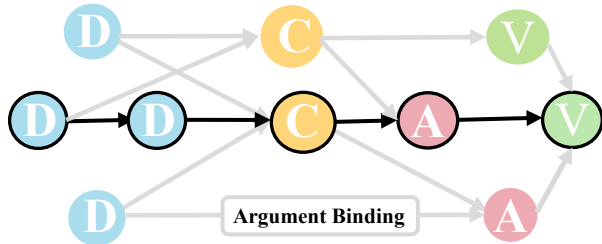


Training Stage: Execution-First Data Synthesis

DB query, Computation, Analysis, Visualization

Program
Graph
Sampling



Execution Engine

Golden Trace : $\tau^* = \dots V_i X_i r_i S_i \dots$

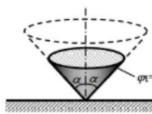
Error-Recovery: $\tau_{\text{aug}}^* \dots V_i X_i e_i S_i V_i X_i O_i \dots$

Trace-to-Question
Generation

\dots Natural Problems with Information Control \dots

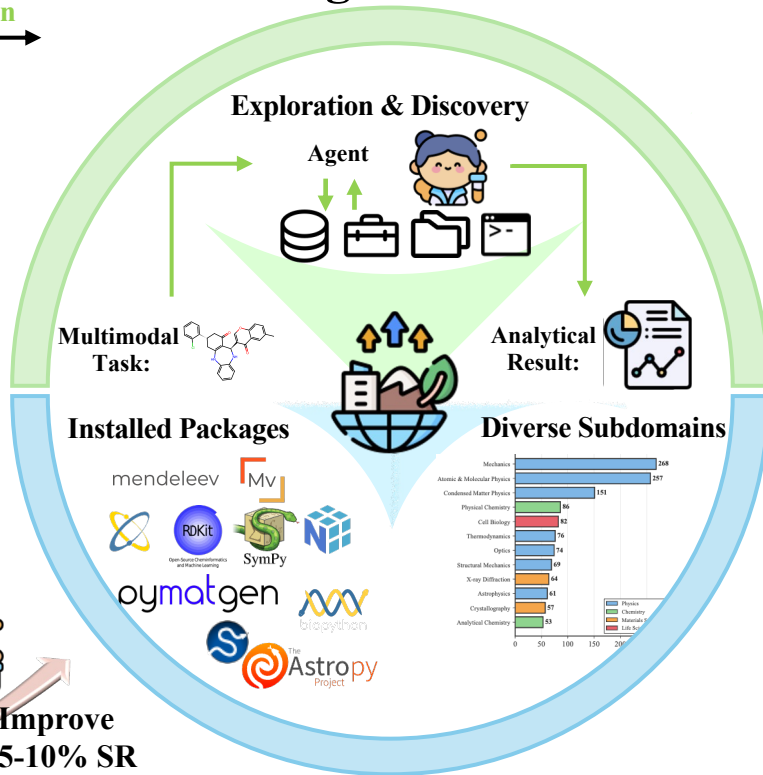
Problem Q & Golden Traces

An infinitely long conducting
circular cone with an apex angle
 $2\alpha = n\pi/6 \dots$



Improve
5-10% SR

SciAgentGYM



SciAgent-Bench

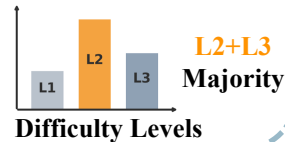
Task Suite:



Biology Chemistry Physics Material



259 Tasks
1134 Subquestions



Execution Engine

Database



Computation



Analysis



Visualize



Sandboxed
Runtime



Error
Capture



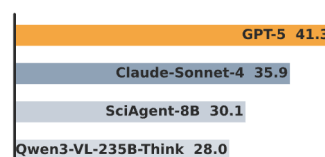
State
Management

Evaluation

SR
Correctness



SPL
Efficiency



Success Rate