

Experiment ‘p6_9_arrays_equal’ Results

December 18, 2025

Experiment outcome: SUCCESS

Bad responses: 0

Responses containing assume : 0

Resolution attempts: 1

Hard fails (resolution): 0

Soft fails (resolution): 0

Verification attempts: 1

Problem Specification

Problem name: p6_9_arrays_equal

Natural language statement: Write a method that checks whether two arrays have the same elements in the same order.

Method signature: p6_9_arrays_equal(arr1: seq<int>, arr2: seq<int>) returns (areEqual: bool)

Ensures

- areEqual == (|arr1| == |arr2| && forall i :: 0 <= i < |arr1| ==> arr1[i] == arr2[i])

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 checks whether two arrays have the same elements in the same order.

The signature should be:

```
method p6_9_arrays_equal(arr1: seq<int>, arr2: seq<int>) returns (areEqual: bool)
```

The method should respect the following contract:

```
ensures areEqual == (|arr1| == |arr2| && forall i :: 0 <= i < |arr1| ==> arr1[i] == arr2[  
  ↪ i])
```

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-Co5BjviQ2pUicWKK8pLuXnSq1fexS

Model: gpt-5.1-chat-latest

Created at: 1766052379

Finish reason: stop

Usage: 330 tokens in, and 141 tokens out

```
{  
  if |arr1| != |arr2| {  
    areEqual := false;  
  } else {  
    areEqual := (forall i :: 0 <= i < |arr1| ==> arr1[i] == arr2[i]);  
  }  
}
```

Final Program

```
method p6_9_arrays_equal(arr1: seq<int>, arr2: seq<int>) returns (areEqual: bool)  
  ensures areEqual == (|arr1| == |arr2| && forall i :: 0 <= i < |arr1| ==> arr1[i]  
    ↪ == arr2[i])  
{  
  if |arr1| != |arr2| {  
    areEqual := false;  
  } else {  
    areEqual := (forall i :: 0 <= i < |arr1| ==> arr1[i] == arr2[i]);  
  }  
}
```

```
}  
}
```

Total Token Usage

Input tokens: 330

Output tokens: 141

Reasoning tokens: 64

Sum of ‘total tokens’: 471

Experiment Timings

Overall Experiment started at 1766052381401, ended at 1766052385262, lasting 3861ms (3.86 seconds)

Iteration #1 started at 1766052381402, ended at 1766052385262, lasting 3860ms (3.86 seconds)