

**Input: A single, long prompt**

Containing: 1. A puzzle in an unknown token language,  
2. Grammar rules for the language, and  
3. An example solution using a backtracking algorithm

LLM

**Internal Tasks for the LLM****Task 1: Grammar Learning**

Interpret token language rules

**Task 2: Algorithm Learning**

Extract backtracking structure from the example

**Task 3: Code Implementation**

Combine grammar and algorithm to implement a Python solver

**Output: Final Python solver code****Automated Verification by Pytest**

Runtime Error (e.g., KeyError)

**Result: Failure**