

Experiment ‘p4_3_a_extract_uppercase’ Results

December 18, 2025

Experiment outcome: SUCCESS

Bad responses: 0

Responses containing assume : 0

Resolution attempts: 2

Hard fails (resolution): 1

Soft fails (resolution): 0

Verification attempts: 1

Problem Specification

Problem name: p4_3_a_extract_uppercase

Natural language statement: Write a method that takes a string and returns only the uppercase letters in the string.

Method signature: p4_3_a_extract_uppercase(s: string) returns (uppercase: string)

Ensures

- forall c :: c in uppercase ==> 'A' <= c <= 'Z'
- forall c :: c in s && 'A' <= c <= 'Z' ==> c in uppercase

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 only the uppercase letters in the string.

The signature should be:

method p4_3_a_extract_uppercase(s: string) returns (uppercase: string)

The method should respect the following contract:

ensures forall c :: c in uppercase ==> 'A' <= c <= 'Z', ensures forall c :: c in s && 'A' <= c <= 'Z' ==> c in uppercase

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-Co4cIu7pw8p5W55YUb5VBnYsXZj8p

Model: gpt-5.1-chat-latest

Created at: 1766050182

Finish reason: stop

Usage: 320 tokens in, and 283 tokens out

```
{
  var res := "";
  var i := 0;
  while i < s.Length
    invariant 0 <= i <= s.Length
    invariant forall c :: c in res ==> 'A' <= c <= 'Z'
    invariant forall j :: 0 <= j < i && 'A' <= s[j] <= 'Z' ==> s[j] in res
  {
    if 'A' <= s[i] <= 'Z' {
      res := res + s[i..i+1];
    }
    i := i + 1;
  }
  uppercase := res;
}
```

Program → GenAI

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

```
problem_data/p4_3_a_extract_uppercase/program.dfy(7,14): Error: type seq<char> does not
  ↳ have a member Length
```

```
7 | while i < s.Length
  | ~~~~~
```

```
problem_data/p4_3_a_extract_uppercase/program.dfy(8,26): Error: type seq<char> does not
  ↳ have a member Length
```

```
8 | 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-Co4cM9F2dvi7tWJgbJUsyyMJc1bJB

Model: gpt-5.1-chat-latest

Created at: 1766050186

Finish reason: stop

Usage: 678 tokens in, and 157 tokens out

```
{
  var res := "";
  var i := 0;
  while i < |s|
    invariant 0 <= i <= |s|
    invariant forall c :: c in res ==> 'A' <= c <= 'Z'
    invariant forall j :: 0 <= j < i && 'A' <= s[j] <= 'Z' ==> s[j] in res
  {
    if 'A' <= s[i] <= 'Z' {
      res := res + s[i..i+1];
    }
    i := i + 1;
  }
  uppercase := res;
}
```

Final Program

```
method p4_3_a_extract_uppercase(s: string) returns (uppercase: string)
  ensures forall c :: c in uppercase ==> 'A' <= c <= 'Z'
  ensures forall c :: c in s && 'A' <= c <= 'Z' ==> c in uppercase
{
```

```
var res := "";
var i := 0;
while i < |s|
  invariant 0 <= i <= |s|
  invariant forall c :: c in res ==> 'A' <= c <= 'Z'
  invariant forall j :: 0 <= j < i && 'A' <= s[j] <= 'Z' ==> s[j] in res
{
  if 'A' <= s[i] <= 'Z' {
    res := res + s[i..i+1];
  }
  i := i + 1;
}
uppercase := res;
}
```

Total Token Usage

Input tokens: 998

Output tokens: 440

Reasoning tokens: 128

Sum of ‘total tokens’: 1438

Experiment Timings

Overall Experiment started at 1766050183868, ended at 1766050192462, lasting 8594ms (8.59 seconds)

Iteration #1 started at 1766050183868, ended at 1766050188030, lasting 4162ms (4.16 seconds)

Iteration #2 started at 1766050188030, ended at 1766050192462, lasting 4432ms (4.43 seconds)