Experiment 'P_2_4_3_product' Results

December 18, 2025

Experiment outcome: SUCCESS
Bad responses: 0
Responses containing `assume` **:** 0
Resolution attempts: 1
Hard fails (resolution): 0
Soft fails (resolution): 0
Verification attempts: 1

Problem Specification

Problem name: P_2_4_3_product
Natural language statement: Write a method that takes two integers and then returns the product.
Method signature: product(x: int, y: int) returns (result:int)

Ensures

- `result == x * y`

GenAI interactions

Below you will find all interactions between the 'user' (program) and the 'assistant' (OpenAI).

Program → GenAI

```
You are given the following task to perform in Dafny:

Write a method that takes two integers and then returns the product.

The signature should be:

method product(x: int, y: int) returns (result:int)

The method should respect the following contract:

ensures result == x * y

Produce and show only the Dafny body of this method, including the curly braces that
    ↪ surround it. Do not show the signature nor contract. You must not use 'assume'
    ↪ anywhere in your code.
```

GenAI → Program

System fingerprint:
ID: chatcmpl-Co4P9sOlTK1vMpzms2OEHOxKAxtsP
Model: gpt-5.1-chat-latest
Created at: 1766049367
Finish reason: stop
Usage: 240 tokens in, and 92 tokens out

```
{
  result := x * y;
}
```

Final Program

```
method product(x: int, y: int) returns (result:int)
      ensures result == x * y
{
  result := x * y;
}
```

Total Token Usage

Input tokens: 240
Output tokens: 92
Reasoning tokens: 64
Sum of 'total tokens': 332

Experiment Timings

Overall Experiment started at 1766049368593, ended at 1766049371893, lasting 3300ms (3.30 seconds)
Iteration #1 started at 1766049368593, ended at 1766049371893, lasting 3300ms (3.30 seconds)