Experiment 'p5_2_a_allTheSame' 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: p5_2_a_allTheSame
Natural language statement: Write a method returning true if the arguments are all the same.
Method signature: p5_2_a_allTheSame(x: real, y: real, z: real) returns (allSame: bool)

Ensures

- `allSame == (x == y && y == z)`

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 returning true if the arguments are all the same.

The signature should be:

method p5_2_a_allTheSame(x: real, y: real, z: real) returns (allSame: bool)

The method should respect the following contract:

ensures allSame == (x == y && y == z)

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-Co4tVVYWhIoMzpSHTWmSPBCqh8R0u
Model: gpt-5.1-chat-latest
Created at: 1766051249
Finish reason: stop
Usage: 278 tokens in, and 33 tokens out

```
{
  allSame := x == y && y == z;
}
```

Final Program

```
method p5_2_a_allTheSame(x: real, y: real, z: real) returns (allSame: bool)
       ensures allSame == (x == y && y == z)
{
  allSame := x == y && y == z;
}
```

Total Token Usage

Input tokens: 278
Output tokens: 33
Reasoning tokens: 0
Sum of 'total tokens': 311

Experiment Timings

Overall Experiment started at 1766051250868, ended at 1766051253741, lasting 2873ms (2.87 seconds)
Iteration #1 started at 1766051250868, ended at 1766051253741, lasting 2873ms (2.87 seconds)