

Experiment ‘2_8_rectangle_properties’ 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: 2_8_rectangle_properties

Natural language statement: Write a method that takes the lengths of the sides of a rectangle. Then return the area and perimeter of the rectangle and the length of the diagonal (use the Pythagorean theorem).

Method signature: rectangle_properties(length: real, width: real) returns (area: real, perimeter: real, diagonal: real)

Ensures

- $\text{area} == \text{length} * \text{width}$
- $\text{perimeter} == 2.0 * (\text{length} + \text{width})$
- $\text{diagonal} * \text{diagonal} == \text{length} * \text{length} + \text{width} * \text{width}$

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 the lengths of the sides of a rectangle. Then return the area
→ and perimeter of the rectangle and the length of the diagonal (use the Pythagorean
→ theorem).

The signature should be:

```
method rectangle_properties(length: real, width: real) returns (area: real, perimeter:  
→ real, diagonal: real)
```

The method should respect the following contract:

```
ensures area == length * width, ensures perimeter == 2.0 * (length + width), ensures  
→ diagonal * diagonal == length * length + width * width
```

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-Co4Pd2JrQ0jX4NXyW62vLmR9l0L3h

Model: gpt-5.1-chat-latest

Created at: 1766049397

Finish reason: stop

Usage: 362 tokens in, and 119 tokens out

```
{  
    area := length * width;  
    perimeter := 2.0 * (length + width);  
    diagonal := sqrt(length * length + width * width);  
}
```

Program → GenAI

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

```
problem_data/2_8_rectangle_properties/program.dfy(8,14): Error: unresolved identifier:  
→ sqrt  
|  
8 | diagonal := sqrt(length * length + width * width);  
| ^~~~
```

```
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-Co4Pg7fOZUsscsEr0zZ4O1KMtnENY

Model: gpt-5.1-chat-latest

Created at: 1766049400

Finish reason: stop

Usage: 491 tokens in, and 379 tokens out

```
{  
    area := length * width;  
    perimeter := 2.0 * (length + width);  
    diagonal :| diagonal * diagonal == length * length + width * width;  
}
```

Program → GenAI

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

problem_data/2_8_rectangle_properties/program.dfy(8,11): 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.

```
|  
8 | diagonal :| diagonal * diagonal == length * length + width * width;  
| ^~
```

Dafny program verifier finished with 0 verified, 1 error

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-Co4PneNWyV2XiBN4rnn7c0XBjJDR3

Model: gpt-5.1-chat-latest

Created at: 1766049407

Finish reason: stop

Usage: 809 tokens in, and 257 tokens out

```
{
    area := length * width;
    perimeter := 2.0 * (length + width);
    diagonal :| {:trigger diagonal * diagonal} diagonal * diagonal == length * length +
        ↪ width * width;
}
```

Program → GenAI

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

```
problem_data/2_8_rectangle_properties/program.dfy(8,14): Error: invalid UnaryExpression
|
8 | diagonal :| {:trigger diagonal * diagonal} diagonal * diagonal == length * length +
    ↪ width * width;
| ^
|
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-Co4PrUr0b2UFU1fWFD8rl76cnQcGH

Model: gpt-5.1-chat-latest

Created at: 1766049411

Finish reason: stop

Usage: 937 tokens in, and 129 tokens out

```
{
    area := length * width;
    perimeter := 2.0 * (length + width);
    diagonal :| diagonal * diagonal == length * length + width * width {:trigger diagonal *
        ↪ diagonal};
}
```

Program → GenAI

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

```
problem_data/2_8_rectangle_properties/program.dfy(8,69): Error: invalid AssignStatement
|
8 | diagonal :| diagonal * diagonal == length * length + width * width {:trigger diagonal *
    ↪ * diagonal};
| ^
```

```
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-Co4Ptu41LLbNTp4Nt7eBBu9MBLDZ4

Model: gpt-5.1-chat-latest

Created at: 1766049413

Finish reason: stop

Usage: 1126 tokens in, and 253 tokens out

```
{  
    area := length * width;  
    perimeter := 2.0 * (length + width);  
    diagonal := (length * length + width * width) ** 0.5;  
}
```

Program → GenAI

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

```
problem_data/2_8_rectangle_properties/program.dfy(8,49): Error: invalid UnaryExpression  
|  
8 | diagonal := (length * length + width * width) ** 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-Co4Px5daXt15t5quBYsoUZICgLbMv

Model: gpt-5.1-chat-latest

Created at: 1766049417

Finish reason: stop

Usage: 1306 tokens in, and 196 tokens out

```
{  
    area := length * width;  
    perimeter := 2.0 * (length + width);  
    diagonal :| (diagonal * diagonal == length * length + width * width) {:trigger diagonal  
    ↳ * diagonal};  
}
```

Program → GenAI

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

```
problem_data/2_8_rectangle_properties/program.dfy(8,71): Error: invalid AssignStatement
|
8 | diagonal :| (diagonal * diagonal == length * length + width * width) {:trigger
  |   ↪ diagonal * diagonal};
  |
 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-Co4Q1MG6RAG6jeWTf6vnZZ0VcdntC

Model: gpt-5.1-chat-latest

Created at: 1766049421

Finish reason: stop

Usage: 1508 tokens in, and 133 tokens out

```
{
    area := length * width;
    perimeter := 2.0 * (length + width);
    diagonal :| {:trigger (diagonal * diagonal)} (diagonal * diagonal == length * length +
        ↪ width * width);
}
```

Program → GenAI

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

```
problem_data/2_8_rectangle_properties/program.dfy(8,14): Error: invalid UnaryExpression
|
8 | diagonal :| {:trigger (diagonal * diagonal)} (diagonal * diagonal == length * length
  |   ↪ + width * width);
  |
 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-Co4Q3WKcrENnJIyBEhjRMT0wI7Z6w

Model: gpt-5.1-chat-latest

Created at: 1766049423

Finish reason: stop

Usage: 1709 tokens in, and 257 tokens out

```
{  
    area := length * width;  
    perimeter := 2.0 * (length + width);  
    diagonal {::trigger diagonal * diagonal} ::| diagonal * diagonal == length * length +  
        ↪ width * width;  
}
```

Final Program

```
method rectangle_properties(length: real, width: real) returns (area: real, perimeter:  
    ↪ real, diagonal: real)  
    ensures area == length * width  
    ensures perimeter == 2.0 * (length + width)  
    ensures diagonal * diagonal == length * length + width * width  
{  
    area := length * width;  
    perimeter := 2.0 * (length + width);  
    diagonal {::trigger diagonal * diagonal} ::| diagonal * diagonal == length * length +  
        ↪ width * width;  
}
```

Total Token Usage

Input tokens: 8248

Output tokens: 1723

Reasoning tokens: 1216

Sum of ‘total tokens’: 9971

Experiment Timings

Iteration #8 started at 1766049425635, ended at 1766049429361, lasting 3726ms (3.73 seconds)
Overall Experiment started at 1766049399676, ended at 1766049429361, lasting 29685ms (29.69 seconds)
Iteration #4 started at 1766049412784, ended at 1766049415694, lasting 2910ms (2.91 seconds)
Iteration #5 started at 1766049415694, ended at 1766049419728, lasting 4034ms (4.03 seconds)
Iteration #6 started at 1766049419728, ended at 1766049423008, lasting 3280ms (3.28 seconds)
Iteration #7 started at 1766049423008, ended at 1766049425635, lasting 2627ms (2.63 seconds)
Iteration #1 started at 1766049399676, ended at 1766049402482, lasting 2806ms (2.81 seconds)
Iteration #2 started at 1766049402482, ended at 1766049408686, lasting 6204ms (6.20 seconds)
Iteration #3 started at 1766049408686, ended at 1766049412784, lasting 4098ms (4.10 seconds)