

Experiment ‘p5_9_coneSurface’ Results

December 18, 2025

Experiment outcome: FAILURE_VERIFY

Bad responses: 0

Responses containing assume : 0

Resolution attempts: 8

Hard fails (resolution): 7

Soft fails (resolution): 0

Verification attempts: 1

Problem Specification

Problem name: p5_9_coneSurface

Natural language statement: Write a method that computes the surface area of a cone with a circular base, given radius and height.

Method signature: p5_9_coneSurface(r: real, h: real) returns (surface_area: real)

Ensures

- $(-r + (\text{surface_area} / (3.14159 * r))) * (-r + (\text{surface_area} / (3.14159 * r))) == h * h + r * r$

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 computes the surface area of a cone with a circular base, given
→ radius and height.

The signature should be:

```
method p5_9_coneSurface(r: real, h: real) returns (surface_area: real)
```

The method should respect the following contract:

```
ensures (-r + (surface_area / (3.14159 * r))) * (-r + (surface_area / (3.14159 * r))) ==  
→ h * h + r * r
```

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-Co4y9aVJwZZbIT7RURUseiLyZw0A9

Model: gpt-5.1-chat-latest

Created at: 1766051537

Finish reason: stop

Usage: 340 tokens in, and 182 tokens out

```
{  
    var slant := sqrt(h * h + r * r);  
    surface_area := 3.14159 * r * (r + slant);  
}
```

Program → GenAI

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

```
problem_data/p5_9_coneSurface/program.dfy(4,15): Error: unresolved identifier: sqrt  
|  
4 | var slant := sqrt(h * h + r * r);  
| ^^^^  
  
1 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: chatmpl-Co4yCRfppRSh5Ha429lIyxIq23Ca0

Model: gpt-5.1-chat-latest

Created at: 1766051540

Finish reason: stop

Usage: 485 tokens in, and 60 tokens out

```
{  
    var slant := (h * h + r * r) ** 0.5;  
    surface_area := 3.14159 * r * (r + slant);  
}
```

Program → GenAI

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

```
problem_data/p5_9_coneSurface/program.dfy(4,32): Error: invalid UnaryExpression  
|  
4 | var slant := (h * h + r * r) ** 0.5;  
| ^
```

1 parse 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: chatmpl-Co4yEa5QGzfL4WrfZ8b2tFGGkoWV3

Model: gpt-5.1-chat-latest

Created at: 1766051542

Finish reason: stop

Usage: 673 tokens in, and 503 tokens out

```
{  
    var slant := Math.Sqrt(h * h + r * r);  
    surface_area := 3.14159 * r * (r + slant);  
}
```

Program → GenAI

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

```
problem_data/p5_9_coneSurface/program.dfy(4,15): Error: unresolved identifier: Math
```

```

4 | var slant := Math.Sqrt(h * h + r * r);
| ^^^^

1 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-Co4yK2EUWVLx2BHRKW9dIqy9Pmb8M

Model: gpt-5.1-chat-latest

Created at: 1766051548

Finish reason: stop

Usage: 849 tokens in, and 760 tokens out

```
{
    var slant := RealMath.Sqrt(h * h + r * r);
    surface_area := 3.14159 * r * (r + slant);
}
```

Program → GenAI

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

```
problem_data/p5_9_coneSurface/program.dfy(4,15): Error: unresolved identifier: RealMath
|
4 | var slant := RealMath.Sqrt(h * h + r * r);
| ^^^^^^^^

1 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-Co4yTlpYi78uzE66UX8OJGXS1Xvs8

Model: gpt-5.1-chat-latest

Created at: 1766051557

Finish reason: stop

Usage: 1034 tokens in, and 695 tokens out

```
{
    var slant := Sqrt(h * h + r * r);
    surface_area := 3.14159 * r * (r + slant);
```

```
}
```

Program → GenAI

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

```
problem_data/p5_9_coneSurface/program.dfy(4,15): Error: unresolved identifier: Sqrt
|
4 | var slant := Sqrt(h * h + r * r);
| ^^^^
1 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-Co4ybGkJalWptArpvRoCTfOcPuI5x

Model: gpt-5.1-chat-latest

Created at: 1766051565

Finish reason: stop

Usage: 1211 tokens in, and 508 tokens out

```
{
    var slant :| slant * slant == h * h + r * r;
    surface_area := 3.14159 * r * (r + slant);
}
```

Program → GenAI

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

```
problem_data/p5_9_coneSurface/program.dfy(2,29): Error: possible division by zero
|
2 | ensures (-r + (surface_area / (3.14159 * r))) * (-r + (surface_area / (3.14159 * r)))
|   ↳ ) == h * h + r * r
| ^
problem_data/p5_9_coneSurface/program.dfy(3,0): Error: a postcondition could not be
|   ↳ proved on this return path
|
3 | {
| ^
problem_data/p5_9_coneSurface/program.dfy(2,87): Related location: this is the
|   ↳ postcondition that could not be proved
|
```

```

2 | ensures (-r + (surface_area / (3.14159 * r))) * (-r + (surface_area / (3.14159 * r)))
  |   ↪ ) == h * h + r * r
  |   ^
  |
problem_data/p5_9_coneSurface/program.dfy(4,12): Error: cannot establish the existence of
  ↪ LHS values that satisfy the such-that predicate. Note, no trigger was found for
  ↪ the such-that predicate, which may be the reason the proof failed. To give a
  ↪ trigger explicitly, use the {::trigger} attribute. For more information, see the
  ↪ section on quantifier instantiation rules in the reference manual.
  |
4 | var slant :| slant * slant == h * h + r * r;
  | ~~~~~
  |

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-Co4yjutNXGPgJWII08C3oDw3t7uow

Model: gpt-5.1-chat-latest

Created at: 1766051573

Finish reason: stop

Usage: 1945 tokens in, and 534 tokens out

```
{
  if r == 0.0 {
    assert false;
  }
  var slant :| {::trigger slant * slant} slant * slant == h * h + r * r;
  surface_area := 3.14159 * r * (r + slant);
}
```

Program → GenAI

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

```

problem_data/p5_9_coneSurface/program.dfy(7,15): Error: invalid UnaryExpression
  |
7 | var slant :| {::trigger slant * slant} slant * slant == h * h + r * r;
  | ^
  |
1 parse 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-Co4yqealE0pGxur3zuStMLeUv8fJq

Model: gpt-5.1-chat-latest

Created at: 1766051580

Finish reason: stop

Usage: 1898 tokens in, and 88 tokens out

```
{  
    if r == 0.0 {  
        assert false;  
    }  
    {:trigger slant * slant}  
    var slant :| slant * slant == h * h + r * r;  
    surface_area := 3.14159 * r * (r + slant);  
}
```

Final Program

```
method p5_9_coneSurface(r: real, h: real) returns (surface_area: real)  
    ensures (-r + (surface_area / (3.14159 * r))) * (-r + (surface_area / (3.14159 * r  
        ↪ ))) == h * h + r * r  
{  
    if r == 0.0 {  
        assert false;  
    }  
    {:trigger slant * slant}  
    var slant :| slant * slant == h * h + r * r;  
    surface_area := 3.14159 * r * (r + slant);  
}
```

Total Token Usage

Input tokens: 8435

Output tokens: 3330

Reasoning tokens: 2816

Sum of ‘total tokens’: 11765

Experiment Timings

Iteration #8 started at 1766051582534, ended at 1766051584932, lasting 2398ms (2.40 seconds)

Overall Experiment started at 1766051538961, ended at 1766051584932, lasting 45971ms (45.97 seconds)

Iteration #4 started at 1766051550671, ended at 1766051559492, lasting 8821ms (8.82 seconds)

Iteration #5 started at 1766051559492, ended at 1766051567676, lasting 8184ms (8.18 seconds)

Iteration #6 started at 1766051567676, ended at 1766051575171, lasting 7495ms (7.50 seconds)

Iteration #7 started at 1766051575171, ended at 1766051582534, lasting 7363ms (7.36 seconds)

Iteration #1 started at 1766051538961, ended at 1766051542050, lasting 3089ms (3.09 seconds)

Iteration #2 started at 1766051542050, ended at 1766051544113, lasting 2063ms (2.06 seconds)

Iteration #3 started at 1766051544113, ended at 1766051550671, lasting 6558ms (6.56 seconds)