

# Experiment 'p\_1\_1' 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\_1

**Natural language statement:** Write a method that returns the greeting "Hugo".

**Method signature:** p\_1\_1() returns (s : string)

### Ensures

- s == "Hugo"

## 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 greeting "Hugo".

The signature should be:

method p\_1\_1() returns (s : string)

The method should respect the following contract:

ensures s == "Hugo"

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-Co4LdRR3VxcRiRVTsN4rLmZlDqQEd

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049149

**Finish reason:** stop

**Usage:** 232 tokens in, and 29 tokens out

```
{  
  s := "Hugo";  
}
```

## Final Program

```
method p_1_1() returns (s : string)  
  ensures s == "Hugo"  
{  
  s := "Hugo";  
}
```

## Total Token Usage

Input tokens: 232

Output tokens: 29

Reasoning tokens: 0

Sum of ‘total tokens’: 261

## Experiment Timings

**Overall Experiment** started at 1766049150828, ended at 1766049154121, lasting 3293ms (3.29 seconds)

**Iteration #1** started at 1766049150831, ended at 1766049154120, lasting 3289ms (3.29 seconds)



# Experiment 'p\_1\_10' 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\_10

**Natural language statement:** Write a method that returns an animal speaking a greeting in ASCII art.

**Method signature:** p\_1\_10() returns (s : string)

## Ensures

- `s == " /\_\_/\ /---\\\n ( ' ' ) / Hello \\n ( - ) | New |\n | | | \\ Coder /\n ( __|__ )  
\\-----/\n"`

## 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 an animal speaking a greeting in ASCII art.

The signature should be:

```
method p_1_10() returns (s : string)
```

The method should respect the following contract:

```
ensures s == " /\_\_/\_\ /---\\n ( ' ' ) / Hello \\n ( - ) | New |n | | | \\ Coders /\n (
  ↳ __|__ ) \\----/\n"
```

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-Co4MK08KKHXMsafAy0889GKO5zReU

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049192

**Finish reason:** stop

**Usage:** 336 tokens in, and 790 tokens out

```
{
  s := " /\_\_/\_\ /---\\n ( ' ' ) / Hello \\n ( - ) | New |n | | | \\ Coders /\n (
    ↳ __|__ ) \\----/\n";
}
```

## Final Program

```
method p_1_10() returns (s : string)
  ensures s == " /\_\_/\_\ /---\\n ( ' ' ) / Hello \\n ( - ) | New |n | | | \\
    ↳ Coders /\n ( __|__ ) \\----/\n"
{
  s := " /\_\_/\_\ /---\\n ( ' ' ) / Hello \\n ( - ) | New |n | | | \\ Coders /\n (
    ↳ __|__ ) \\----/\n";
}
```

## Total Token Usage

Input tokens: 336

Output tokens: 790

Reasoning tokens: 704

Sum of ‘total tokens’: 1126

## Experiment Timings

**Overall Experiment** started at 1766049193764, ended at 1766049203653, lasting 9889ms (9.89 seconds)

**Iteration #1** started at 1766049193764, ended at 1766049203653, lasting 9889ms (9.89 seconds)



# Experiment 'p\_1\_2' 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\_2

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

**Method signature:** p\_1\_2() returns (i : int)

### Ensures

- $i == (10 * 11) / 2$

## 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 sum of the first ten positive integers,  $1 + 2 + \dots + 10$ .

The signature should be:

method p\_1\_2() returns (i : int)

The method should respect the following contract:

ensures i == (10 \* 11) / 2

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-Co4LhWDjlOmq3NW72OJRVh9RQMxdj

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049153

**Finish reason:** stop

**Usage:** 272 tokens in, and 228 tokens out

```
{
  var sum := 0;
  var k := 1;
  while k <= 10
    invariant 1 <= k && k <= 11
    invariant sum == (k - 1) * k / 2
  {
    sum := sum + k;
    k := k + 1;
  }
  i := sum;
}
```

## Final Program

```
method p_1_2() returns (i : int)
  ensures i == (10 * 11) / 2
{
  var sum := 0;
```

```
var k := 1;
while k <= 10
  invariant 1 <= k && k <= 11
  invariant sum == (k - 1) * k / 2
{
  sum := sum + k;
  k := k + 1;
}
i := sum;
}
```

## Total Token Usage

Input tokens: 272

Output tokens: 228

Reasoning tokens: 128

Sum of 'total tokens': 500

## Experiment Timings

**Overall Experiment** started at 1766049154896, ended at 1766049158737, lasting 3841ms (3.84 seconds)

**Iteration #1** started at 1766049154897, ended at 1766049158737, lasting 3840ms (3.84 seconds)



# 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

- $i == \text{factorial}(10)$

### 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,  $1 * 2 * \dots *$   
↪ 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:** chatcmpl-Co4Ll7tntL5DAeCiwgpkSfVgPFZtY

**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)



# Experiment 'p\_1\_4' 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\_4

**Natural language statement:** Write a method that returns the balance of an account after the first, second, and third year. The account has an initial balance of \$1,000 and earns 5 percent interest per year.

**Method signature:** p\_1\_4() returns (first : real, second : real, third : real)

## Ensures

- `first == 1000.0 + 1000.0 * 0.05`
- `second == first + first * 0.05`
- `third == second + second * 0.05`

## 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 balance of an account after the first, second, and third  
→ year. The account has an initial balance of \$1,000 and earns 5 percent interest  
→ per year.

The signature should be:

method p\_1\_4() returns (first : real, second : real, third : real)

The method should respect the following contract:

ensures first == 1000.0 + 1000.0 \* 0.05, ensures second == first + first \* 0.05, ensures  
→ third == second + second \* 0.05

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-Co4LtLiqDf5lt933fq7niQll85kbt

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049165

**Finish reason:** stop

**Usage:** 380 tokens in, and 131 tokens out

```
{
  first := 1000.0 + 1000.0 * 0.05;
  second := first + first * 0.05;
  third := second + second * 0.05;
}
```

## Final Program

```
method p_1_4() returns (first : real, second : real, third : real)
  ensures first == 1000.0 + 1000.0 * 0.05
  ensures second == first + first * 0.05
  ensures third == second + second * 0.05
{
  first := 1000.0 + 1000.0 * 0.05;
  second := first + first * 0.05;
  third := second + second * 0.05;
}
```

}

## Total Token Usage

Input tokens: 380

Output tokens: 131

Reasoning tokens: 64

Sum of 'total tokens': 511

## Experiment Timings

**Overall Experiment** started at 1766049166007, ended at 1766049170434, lasting 4427ms (4.43 seconds)

**Iteration #1** started at 1766049166007, ended at 1766049170434, lasting 4427ms (4.43 seconds)



# Experiment 'p\_1\_5' 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\_5

**Natural language statement:** Write a method that returns the name Dave inside a box on the screen.

Do your best to approximate lines with characters such as | - +.

**Method signature:** p\_1\_5() returns (s : string)

## Ensures

- s == "+-----+\n| Dave |\n+-----+\n"

## 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 name Dave inside a box on the screen. Do your best to  
↪ approximate lines with characters such as | - +.

The signature should be:

method p\_1\_5() returns (s : string)

The method should respect the following contract:

ensures s == "+-----+\n| Dave |\n+-----+\n"

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-Co4Ly2EE7UcPQdUrNnUrc9PPtZXq9

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049170

**Finish reason:** stop

**Usage:** 286 tokens in, and 106 tokens out

```
{  
  s := "+-----+\n| Dave |\n+-----+\n";  
}
```

### Final Program

```
method p_1_5() returns (s : string)  
  ensures s == "+-----+\n| Dave |\n+-----+\n"  
{  
  s := "+-----+\n| Dave |\n+-----+\n";  
}
```

### Total Token Usage

**Input tokens:** 286

**Output tokens:** 106

**Reasoning tokens:** 64

**Sum of ‘total tokens’:** 392

## Experiment Timings

**Overall Experiment** started at 1766049171097, ended at 1766049175176, lasting 4079ms (4.08 seconds)

**Iteration #1** started at 1766049171098, ended at 1766049175176, lasting 4078ms (4.08 seconds)



# Experiment ‘p\_1\_6’ 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\_6

**Natural language statement:** Write a method that returns the name Harry in large letters.

**Method signature:** p\_1\_6() returns (s : string)

## Ensures

- `s == "* * * ***** * *\n* * * * * * * *\n***** * * ***** * *\n* * * * * * * *"`

## 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 name Harry in large letters.

The signature should be:

method p\_1\_6() returns (s : string)

The method should respect the following contract:

```
ensures s == "* * * ***** ***** * *\n* * * * * * * * *\n***** * * ***** ***** * *\n* *  
↪ ***** * * * * *\n* * * * * * * * *"
```

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-Co4M2r5LeMth2I2iZg4pzXC6QASB8

Model: gpt-5.1-chat-latest

Created at: 1766049174

Finish reason: stop

Usage: 396 tokens in, and 239 tokens out

```
{  
  s := "* * * ***** ***** * *\n* * * * * * * * *\n***** * * ***** ***** * *\n* *  
↪ ***** * * * * *\n* * * * * * * * *";  
}
```

## Final Program

```
method p_1_6() returns (s : string)  
  ensures s == "* * * ***** ***** * *\n* * * * * * * * *\n***** * * ***** ***** * *\n* *  
↪ *\n* * ***** * * * *\n* * * * * * * * *"  
{  
  s := "* * * ***** ***** * *\n* * * * * * * * *\n***** * * ***** ***** * *\n* *  
↪ ***** * * * * *\n* * * * * * * * *";  
}
```

## Total Token Usage

Input tokens: 396

Output tokens: 239

Reasoning tokens: 128

Sum of 'total tokens': 635

## Experiment Timings

**Overall Experiment** started at 1766049175835, ended at 1766049180112, lasting 4277ms (4.28 seconds)

**Iteration #1** started at 1766049175840, ended at 1766049180112, lasting 4272ms (4.27 seconds)



# Experiment 'p\_1\_7' 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\_7

**Natural language statement:** Write a method that returns a face in ASCII art.

**Method signature:** p\_1\_7() returns (s : string)

### Ensures

- `s == " ////\n +\"\"\"\"\"\"+\\n(| o o |)\n | ^ |\n | ' - ' |\n +-----+\n"`

## 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 a face in ASCII art.

The signature should be:

method p\_1\_7() returns (s : string)

The method should respect the following contract:

ensures s == " /////\n +\"\\\"\\\"\\\"\\\"+\\n(| o o |)\n | ^ |\n | ‘’- |\n +-----+\n"

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-Co4M7uDzWhzzdgh6cKDukaT8W5rj4

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049179

**Finish reason:** stop

**Usage:** 296 tokens in, and 196 tokens out

```
{ s := " /////\n +\"\\\"\\\"\\\"\\\"\\\"+\\n(| o o |)\n | ^ |\n | ‘’- |\n +-----+\n"; }
```

## Final Program

```
method p_1_7() returns (s : string)
    ensures s == " /////\n +\"\\\"\\\"\\\"\\\"\\\"+\\n(| o o |)\n | ^ |\n | ‘’- |\n +-----+\n"
{ s := " /////\n +\"\\\"\\\"\\\"\\\"\\\"+\\n(| o o |)\n | ^ |\n | ‘’- |\n +-----+\n"; }
```

## Total Token Usage

**Input tokens:** 296

**Output tokens:** 196

**Reasoning tokens:** 128

**Sum of ‘total tokens’:** 492

## Experiment Timings

**Overall Experiment** started at 1766049180796, ended at 1766049185078, lasting 4282ms (4.28 seconds)

**Iteration #1** started at 1766049180796, ended at 1766049185078, lasting 4282ms (4.28 seconds)



# Experiment ‘p\_1\_8’ 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\_8

**Natural language statement:** Write a method that returns an imitation of a Piet Mondrian painting. (Search the Internet if you are not familiar with his paintings.) Use character sequences such as @@@ or ::: to indicate different colors, and use - and | to form lines.

**Method signature:** p\_1\_8() returns (s : string)

## Ensures

- s == mystring(" !!!@@@@@@@@@@@@@@@@@@@@\n", 5)

## Functional Code Given

```
function mystring(s : string, n : nat): string
{
  if n == 0 then "" else s + mystring(s, 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 an imitation of a Piet Mondrian painting. (Search the  
→ Internet if you are not familiar with his paintings.) Use character sequences such  
→ as @@@ or ::: to indicate different colors, and use - and | to form lines.

The signature should be:

method p\_1\_8() returns (s : string)

The method should respect the following contract:

ensures s == mystring(" !!!@@@@@@@@@@@@@@@@@@@@\n", 5)

The contract uses the following dafny code:

```
function mystring(s : string, n : nat): string
{
  if n == 0 then "" else s + mystring(s, 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:** chatcmpl-Co4MC9bcfcGM3w9MdX1WirICjmd99

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049184

**Finish reason:** stop

**Usage:** 448 tokens in, and 113 tokens out

```
{
  var line := " !!!@@@@@@@@@@@@@@@@@@@@\n";
  s := line + line + line + line + line;
}
```

## Final Program

```
function mystring(s : string, n : nat): string
```

```
{
  if n == 0 then "" else s + mystring(s, n - 1)
}

method p_1_8() returns (s : string)
  ensures s == mystring(" !!!@!!!!!!!!!!!!!!!!!!!!!!!!!!!!\n", 5)
{
  var line := " !!!@!!!!!!!!!!!!!!!!!!!!!!!!!!!!\n";
  s := line + line + line + line + line;
}
```

## Total Token Usage

Input tokens: 448

Output tokens: 113

Reasoning tokens: 64

Sum of 'total tokens': 561

## Experiment Timings

**Overall Experiment** started at 1766049185847, ended at 1766049189504, lasting 3657ms (3.66 seconds)

**Iteration #1** started at 1766049185847, ended at 1766049189504, lasting 3657ms (3.66 seconds)



# Experiment 'p\_1\_9' 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\_9

**Natural language statement:** Write a method that returns a house in ASCII art.

**Method signature:** p\_1\_9() returns (s : string)

### Ensures

- s == " +\n + +\n + +\n+-----+\n| .-. | \n| | | \n+---+\n"

## 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 a house in ASCII art.

The signature should be:

method p\_1\_9() returns (s : string)

The method should respect the following contract:

ensures s == " +\n + +\n + +\n+-----+\n| .-. |\n| | | |\n+---+++\n"

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-Co4MGJ6jMxX75fbOl9E9KKHurpE2L

**Model:** gpt-5.1-chat-latest

**Created at:** 1766049188

**Finish reason:** stop

**Usage:** 292 tokens in, and 127 tokens out

```
{  
  s := " +\n + +\n + +\n+-----+\n| .-. |\n| | | |\n+---+++\n";  
}
```

## Final Program

```
method p_1_9() returns (s : string)  
  ensures s == " +\n + +\n + +\n+-----+\n| .-. |\n| | | |\n+---+++\n"  
{  
  s := " +\n + +\n + +\n+-----+\n| .-. |\n| | | |\n+---+++\n";  
}
```

## Total Token Usage

Input tokens: 292

Output tokens: 127

Reasoning tokens: 64

Sum of ‘total tokens’: 419

## Experiment Timings

**Overall Experiment** started at 1766049190180, ended at 1766049193070, lasting 2890ms (2.89 seconds)

**Iteration #1** started at 1766049190180, ended at 1766049193070, lasting 2890ms (2.89 seconds)

