

Experiment ‘P_2_4_1_sum’ 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_1_sum

Natural language statement: Write a method that takes two integers and then returns the sum.

Method signature: sum(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 sum.

The signature should be:

```
method sum(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-Co4P2PgMlzPdv97jY9XLHRD0rRnz5

Model: gpt-5.1-chat-latest

Created at: 1766049360

Finish reason: stop

Usage: 240 tokens in, and 28 tokens out

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

Final Program

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

Total Token Usage

Input tokens: 240

Output tokens: 28

Reasoning tokens: 0

Sum of ‘total tokens’: 268

Experiment Timings

Overall Experiment started at 1766049360845, ended at 1766049364085, lasting 3240ms (3.24 seconds)
Iteration #1 started at 1766049360845, ended at 1766049364085, lasting 3240ms (3.24 seconds)