

Experiment ‘p4_3_d_count_vowels’ Results

December 18, 2025

Experiment outcome: FAILURE_VERIFY

Bad responses: 0

Responses containing assume : 0

Resolution attempts: 8

Hard fails (resolution): 1

Soft fails (resolution): 1

Verification attempts: 6

Problem Specification

Problem name: p4_3_d_count_vowels

Natural language statement: Write a method that takes a string and returns the number of vowels in the string.

Method signature: p4_3_d_count_vowels(s: string) returns (count: int)

Ensures

- count ≥ 0
- $\forall c :: c \in s \ \&\& \ c \in ['a', 'e', 'i', 'o', 'u', 'A', 'E', 'I', 'O', 'U'] \implies \text{count} \geq 1$
- $(\forall c :: c \in s \implies c \notin ['a', 'e', 'i', 'o', 'u', 'A', 'E', 'I', 'O', 'U']) \implies \text{count} = 0$

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 a string and returns the number of vowels in the string.

The signature should be:

```
method p4_3_d_count_vowels(s: string) returns (count: int)
```

The method should respect the following contract:

```
ensures count >= 0, ensures forall c :: c in s && c in ['a','e','i','o','u','A','E','I','  
↪ O','U'] ==> count >= 1, ensures (forall c :: c in s ==> c !in ['a','e','i','o','u'  
↪ ','A','E','I','O','U']) ==> count == 0
```

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-Co4dFAs8Tepw2tLaUmSkfKtnxt66u

Model: gpt-5.1-chat-latest

Created at: 1766050241

Finish reason: stop

Usage: 406 tokens in, and 169 tokens out

```
{  
    count := 0;  
    var i := 0;  
    while i < s.Length  
        invariant 0 <= i <= s.Length  
        invariant count >= 0  
        invariant count <= i  
    {  
        var c := s[i];  
        if c == 'a' || c == 'e' || c == 'i' || c == 'o' || c == 'u' ||  
            c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U'  
        {  
            count := count + 1;  
        }  
        i := i + 1;  
    }  
}
```

Program → GenAI

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

```
problem_data/p4_3_d_count_vowels/program.dfy(8,14): Error: type seq<char> does not have a
    ↪ member Length
|
8 | while i < s.Length
| ~~~~~

problem_data/p4_3_d_count_vowels/program.dfy(9,26): Error: type seq<char> does not have a
    ↪ member Length
|
9 | invariant 0 <= i <= s.Length
| ~~~~~

2 resolution/type errors detected in program.dfy
```

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-Co4dI7G6mhGiKt8EDUtvem7H6dQYU

Model: gpt-5.1-chat-latest

Created at: 1766050244

Finish reason: stop

Usage: 735 tokens in, and 171 tokens out

```
{
  count := 0;
  var i := 0;
  while i < |s|
    invariant 0 <= i <= |s|
    invariant count >= 0
    invariant count <= i
  {
    var c := s[i];
    if c == 'a' || c == 'e' || c == 'i' || c == 'o' || c == 'u' ||
       c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U'
    {
      count := count + 1;
    }
    i := i + 1;
  }
}
```

Program → GenAI

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

```
problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
    ↪ proved on this return path
|
5 | {
| ^
|
problem_data/p4_3_d_count_vowels/program.dfy(3,9): Related location: this is the
    ↪ postcondition that could not be proved
|
3 | ensures forall c :: c in s && c in ['a','e','i','o','u','A','E','I','O','U'] ==>
    ↪ count >= 1
| ~~~~~
|
problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
    ↪ proved on this return path
|
5 | {
| ^
|
problem_data/p4_3_d_count_vowels/program.dfy(4,92): Related location: this is the
    ↪ postcondition that could not be proved
|
4 | ensures (forall c :: c in s ==> c !in ['a','e','i','o','u','A','E','I','O','U']) ==>
    ↪ count == 0
| ^~
```

Dafny program verifier finished with 0 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-Co4dM9vEAIBdcCHDVGirAsCpBtobo

Model: gpt-5.1-chat-latest

Created at: 1766050248

Finish reason: stop

Usage: 1387 tokens in, and 395 tokens out

```
{
  count := 0;
  var i := 0;
  while i < |s|
    invariant 0 <= i <= |s|
    invariant count >= 0
    invariant count <= i
```

```

invariant (forall c :: c in s[..i] && (c == 'a' || c == 'e' || c == 'i' || c == 'o' ||
    ↪ c == 'u' || c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U') ==>
    ↪ count >= 1)
invariant ((forall c :: c in s[..i] ==> !(c == 'a' || c == 'e' || c == 'i' || c == 'o' ||
    ↪ || c == 'u' || c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U')) ==>
    ↪ count == 0)
{
    var c := s[i];
    if c == 'a' || c == 'e' || c == 'i' || c == 'o' || c == 'u' ||
        c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U'
    {
        count := count + 1;
    }
    i := i + 1;
}
}

```

Program → GenAI

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

```

problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
    ↪ proved on this return path
|
5 | {
| ^
| 

problem_data/p4_3_d_count_vowels/program.dfy(3,9): Related location: this is the
    ↪ postcondition that could not be proved
|
3 | ensures forall c :: c in s && c in ['a','e','i','o','u','A','E','I','O','U'] ==>
    ↪ count >= 1
| ^~~~~~

problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
    ↪ proved on this return path
|
5 | {
| ^
| 

problem_data/p4_3_d_count_vowels/program.dfy(4,92): Related location: this is the
    ↪ postcondition that could not be proved
|
4 | ensures (forall c :: c in s ==> c !in ['a','e','i','o','u','A','E','I','O','U']) ==>
    ↪ count == 0
| ^~

problem_data/p4_3_d_count_vowels/program.dfy(12,14): Error: this invariant could not be
    ↪ proved to be maintained by the loop
Related message: loop invariant violation
|

```

```

12 | invariant (forall c :: c in s[..i] && (c == 'a' || c == 'e' || c == 'i' || c == 'o'
  ↪ || c == 'u' || c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U') ==> count
  ↪ >= 1)
| ^
| ^

problem_data/p4_3_d_count_vowels/program.dfy(13,175): Error: this invariant could not be
  ↪ proved to be maintained by the loop
Related message: loop invariant violation
|
13 | invariant ((forall c :: c in s[..i] ==> !(c == 'a' || c == 'e' || c == 'i' || c == 'o'
  ↪ || c == 'u' || c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U')) ==>
  ↪ count == 0)
| ^^

Dafny program verifier finished with 0 verified, 4 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-Co4dSlw0G42iLViyGTMoy5kxZsw19

Model: gpt-5.1-chat-latest

Created at: 1766050254

Finish reason: stop

Usage: 2557 tokens in, and 270 tokens out

```
{
    count := 0;
    var hasVowel := false;
    var i := 0;
    while i < |s|
        invariant 0 <= i <= |s|
        invariant count >= 0
        invariant hasVowel ==> count >= 1
        invariant (!hasVowel) ==> count == 0
    {
        var c := s[i];
        if c == 'a' || c == 'e' || c == 'i' || c == 'o' || c == 'u' ||
           c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U'
        {
            count := count + 1;
            hasVowel := true;
        }
        i := i + 1;
    }
}
```

Program → GenAI

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

```
problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
  ↪ proved on this return path
|
5 | {
| ^
|
problem_data/p4_3_d_count_vowels/program.dfy(3,9): Related location: this is the
  ↪ postcondition that could not be proved
|
3 | ensures forall c :: c in s && c in ['a','e','i','o','u','A','E','I','O','U'] ==>
  ↪ count >= 1
| ^~~~~~
|
problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
  ↪ proved on this return path
|
5 | {
| ^
|
problem_data/p4_3_d_count_vowels/program.dfy(4,92): Related location: this is the
  ↪ postcondition that could not be proved
|
4 | ensures (forall c :: c in s ==> c !in ['a','e','i','o','u','A','E','I','O','U']) ==>
  ↪ count == 0
| ^~
```

Dafny program verifier finished with 0 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: chatmpl-Co4dXTbVy1RAf6GsrY8Y5UatTQ1Ha

Model: gpt-5.1-chat-latest

Created at: 1766050259

Finish reason: stop

Usage: 2834 tokens in, and 350 tokens out

```
{
  count := 0;
  ghost var hasVowel := false;
  var i := 0;
  while i < |s|
    invariant 0 <= i <= |s|
```

```

invariant count >= 0
invariant hasVowel <=> (exists c :: c in s[..i] && (c == 'a' || c == 'e' || c == 'i'
    ↪ || c == 'o' || c == 'u' || c == 'A' || c == 'E' || c == 'I' || c == 'O' || c ==
    ↪ 'U'))
invariant hasVowel ==> count >= 1
invariant (!hasVowel) ==> count == 0
{
    var c := s[i];
    if c == 'a' || c == 'e' || c == 'i' || c == 'o' || c == 'u' ||
        c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U' {
        count := count + 1;
        hasVowel := true;
    }
    i := i + 1;
}

```

Program → GenAI

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

```

problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
    ↪ proved on this return path
|
5 | {
| ^
| 

problem_data/p4_3_d_count_vowels/program.dfy(3,9): Related location: this is the
    ↪ postcondition that could not be proved
|
3 | ensures forall c :: c in s && c in ['a','e','i','o','u','A','E','I','O','U'] ==>
    ↪ count >= 1
| ^~~~~~

problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
    ↪ proved on this return path
|
5 | {
| ^
| 

problem_data/p4_3_d_count_vowels/program.dfy(4,92): Related location: this is the
    ↪ postcondition that could not be proved
|
4 | ensures (forall c :: c in s ==> c !in ['a','e','i','o','u','A','E','I','O','U']) ==>
    ↪ count == 0
| ^~

problem_data/p4_3_d_count_vowels/program.dfy(12,23): Error: this invariant could not be
    ↪ proved to be maintained by the loop
Related message: loop invariant violation
| 
```

```

12 | invariant hasVowel <=> (exists c :: c in s[..i] && (c == 'a' || c == 'e' || c == 'i'
  ↪ ' || c == 'o' || c == 'u' || c == 'A' || c == 'E' || c == 'I' || c == 'O' || c ==
  ↪ 'U'))
| ^^^^

```

Dafny program verifier finished with 0 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: chatmpl-Co4ddZAWqJUn4B11LdOlpr9IraFen

Model: gpt-5.1-chat-latest

Created at: 1766050265

Finish reason: stop

Usage: 3705 tokens in, and 350 tokens out

```
{
  count := 0;
  ghost var hasVowel := false;
  var i := 0;
  while i < |s|
    invariant 0 <= i <= |s|
    invariant count >= 0
    invariant hasVowel ==> count >= 1
    invariant (!hasVowel) ==> count == 0
    invariant hasVowel ==> (exists c :: c in s[..i] && (c == 'a' || c == 'e' || c == 'i'
      ↪ ' || c == 'o' || c == 'u' || c == 'A' || c == 'E' || c == 'I' || c == 'O' || c ==
      ↪ 'U'))
  {
    var c := s[i];
    if c == 'a' || c == 'e' || c == 'i' || c == 'o' || c == 'u' ||
      c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U'
    {
      count := count + 1;
      hasVowel := true;
    }
    i := i + 1;
  }
}
```

Program → GenAI

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

```
problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
  ↪ proved on this return path
```

```

5 | {
| ^
| problem_data/p4_3_d_count_vowels/program.dfy(3,9): Related location: this is the
|   ↪ postcondition that could not be proved
| ensures forall c :: c in s && c in ['a','e','i','o','u','A','E','I','O','U'] ==>
|   ↪ count >= 1
| ^~~~~~
| problem_data/p4_3_d_count_vowels/program.dfy(5,0): Error: a postcondition could not be
|   ↪ proved on this return path
| ^
5 | {
| ^
| problem_data/p4_3_d_count_vowels/program.dfy(4,92): Related location: this is the
|   ↪ postcondition that could not be proved
| ensures (forall c :: c in s ==> c !in ['a','e','i','o','u','A','E','I','O','U']) ==>
|   ↪ count == 0
| ^~~
| problem_data/p4_3_d_count_vowels/program.dfy(14,28): Error: this invariant could not be
|   ↪ proved to be maintained by the loop
Related message: loop invariant violation
|
14 | invariant hasVowel ==> (exists c :: c in s[..i] && (c == 'a' || c == 'e' || c == 'i'
|   ↪ || c == 'o' || c == 'u' || c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U'))
| ^~~~~~

```

Dafny program verifier finished with 0 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: chatmpl-Co4dirjW9ejXseAybfNGhiXp3j8eb

Model: gpt-5.1-chat-latest

Created at: 1766050270

Finish reason: stop

Usage: 4443 tokens in, and 512 tokens out

```
{
  count := 0;
```

```

ghost var hasVowel := false;
var i := 0;
while i < |s|
    invariant 0 <= i <= |s|
    invariant count >= 0
    invariant hasVowel == (exists j :: 0 <= j < i &&
        (s[j] == 'a' || s[j] == 'e' || s[j] == 'i' || s[j] == 'o' || s[j] == 'u' ||
         s[j] == 'A' || s[j] == 'E' || s[j] == 'I' || s[j] == 'O' || s[j] == 'U'))
    invariant hasVowel ==> count >= 1
    invariant !hasVowel ==> count == 0
{
    var c := s[i];
    var isVowel := c == 'a' || c == 'e' || c == 'i' || c == 'o' || c == 'u' ||
                  c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U';
    if isVowel {
        count := count + 1;
        hasVowel := true;
    }
    i := i + 1;
}
}

```

Program → GenAI

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

```

problem_data/p4_3_d_count_vowels/program.dfy(12,27): Warning: unusual indentation in body
    ↪ of exists (which ends at line 14, column 81); do you perhaps need parentheses?
|
12 | invariant hasVowel == (exists j :: 0 <= j < i &&
| ~~~~~~

```

Compilation failed because warnings were found and --allow-warnings is false

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: chatmpl-Co4dpT13JBlXnNVNDuYqhywR4nhE6

Model: gpt-5.1-chat-latest

Created at: 1766050277

Finish reason: stop

Usage: 4607 tokens in, and 314 tokens out

```
{
    count := 0;
    ghost var hasVowel := false;
    var i := 0;
```

```

while i < |s|
invariant 0 <= i <= |s|
invariant count >= 0
invariant hasVowel == (exists j :: 0 <= j < i && ((s[j] == 'a' || s[j] == 'e' || s[j]
    ↪ == 'i' || s[j] == 'o' || s[j] == 'u' || s[j] == 'A' || s[j] == 'E' || s[j] == 'I'
    ↪ || s[j] == 'O' || s[j] == 'U')))
invariant hasVowel ==> count >= 1
invariant !hasVowel ==> count == 0
{
var c := s[i];
var isVowel := c == 'a' || c == 'e' || c == 'i' || c == 'o' || c == 'u' ||
    c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U';
if isVowel {
    count := count + 1;
    hasVowel := true;
}
i := i + 1;
}
}

```

Final Program

```

method p4_3_d_count_vowels(s: string) returns (count: int)
    ensures count >= 0
    ensures forall c :: c in s && c in ['a','e','i','o','u','A','E','I','O','U'] ==>
        ↪ count >= 1
    ensures (forall c :: c in s ==> c !in ['a','e','i','o','u','A','E','I','O','U'])
        ↪ ==> count == 0
{
count := 0;
ghost var hasVowel := false;
var i := 0;
while i < |s|
    invariant 0 <= i <= |s|
    invariant count >= 0
    invariant hasVowel == (exists j :: 0 <= j < i && ((s[j] == 'a' || s[j] == 'e' || s[j]
        ↪ == 'i' || s[j] == 'o' || s[j] == 'u' || s[j] == 'A' || s[j] == 'E' || s[j] == 'I'
        ↪ || s[j] == 'O' || s[j] == 'U')))
    invariant hasVowel ==> count >= 1
    invariant !hasVowel ==> count == 0
{
var c := s[i];
var isVowel := c == 'a' || c == 'e' || c == 'i' || c == 'o' || c == 'u' ||
    c == 'A' || c == 'E' || c == 'I' || c == 'O' || c == 'U';
if isVowel {
    count := count + 1;
    hasVowel := true;
}
i := i + 1;
}
}

```

}

Total Token Usage

Input tokens: 20674

Output tokens: 2531

Reasoning tokens: 448

Sum of ‘total tokens’: 23205

Experiment Timings

Iteration #8 started at 1766050279198, ended at 1766050284207, lasting 5009ms (5.01 seconds)

Overall Experiment started at 1766050242892, ended at 1766050284207, lasting 41315ms (41.32 seconds)

Iteration #4 started at 1766050256223, ended at 1766050260769, lasting 4546ms (4.55 seconds)

Iteration #5 started at 1766050260769, ended at 1766050267059, lasting 6290ms (6.29 seconds)

Iteration #6 started at 1766050267059, ended at 1766050272647, lasting 5588ms (5.59 seconds)

Iteration #7 started at 1766050272647, ended at 1766050279198, lasting 6551ms (6.55 seconds)

Iteration #1 started at 1766050242892, ended at 1766050246201, lasting 3309ms (3.31 seconds)

Iteration #2 started at 1766050246201, ended at 1766050250176, lasting 3975ms (3.98 seconds)

Iteration #3 started at 1766050250176, ended at 1766050256223, lasting 6047ms (6.05 seconds)