

Experiment ‘p_1_3’ 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_1_3

Natural language statement: Write a method that returns the product of the first ten positive integers,
 $1 * 2 * \dots * 10$.

Method signature: p_1_3() returns (i : int)

Ensures

- i == factorial(10)

Functional Code Given

```
function factorial(n : int) : int
  decreases n
{
  if n <= 1 then 1 else n * factorial(n - 1)
}
```

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 returns the product of the first ten positive integers, $1 * 2 * \dots * \hookrightarrow 10$.

The signature should be:

```
method p_1_3() returns (i : int)
```

The method should respect the following contract:

```
ensures i == factorial(10)
```

The contract uses the following dafny code:

```
function factorial(n : int) : int
  decreases n
{
  if n <= 1 then 1 else n * factorial(n - 1)
}
```

These function(s) must not be used in your implementation of the method.

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-Co4Ll7ntL5DAeCiwgpKSfVgPFZtY

Model: gpt-5.1-chat-latest

Created at: 1766049157

Finish reason: stop

Usage: 374 tokens in, and 374 tokens out

```
{
  i := 1 * 2 * 3 * 4 * 5 * 6 * 7 * 8 * 9 * 10;
}
```

Final Program

```
function factorial(n : int) : int
  decreases n
```

```

{
    if n <= 1 then 1 else n * factorial(n - 1)
}

method p_1_3() returns (i : int)
    ensures i == factorial(10)
{
    i := 1 * 2 * 3 * 4 * 5 * 6 * 7 * 8 * 9 * 10;
}

```

Total Token Usage

Input tokens: 374

Output tokens: 374

Reasoning tokens: 320

Sum of ‘total tokens’: 748

Experiment Timings

Overall Experiment started at 1766049159361, ended at 1766049165251, lasting 5890ms (5.89 seconds)

Iteration #1 started at 1766049159362, ended at 1766049165251, lasting 5889ms (5.89 seconds)