

Experiment ‘2_4_6_max’ 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: 2_4_6_max

Natural language statement: Write a method that takes two integers and then returns the maximum (the larger of the two).

Method signature: max(x: int, y: int) returns (result:int)

Ensures

- `result == x || result == y`
- `if x >= y then result == x else result == 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 maximum (the larger of the ↪ two).

The signature should be:

```
method max(x: int, y: int) returns (result:int)
```

The method should respect the following contract:

```
ensures result == x || result == y, ensures if x >= y then result == x else result == 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: chatmpl-Co4PKVaGzlMAthLdbwSCJhnA3VIV1

Model: gpt-5.1-chat-latest

Created at: 1766049378

Finish reason: stop

Usage: 284 tokens in, and 111 tokens out

```
{
    if x >= y {
        result := x;
    } else {
        result := y;
    }
}
```

Final Program

```
method max(x: int, y: int) returns (result:int)
    ensures result == x || result == y
    ensures if x >= y then result == x else result == y
{
    if x >= y {
        result := x;
    } else {
        result := y;
    }
}
```

}

Total Token Usage

Input tokens: 284

Output tokens: 111

Reasoning tokens: 64

Sum of ‘total tokens’: 395

Experiment Timings

Overall Experiment started at 1766049379764, ended at 1766049382773, lasting 3009ms (3.01 seconds)

Iteration #1 started at 1766049379765, ended at 1766049382773, lasting 3008ms (3.01 seconds)