

Experiment ‘P_3_4_Classification’ 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_3_4_Classification

Natural language statement: Write a method that takes three numbers and returns "all the same" if they are all the same, "all different" if they are all different, and "neither" otherwise.

Method signature: P_3_4_Classification (i:int, j: int, k: int) returns (s: string)

Ensures

- if i == j == k then s == "all the same" else if i != j && j!=k && i!=k then s == "all different" else s== "neither"

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 three numbers and returns "all the same" if they are all the same, "all different" if they are all different, and "neither" otherwise.

The signature should be:

```
method P_3_4_Classification (i:int, j: int, k: int) returns (s: string)
```

The method should respect the following contract:

```
ensures if i == j == k then s == "all the same" else if i != j && j!=k && i!=k then s ==  
    ↪ "all different" else s== "neither"
```

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: chatmpl-Co4SbufU6j92JeBxsuQnKglRbIb2

Model: gpt-5.1-chat-latest

Created at: 1766049581

Finish reason: stop

Usage: 382 tokens in, and 84 tokens out

```
{  
    if i == j && j == k {  
        s := "all the same";  
    } else if i != j && j != k && i != k {  
        s := "all different";  
    } else {  
        s := "neither";  
    }  
}
```

Final Program

```
method P_3_4_Classification (i:int, j: int, k: int) returns (s: string)  
    ensures if i == j == k then s == "all the same" else if i != j && j!=k && i!=k  
        ↪ then s == "all different" else s== "neither"  
{  
    if i == j && j == k {
```

```
s := "all the same";
} else if i != j && j != k && i != k {
    s := "all different";
} else {
    s := "neither";
}
}
```

Total Token Usage

Input tokens: 382

Output tokens: 84

Reasoning tokens: 0

Sum of ‘total tokens’: 466

Experiment Timings

Overall Experiment started at 1766049583094, ended at 1766049586006, lasting 2912ms (2.91 seconds)

Iteration #1 started at 1766049583094, ended at 1766049586006, lasting 2912ms (2.91 seconds)