Multi-Engineer Part Families (`part_family_id`) & Zero-Contamination Holdout Architecture for 3D CAD Direct Preference Optimization

SpacetimeML Embodied & Spatial Systems Group
SpacetimeML Research, San Francisco, CA
`research@spacetimeml.com` · `https://spacetimeml.com`

October 2026 · Technical Report `STM-TR-2026-03`

Abstract

Training long-horizon 3D Computer-Use Agents (CUAs) via supervised behavior cloning alone is suboptimal when human demonstrations vary widely in action efficiency, UI navigation paths, and error-recovery subtrajectories. We introduce the **Multi-Engineer Part Family (`part_family_id`)** architecture across SpacetimeML's 105-session industrial CAD corpus (72.12 hours, 102,821 frames, 502,556 OS GUI events, 0 duplicate deliverable SHA-256 hashes). By routing identical ISO-128 2D engineering blueprints to multiple independent mechanical engineers operating *SolidWorks*, *PTC Creo*, *AutoCAD*, and *FreeCAD*, we obtain 29 multi-trajectory part families where distinct human trajectories $\tau_1, \ldots, \tau_k$ converge to verified ISO-10303-21 `.STEP` solids ($R = 1.000$) while exhibiting up to a $4.9\times$ spread in completion horizon. We formulate **Horizon-Weighted CAD Direct Preference Optimization ($\mathcal{L}_{\text{CAD-DPO}}$)** and cross-application Group Relative Policy Optimization (GRPO), and detail our cryptographic SHA-256 deduplication and air-gapped enterprise holdout protocol.

1 Action-Horizon Variance in 3D CUA

In desktop 3D CAD modeling, two mechanical engineers presented with the same ISO-128 reference blueprint B_f (part family f) may both produce geometrically identical, watertight B-Rep solids ($R(M(\tau), M_{\text{gt}}) = 1.000$), yet follow substantially different GUI trajectories:

- **Expert Minimum-Feature Trajectory (τ_w):** Selects the canonical datum plane, applies symmetric midpoint sketch relations, extrudes the primary boss, and patterns the counterbore array in $T_w = 10.4\,\text{min}$ ($|\tau_w| = 1,840$ GUI actions).
- **Exploratory Trial-and-Error Trajectory (τ_l):** Extrudes individual lobes sequentially, triggers an over-defined sketch rebuild warning, deletes and recreates a fillet loop, and completes the part in $T_l = 46.0\,\text{min}$ ($|\tau_l| = 8,120$ GUI actions).

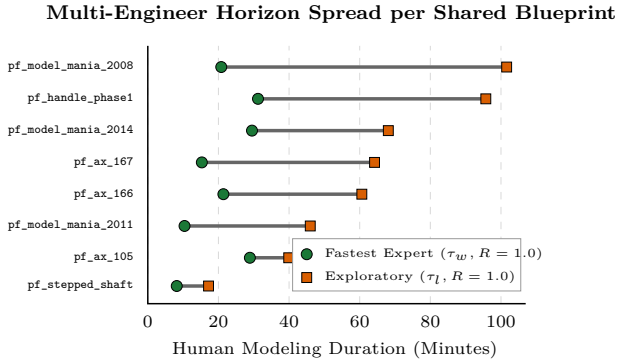


Figure 1: Observed human modeling duration spread (τ_w vs. τ_l) across eight representative multi-engineer part families.

Uniform behavior cloning over $\{\tau_w, \tau_l\}$ assigns equal probability mass to exploratory UI thrashing. Pairing (τ_w, τ_l) under a shared `part_family_id` yields a direct preference signal over multi-thousand-step GUI trajectories (Figure 1).

2 Horizon-Weighted CAD-DPO Objective

For a blueprint B_f in part family f, let $\tau = \{(o_t, a_t)\}_{t=1}^{T}$ denote a trajectory producing final exported `.STEP` solid $M(\tau)$. We define the trajectory utility $U(\tau; B_f)$ by combining the deterministic B-Rep reward $R(M(\tau), M_{\text{gt}})$ from `verify_step.py` with an action-horizon discount:

$$U(\tau; B_f) = R(M(\tau), M_{\text{gt}}) \exp\left(-\alpha \frac{\max(0, |\tau| - |\tau_f^*|)}{|\tau_f^*|}\right), \tag{1}$$

where $|\tau_f^*| = \min_{\tau' \in \mathcal{T}_f : R(\tau') \geq 0.85} |\tau'|$ is the shortest verified human action horizon for part family f, and $\alpha > 0$ is a horizon regularization weight.

Table 1: Representative Multi-Engineer Part Families (`part_family_id`) in the 105-Session SpacetimeML Corpus.

Part Family ID	Spec	Mechanical Component Description	Eng.	Durations (min)	Spread	Frames / Events
`pf_model_mania_2011`	`AX-145 / 018`	Model Mania 2011 Ribbed Clevis & Boss	4×	$10.4, 13.6, 25.2, 46.0$	4.4×	$2,309/15,316$
`pf_ax_167`	`AX-167 / 174`	Multi-Port Hydraulic Valve Manifold	4×	$15.3, 18.7, 32.6, 64.2$	4.2×	$3,748/5,909$
`pf_model_mania_2008`	`AX-023 / 119`	Model Mania 2008 Swept Bellcrank Yoke	3×	$20.8, 68.1, 101.6$	4.9×	$3,815/42,647$
`pf_handle_phase1`	`AX-123 / 106`	Ergonomic Contoured Actuator Handle	3×	$31.2, 80.1, 95.7$	3.1×	$5,204/15,229$
`pf_model_mania_2014`	`AX-118 / 126`	Model Mania 2014 Tri-Boss Trunnion	3×	$29.5, 49.6, 68.1$	2.3×	$4,331/4,364$
`pf_ax_166`	`AX-166 / 176`	Slotted Eccentric Cam Housing	3×	$21.4, 37.6, 60.6$	2.8×	$3,245/3,504$
`pf_ax_105`	`AX-105`	Counterbored Bearing Pillow Block	3×	$28.9, 36.5, 39.9$	1.4×	$2,839/2,849$
`pf_stepped_shaft`	`AX-017 / 014`	Keyed Multi-Diameter Drive Shaft	3×	$8.2, 14.1, 17.2$	2.1×	797/955

Given any pair $(\tau_w, \tau_l) \in \mathcal{T}_f \times \mathcal{T}_f$ with $U(\tau_w; B_f) - U(\tau_l; B_f) \geq \delta$, the Horizon-Weighted CAD-DPO loss [1] is:

$$\mathcal{L}_{\text{CAD-DPO}}(\pi_\theta; \pi_{\text{ref}}) = -\mathbb{E}_{(B_f, \tau_w, \tau_l)} \left[\log \sigma \left(\beta \log \frac{\pi_\theta(\tau_w \mid B_f)}{\pi_{\text{ref}}(\tau_w \mid B_f)} - \beta \log \frac{\pi_\theta(\tau_l \mid B_f)}{\pi_{\text{ref}}(\tau_l \mid B_f)} \right) \right]. \tag{2}$$

3 Cryptographic Deduplication & Holdout Split

To guarantee zero data contamination between public evaluation and enterprise post-training:

1. **Ingestion-Time SHA-256 Deliverable Gate:** Every uploaded native CAD and `.STEP` file is hashed via SHA-256 against the global deliverable index. Byte-identical deliverable re-uploads across sessions are rejected at ingestion (`HTTP 400 duplicate_deliverables`), ensuring 0 duplicate deliverable hashes across all 105 sessions.

2. **Public vs. Private Partition:** While 62 episodes are published openly in `spacetime-ml/cad-cua`, the remaining 43+ verified human sessions and multi-engineer preference pairs are maintained in an air-gapped enterprise vault for frontier lab licensing.

References

[1] R. Rafailov, A. Sharma, E. Mitchell, et al., "Direct Preference Optimization: Your Language Model is Secretly a Reward Model," *NeurIPS*, 2023.

[2] Z. Shao et al., "DeepSeekMath: Pushing the Limits of Mathematical Reasoning in Open Language Models," *arXiv:2402.03300*, 2024.