

Semantic Code Similarity Analysis (BERT)

Antigravity Agent

January 29, 2026

1 Methodology

We utilized the pre-trained **TF-IDF (BERT Fallback)** model to embed C++ code functions and control blocks into a high-dimensional vector space. We then computed Cosine Similarity to identify semantically similar code regions.

2 Top Similar Pairs

The following pairs exhibit the highest semantic similarity (> 0.8):

Unit A	Unit B	Similarity
record_manual_install/if_1_3	prune_orphans/if_1_12	1.0000
prune_orphans/if_2_3	uninstall_package/if_2_3	1.0000
prune_orphans/if_2_3	run_package_tests/if_2_3	1.0000
prune_orphans/if_2_3	run_all_package_tests/if_2_3	1.0000
prune_orphans/if_2_3	freeze_environment/if_2_3	1.0000
prune_orphans/if_2_3	audit_environment/if_2_3	1.0000
prune_orphans/if_2_3	verify_environment/if_2_3	1.0000
prune_orphans/for_1_2	uninstall_package/for_1_2	1.0000
prune_orphans/for_1_2	run_package_tests/for_1_2	1.0000
prune_orphans/for_1_2	run_all_package_tests/for_1_2	1.0000
prune_orphans/for_1_2	freeze_environment/for_1_2	1.0000
prune_orphans/for_1_2	audit_environment/for_1_2	1.0000
prune_orphans/for_1_2	verify_environment/for_1_2	1.0000
uninstall_package/if_2_3	run_package_tests/if_2_3	1.0000
uninstall_package/if_2_3	run_all_package_tests/if_2_3	1.0000
uninstall_package/if_2_3	freeze_environment/if_2_3	1.0000
uninstall_package/if_2_3	audit_environment/if_2_3	1.0000
uninstall_package/if_2_3	verify_environment/if_2_3	1.0000
uninstall_package/for_1_2	run_package_tests/for_1_2	1.0000
uninstall_package/for_1_2	run_all_package_tests/for_1_2	1.0000