Experiment ‘p4_1_d_sum_odd_numbers’ Results

December 18, 2025

Experiment outcome: FAILURE_VERIFY
Bad responses: 0
Responses containing `assume` **:** 0
Resolution attempts: 8
Hard fails (resolution): 0
Soft fails (resolution): 0
Verification attempts: 8

Problem Specification

Problem name: p4_1_d_sum_odd_numbers
Natural language statement: Write a method with loops that computes the sum of all odd numbers between a and b (inclusive), where a and b are inputs.
Method signature: p4_1_d_sum_odd_numbers(a: int, b: int) returns (sum: int)

Ensures

- `sum == sum_odd(a, b)`

Functional Code Given

```
function sum_odd(a: int, b: int): int
decreases b - a
{
  if a > b then 0
  else if a % 2 == 1 then a + sum_odd(a + 1, b)
  else sum_odd(a + 1, b)
}
```

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 with loops that computes the sum of all odd numbers between a and b (
    ↪ inclusive), where a and b are inputs.

The signature should be:

method p4_1_d_sum_odd_numbers(a: int, b: int) returns (sum: int)

The method should respect the following contract:

ensures sum == sum_odd(a, b)

The contract uses the following dafny code:

function sum_odd(a: int, b: int): int
decreases b - a
{
  if a > b then 0
  else if a % 2 == 1 then a + sum_odd(a + 1, b)
  else sum_odd(a + 1, b)
}

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-Co4ZiKMCEFtZ1FP00A6DvbFU43u20
Model: gpt-5.1-chat-latest
Created at: 1766050022
Finish reason: stop
Usage: 472 tokens in, and 173 tokens out

```
{
  var i := a;
  sum := 0;
  while i <= b
    invariant a <= i <= b + 1
    invariant sum == sum_odd(a, i - 1)
    decreases b - i
  {
    if i % 2 == 1 {
      sum := sum + i;
    }
    i := i + 1;
  }
}
```

Program → GenAI

```
When using dafny verify, the below error is emitted and verify fails:

problem_data/p4_1_d_sum_odd_numbers/program.dfy(15,21): Error: this loop invariant could
    ↪ not be proved on entry
 Related message: loop invariant violation
   |
15 | invariant a <= i <= b + 1
   | ^^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(16,18): Error: this invariant could not
    ↪ be proved to be maintained by the loop
 Related message: loop invariant violation
   |
16 | invariant sum == sum_odd(a, i - 1)
   | ^^


Dafny program verifier finished with 1 verified, 2 errors

Correct the error by altering only the method body. Produce and show only the Dafny body,
    ↪  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-Co4Zm8t2izFRDhxUOopKks9QouAW3
Model: gpt-5.1-chat-latest
Created at: 1766050026
Finish reason: stop
Usage: 762 tokens in, and 424 tokens out

```
{
  var i := a;
  sum := 0;
  while i <= b
    invariant i >= a
    invariant sum == sum_odd(a, i - 1)
    decreases b - i
  {
    if i % 2 == 1 {
      sum := sum + i;
    }
    i := i + 1;
  }
}
```

Program → GenAI

```
When using dafny verify, the below error is emitted and verify fails:

problem_data/p4_1_d_sum_odd_numbers/program.dfy(11,0): Error: a postcondition could not
    ↪ be proved on this return path
   |
11 | {
   | ^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(10,13): Related location: this is the
    ↪ postcondition that could not be proved
   |
10 |  ensures sum == sum_odd(a, b)
   | ^^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(16,18): Error: this invariant could not
    ↪ be proved to be maintained by the loop
 Related message: loop invariant violation
   |
16 | invariant sum == sum_odd(a, i - 1)
   | ^^


Dafny program verifier finished with 1 verified, 2 errors

Correct the error by altering only the method body. Produce and show only the Dafny body,
    ↪  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-Co4ZsK9WficSXxyKZ998iUsZifWed
Model: gpt-5.1-chat-latest
Created at: 1766050032
Finish reason: stop
Usage: 1144 tokens in, and 241 tokens out

```
{
  var i := a;
  sum := 0;
  while i <= b
    invariant a <= i
    invariant i <= b + 1
    invariant sum == sum_odd(a, i - 1)
    decreases b - i
  {
    if i % 2 == 1 {
      sum := sum + i;
    }
    i := i + 1;
  }
}
```

Program → GenAI

```
When using dafny verify, the below error is emitted and verify fails:

problem_data/p4_1_d_sum_odd_numbers/program.dfy(16,16): Error: this loop invariant could
    ↪ not be proved on entry
 Related message: loop invariant violation
   |
16 | invariant i <= b + 1
   | ^^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(17,18): Error: this invariant could not
    ↪ be proved to be maintained by the loop
 Related message: loop invariant violation
   |
17 | invariant sum == sum_odd(a, i - 1)
   | ^^


Dafny program verifier finished with 1 verified, 2 errors

Correct the error by altering only the method body. Produce and show only the Dafny body,
    ↪  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-Co4ZwXOZDCJzYSNdN7RKwQalG2bfq
Model: gpt-5.1-chat-latest
Created at: 1766050036
Finish reason: stop
Usage: 1414 tokens in, and 119 tokens out

```
{
  var i := a;
  sum := 0;
  while i < b + 1
    invariant i >= a
    invariant i <= b + 1
    invariant sum == sum_odd(a, i - 1)
    decreases b + 1 - i
  {
    if i % 2 == 1 {
      sum := sum + i;
    }
    i := i + 1;
  }
}
```

Program → GenAI

```
When using dafny verify, the below error is emitted and verify fails:

problem_data/p4_1_d_sum_odd_numbers/program.dfy(16,16): Error: this loop invariant could not be proved on entry
 Related message: loop invariant violation
   |
16 | invariant i <= b + 1
   | ^^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(17,18): Error: this invariant could not be proved to be maintained by the loop
 Related message: loop invariant violation
   |
17 | invariant sum == sum_odd(a, i - 1)
   | ^^


Dafny program verifier finished with 1 verified, 2 errors

Correct the error by altering only the method body. Produce and show only the Dafny body,  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-Co4a0kp3s9iORKfFmhSqwnYjeIrmD
Model: gpt-5.1-chat-latest
Created at: 1766050040
Finish reason: stop
Usage: 1731 tokens in, and 168 tokens out

```
{
  var i := a;
  sum := 0;
  while i <= b
    invariant i >= a
    invariant sum == sum_odd(a, i - 1)
    decreases b - i
  {
    if i % 2 == 1 {
      sum := sum + i;
    }
    i := i + 1;
  }
}
```

Program → GenAI

```
When using dafny verify, the below error is emitted and verify fails:

problem_data/p4_1_d_sum_odd_numbers/program.dfy(11,0): Error: a postcondition could not
    ↪ be proved on this return path
   |
11 | {
   | ^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(10,13): Related location: this is the
    ↪ postcondition that could not be proved
   |
10 |  ensures sum == sum_odd(a, b)
   | ^^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(16,18): Error: this invariant could not
    ↪ be proved to be maintained by the loop
 Related message: loop invariant violation
   |
16 | invariant sum == sum_odd(a, i - 1)
   | ^^


Dafny program verifier finished with 1 verified, 2 errors

Correct the error by altering only the method body. Produce and show only the Dafny body,
    ↪  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-Co4a4IPuOvEziNjjdVzKoXzHOUtW3
Model: gpt-5.1-chat-latest
Created at: 1766050044
Finish reason: stop
Usage: 2115 tokens in, and 126 tokens out

```
{
  var i := a;
  sum := 0;

  while i <= b + 1
    invariant a <= i <= b + 1
    invariant sum == sum_odd(a, i - 1)
    decreases b + 1 - i
  {
    if i <= b {
      if i % 2 == 1 {
        sum := sum + i;
      }
    }
    i := i + 1;
  }
}
```

Program → GenAI

```
When using dafny verify, the below error is emitted and verify fails:

problem_data/p4_1_d_sum_odd_numbers/program.dfy(16,21): Error: this loop invariant could
    ↪ not be proved on entry
 Related message: loop invariant violation
   |
16 | invariant a <= i <= b + 1
   | ^^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(16,21): Error: this invariant could not
    ↪ be proved to be maintained by the loop
 Related message: loop invariant violation
   |
16 | invariant a <= i <= b + 1
   | ^^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(17,18): Error: this invariant could not
    ↪ be proved to be maintained by the loop
 Related message: loop invariant violation
   |
17 | invariant sum == sum_odd(a, i - 1)
   | ^^


Dafny program verifier finished with 1 verified, 3 errors

Correct the error by altering only the method body. Produce and show only the Dafny body,
    ↪  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-Co4a8dWOGlSt6HgKs7zs6dgTBrI4K
Model: gpt-5.1-chat-latest
Created at: 1766050048

Finish reason: stop
Usage: 2522 tokens in, and 320 tokens out

```
{
  if a > b {
    sum := 0;
  } else {
    var i := a;
    sum := 0;
    while i <= b
      invariant a <= i <= b
      invariant sum == sum_odd(a, i - 1)
      decreases b - i
    {
      if i % 2 == 1 {
        sum := sum + i;
      }
      i := i + 1;
    }
  }
}
```

Program → GenAI

```
When using dafny verify, the below error is emitted and verify fails:

problem_data/p4_1_d_sum_odd_numbers/program.dfy(18,23): Error: this invariant could not
    ↪ be proved to be maintained by the loop
 Related message: loop invariant violation
   |
18 | invariant a <= i <= b
   | ^^

problem_data/p4_1_d_sum_odd_numbers/program.dfy(19,20): Error: this invariant could not
    ↪ be proved to be maintained by the loop
 Related message: loop invariant violation
   |
19 | invariant sum == sum_odd(a, i - 1)
   | ^^


Dafny program verifier finished with 1 verified, 2 errors

Correct the error by altering only the method body. Produce and show only the Dafny body,
    ↪  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-Co4aDd49xso3oOkpN0odgZnl7mxgT

Model: gpt-5.1-chat-latest
Created at: 1766050053
Finish reason: stop
Usage: 2790 tokens in, and 202 tokens out

```
{
  if a > b {
    sum := 0;
  } else {
    var i := a;
    sum := 0;
    while i < b + 1
      invariant i >= a
      invariant i <= b
      invariant sum == sum_odd(a, i - 1)
      decreases b + 1 - i
    {
      if i % 2 == 1 {
        sum := sum + i;
      }
      i := i + 1;
    }
  }
}
```

Final Program

```
function sum_odd(a: int, b: int): int
decreases b - a
{
  if a > b then 0
  else if a % 2 == 1 then a + sum_odd(a + 1, b)
  else sum_odd(a + 1, b)
}

method p4_1_d_sum_odd_numbers(a: int, b: int) returns (sum: int)
        ensures sum == sum_odd(a, b)
{
  if a > b {
    sum := 0;
  } else {
    var i := a;
    sum := 0;
    while i < b + 1
      invariant i >= a
      invariant i <= b
      invariant sum == sum_odd(a, i - 1)
      decreases b + 1 - i
    {
      if i % 2 == 1 {
        sum := sum + i;
      }
      i := i + 1;
    }
  }
}
```

Total Token Usage

Input tokens: 12950
Output tokens: 1773
Reasoning tokens: 832
Sum of ‘total tokens’: 14723

Experiment Timings

Iteration #8 started at 1766050055171, ended at 1766050059783, lasting 4612ms (4.61 seconds)
Overall Experiment started at 1766050024096, ended at 1766050059783, lasting 35687ms (35.69 seconds)
Iteration #4 started at 1766050038314, ended at 1766050041869, lasting 3555ms (3.56 seconds)
Iteration #5 started at 1766050041869, ended at 1766050045907, lasting 4038ms (4.04 seconds)
Iteration #6 started at 1766050045907, ended at 1766050049612, lasting 3705ms (3.71 seconds)
Iteration #7 started at 1766050049612, ended at 1766050055171, lasting 5559ms (5.56 seconds)
Iteration #1 started at 1766050024097, ended at 1766050027840, lasting 3743ms (3.74 seconds)
Iteration #2 started at 1766050027840, ended at 1766050034042, lasting 6202ms (6.20 seconds)
Iteration #3 started at 1766050034042, ended at 1766050038314, lasting 4272ms (4.27 seconds)