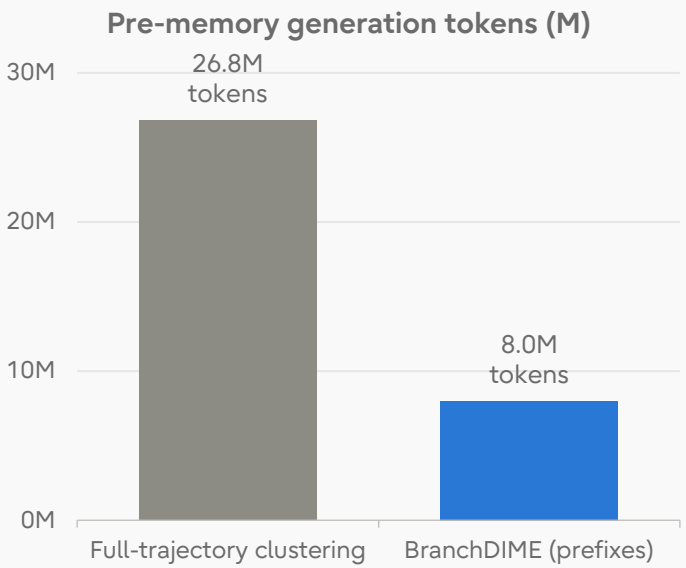
More samples alone

Not enough

Random T ≈ Single T
Consensus T: no better, 4× latency

Prefix clustering at 30% of the tokens

Pre-memory generation tokens (GPQA, DSL-8B)



Full-trajectory clustering

0.539

accuracy · generate all n trajectories, then cluster

BranchDIME

0.534

accuracy · cluster prefixes, expand K representatives

−70.2% tokens for **−0.005** accuracy

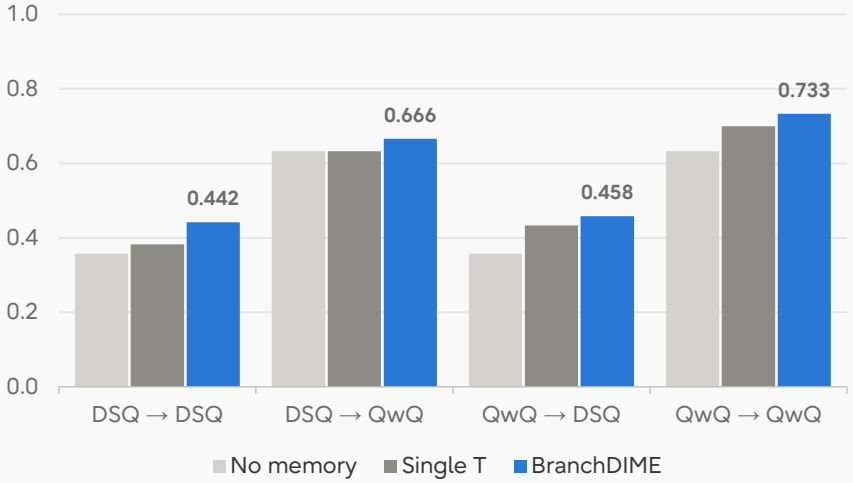
Offline memory-construction cost (one-time, amortized over all queries)

Method	Tokens generated	Note
Single T	t_1	one full trajectory
Random T	$\Sigma_{i=1..K}\, t_i$	K full trajectories
Consensus T	$\Sigma_{i=1..n}\, t_i$	all n finished first
BranchDIME	$\mathbf{n{\cdot}k + \Sigma_{i=1..K}\,(t_i - k)}$	**≈ Random + (n − K)·k**

Memories transfer across tasks and models

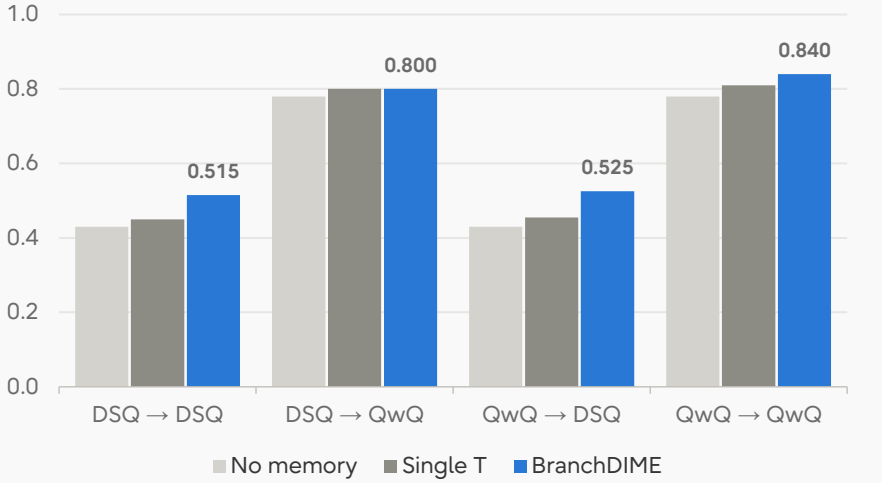
Memory built from MATH-M, tested on the disjoint AIME25 and OlymMATH benchmarks.

AIME25



OlymMATH

x-axis: $LLM_{trj} \rightarrow LLM_{res}$ · DSQ = DSQ-7B, QwQ = QwQ-32B



+15.4% relative on AIME25

0.383 → 0.442 with DSQ-7B building and using memory

Strong → weak transfer works

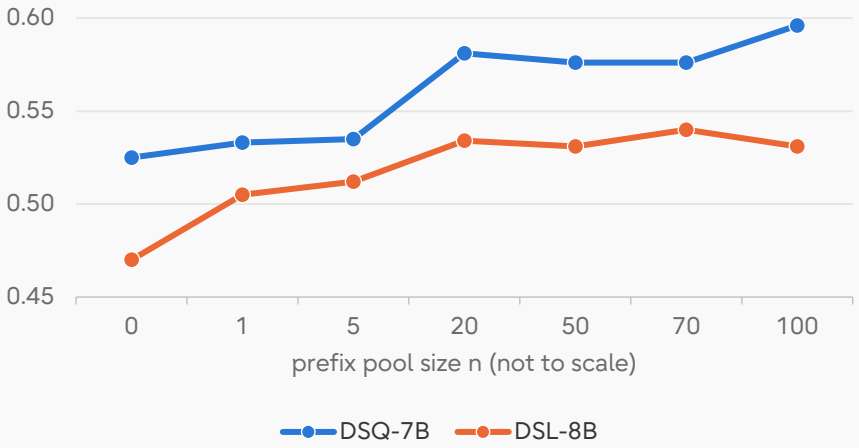
QwQ-32B memories lift DSQ-7B: 0.455 → 0.525 on OlymMATH

Smaller gains for QwQ-32B

Stronger responders recover some structure themselves

Ablations: prefix pool and memory types

Prefix pool size n (K = 5 fixed, GPQA-Diamond)



n = 0 is no memory and n = 1 is Single T.

The big jump is **n = 5 → 20**, where clustering can finally be selective; gains plateau after. n = 20 is the cost-effective default.

Memory types: one type vs. all three (Table 9)

Benchmark	Model	Pos.	Neg.	Contr.	All
GPQA-D	DSQ-7B	0.529	0.540	0.544	**0.581**
GPQA-D	Q2.5-7B	0.355	0.372	0.391	**0.392**
GPQA-D	DSL-8B	0.510	0.505	0.517	**0.534**
MATH500	DSQ-7B	0.884	0.872	0.884	**0.896**
MATH500	Q2.5-7B	0.763	0.751	0.733	**0.795**
MATH500	DSL-8B	0.888	0.880	0.890	**0.902**

Using all three types wins in **6 / 6** settings, by +0.1 to +3.7 points over the best single type.

Contrastive is the strongest single type on GPQA-Diamond; no single type is consistently enough.

Ablations: generator, clustering, features

Memory generator

Table 6 · GPQA-Diamond

Generator	Memory	DSQ	DSL
GPT-5-Mini	Single T	0.533	0.505
	BranchDIME	**0.581**	**0.534**
Qwen3-30B	Single T	0.527	0.502
	BranchDIME	**0.539**	**0.520**
Llama-70B	Single T	0.515	0.515
	BranchDIME	**0.530**	**0.520**

Open-weight generators shrink the gain, but BranchDIME beats Single T with every generator.

Adaptive clustering

Table 7 · GPQA-M, DSL-8B

Method	Reps	Avg B	Acc.
Single T	250	1.00	0.505
Fixed B = 5	**1250**	**5.00**	**0.534**
Agglo. $\delta = 0.20$	**1125**	**4.50**	**0.532**
Agglo. $\delta = 0.30$	**956**	**3.82**	**0.523**

Agglomerative clustering picks B per task: −10% expansions for −0.002 accuracy; −23.5% still beats Single T.

Prefix representation

Table 8 · GPQA-Diamond, SC@5

Model	No mem.	TF-IDF	Jina
DSQ-7B	0.525	0.581	0.574
DSL-8B	0.470	0.534	0.539
Average	0.498	0.558	0.557

Neural embeddings give no consistent gain over TF-IDF (0.1 pt), so the lightweight default stays.

Limitations and future work

FUJITSU

Limitations

- **Offline construction stage**: prefix sampling, clustering and expansion. Paid once and amortized, but adds cost for very large memory banks.
- **Surface-level branch discovery**: TF-IDF may miss branches that differ in substance but not in wording.
- **Construction hyperparameters**: prefix length k, pool size n, branches B and budget K.
- **Generator dependence**: the size of the gain varies with the memory generator.
- **Scope**: math and STEM reasoning; not yet long-horizon tool use or interactive agents.

Future work

- Adaptive prefix budgets, early stopping and dynamic branch allocation.
- Learned similarity functions and task-adaptive clustering for branch discovery.
- Choosing n, B and K automatically from the observed prefix distribution.
- Better distillation: verification and uncertainty-aware memory filtering.
- Tool use, interactive environments, multilingual tasks, delayed feedback.

Takeaways

FUJITSU

From **trajectory as experience** → **trajectory distribution as experience**

GPQA-Diamond	MATH500	Branch coverage	Construction	Transfer
+8.7%	**+2.2%**	**100%**	**−70%**	**+15.4%**
relative accuracy vs. single-trajectory memory	relative accuracy vs. single-trajectory memory	vs. 64.9% random, 20.0% consensus	tokens vs. full-trajectory clustering, within 0.005 accuracy	relative, MATH-M memory on AIME25 / OlymMATH

- Under stochastic decoding, a single realized trajectory is an **incomplete unit of experience**.
- BranchDIME samples cheap prefixes, discovers branches, expands one representative per branch and distills **positive, anti-pattern and contrastive** memories, all at a fixed trajectory budget.
- Multiple trajectories are not redundant samples to aggregate: they are observations of a reasoning structure that can be **discovered, compressed and reused**.

NEURAL INFORMATION PROCESSING SYSTEMS

NeurIPS 2026

FUJITSU

Thank you

The Butterfly Effect in Reasoning: Branch-Structured Distributional Memory for Stochastic Agents

Ishan Jindal, Vivek Dhamale, Mahesh Chandran

Fujitsu Research of India