Cross-Model Cognitive Task Combiner: Ensemble Architectures for Heterogeneous LLM Outputs

ForceDream Research Team, Cognitive Architecture Division | 2026-01-30 | v1.1 | 3 pages
Category: Inference | Layers: L1, L4, L3
URL: https://forcedream.com/research/cross-model-cognitive-task-combiner-llm-ensemble-architecture

WORM ACCESS SEAL | L828
fd2026007a1e9b3c

Abstract

CMCTC achieves 9.4% hallucination reduction and 7.2% factual grounding improvement across 12,000 diverse tasks by aggregating outputs from multiple heterogeneous LLM providers using confidence-weighted voting calibrated to task type. Mean additional latency: 34ms for soft-weighted synthesis mode.

1. Motivation

Single-model inference pipelines exhibit systematic failure modes: hallucination rates of 3-8% on factual tasks, distributional shift under prompt variation, and task-type miscalibration. Scaling alone does not eliminate these failure modes. CMCTC addresses them through ensemble aggregation.

2. Confidence-Weighted Voting

For each task, CMCTC queries k providers (typically k=3) and assigns each output a confidence score using a lightweight calibration model. The final output is produced by a confidence-weighted voting function optimising over discrete or continuous output spaces.

3. Three Operating Modes

Hard-majority: highest weighted vote count selected, used for classification. Soft-weighted: confidence-weighted combination, used for generation. Adversarial: contradiction detection against WORM-sealed Atlas records, triggering re-query with explicit contradiction context, used for factual recall.

4. Evaluation

Across 12,000 diverse tasks: hallucination rate 9.4% reduction ($p<0.001$), factual grounding 7.2% improvement vs Atlas ground truth, task success rate 4.1% improvement, additional latency (soft-weighted) 34ms, throughput impact 2.1% reduction.

5. Conclusions

CMCTC demonstrates reliable quality improvements across diverse task types. The adversarial mode provides grounding in WORM-sealed facts, critical for compliance and fraud detection where hallucination is a regulatory risk.

Live API Endpoints

```
POST /v1/inference/race
POST /v1/inference/route
GET /v1/models/list
POST /v1/inference/ensemble
```

Citation

ForceDream Research Team (2026). Cross-Model Cognitive Task Combiner. ForceDream Intelligence OS Research Series, FD-2026-007. https://forcedream.com/research/cross-model-cognitive-task-combiner-llm-ensemble-architecture

---