Program Verification

Assignment 2

Niek Haarman, Tanja de Jong, Tinus Pool

July 3, 2015

For this assignment we specified and verified a sequential and concurrent `Queue` implementation. Below is a report about our findings.

1 Sequential Queue implementation

For the sequential Queue, we decided to go with a `LinkedList` implementation, which can be seen as a FIFO queue where items are appended at the tail, and taken from the head of the queue. We took the Java implementation of `LinkedList` and stripped it down to contain only code relevant to the exercise, and created a PVL implementation based on that.

The Java implementation keeps a reference to the first and last `Node` of the list as fields. This allows for easy appending and extraction of values. We chose to create a `contents()` function which describes the current contents of the list as a `seq<int>`. That way we can easily ensure items are added and removed properly. To do that, we declared a recursive `contents()` function on `Node`. This approach requires that the first `Node` has (recursively) at least a read permission on the `val` and `next` fields of all subsequent `Nodes`, as seen in Listing **??**. This immediately leads to a problem, since we cannot easily obtain a write permission to append an item to the queue: either we have a full write permissions on the contents of the `LinkedList`'s `last` field, or we have recursive read permissions for all `Nodes`. To solve this, we dropped the `last` field as a whole, and gave full recursive *write* permission to the `Nodes`. Appending an item now traverses the entire list, starting from the head.

While working at the sequential exercises we did want to hold the permission stick on both ends. The problem however is that you can only release one end of the stick at a time, while you have to release both ends to make the stick float in the air. Only when the stick floats in the air you can adjust it.

This is a problem we encountered in our first try with the queue. We had a node last and first and the resource state required the next state of the node. This way we created a chain from the first till the last. So if you wanted to edit the last you where not allowed because last permission was also locked by the chain created by implied by the first state.

```
class Node {
  Node next;
  int val;

  resource state() = Value(val) ** Value(next) ** next->state();

  requires state();
  seq<int> contents() =
  unfolding state() in (
    next == null
     ? seq<int>{val}
     : seq<int>{val} + next.contents()
  );
}
```

Listing 1: Basic Node specification

This was easily solved by just removing the last, and following the chain till the end to change the last element of the queue. An other possible option was creating all the nodes in a factory and watch the permissions in this factory. However because this concept had some troubles implementing at first and the other implementation was faster implemented, we decided to not use a factory but remove the last node. We still could argue what solution would be better because they both have there nice things and ugly things. Another reason why the factory was not added in the end is because Vercors can be a bit fragile in times. We did not want to burn our fingers when there is already something workable.

2 Concurrent Queue implementation

For the concurrent Queue, the assignment specified to use the LinkedBlockingQueue. Again, we took the Java implementation and created a PVL implementation of the required methods.

A big difference between the implementation of LinkedBlockedQueue and LinkedList, which we used for the sequential queue, is that the constructor of the LinkedBlockingQueue constructs a `null` Node, so that the queue is never actually empty, while the LinkedList simply starts with an empty list. At our first try, we did not understand why this approach was taken, so we decided to implement this part similarly to the LinkedList, by constructing the Queue without any nodes. However, during the process of verifying the LinkedBlockingQueue, we realized that this resulted in a problem for a concurrent queue if one thread attempts to put a value in the queue, while an another thread tries to read a value from that queue at the same time, when there is exactly one element in the queue.

The main problem we encountered while verifying the concurrent implementation was the following: How can you build in two locks in a queue that refer to two different parts of the queue? This problem is created because of the restriction that the queue can contain only one lock_invariant. At first, we tried to solve this problem by using kind of symbolic_locks. This way, nothing was actually locked, but because the implementation would be consistent in the usage and basically have barrier against violating what we would actually wanting to lock, it was a good start and in our opinion valid for the moment. Using this approach, one of the locks prevents that elements are taken when the queue is empty, while the other locks prevents that elements are added when the queue has reached its maximum capacity.

In this way, we were able to write all the method implementations. However, for inserting and removing objects, we used the same method as we used in the sequential implementation, which required to have full permission over the queue and it's node. So that implementation is like one stick or chain that is only allowed to be held by one hand. This is a problem for concurrent programs, where you have to be able to hold the same side of the stick/chain and be able to hold it with multiple hands. So, this implementation did not allow for concurrent access to the queue.

Luckily for us, the lock_invariant is something magical that spawns invisible hand and holds up parts of the chain. So, instead of creating the chain by saying that you also want to ensure the next state, you can now only know the next, and the lock_invariant of that node will know its own state. The lock_invariant will preserve the state, so you don't have to save that state by having to remember the next node's state. This enabled us to introduce the head and last node again. The queue became a thing where we only want to ensure that it has a value. We don't need write permission of the queue itself to change elements. We changed the symbolic locks to elements that do actually know the queue instance. One of the locks will lock the first element of the queue, which prevents other threads from removing elements from the queue. The other lock will lock the last element of the queue, which prevents other threads from adding elements to the queue.

3 Other problems encountered:

- Cannot test with java, because java doesn't recognize seq¡..¿ as data type.
- Need a `lockAndUnlock()` action for some methods, to create a kind of atomic `get()`. In our atomic integer we where not allowed to use our `get()` in verifying our code, because a normal method might change code.
- Have to catch return values. Making this a habit makes people doubt your developer skills.
- Error messages for not catching return values and using reserved variables can be confusing and hard to discover if you write a lot of code in a flow.
- Assigning null to a return or variable was troubling.

It is possible that along the way some of the problems mentioned above where solved.

4 Things we still could do to verify concurrent code?

We could create local data storage objects and store snapshots of situations in this local data storage. Based on what is inside those snapshots you could argue what was happening inside the concurrent code and based on those things what the end result could be. However because the border between what is used to verify code in PVL and the actual code is already fragile, we didn't want to push the developer by giving an extra object to a method what only would be used to verify the code instead of helping to get the actual result. Maybe it would be nice to create a subtle border between what is used to verify code and the actual functionality of the code. Instead of a history, you could give a local box that would help reason about what happened inside the code. History would probably be easier to use and less complex to evaluate what would be the end result.

5 Possibilties for adding functional properties

In the current specification, it is not possible to ensure the state of the queue after any method, because no assumptions can be made about other threads that might have changed the queue after the lock has been released. For example, after the `take()` method is called, it can not be ensured that the contents of the queue are equal to the contents of the queue, before the method was called, without the first element.

However, it is possible to use so-called *history* to be able to ensure more properties of the methods. By doing this, the history keeps track of all elements that have been added to or removed from the queue. For example, if we would add this specification to the `put()` and `take()` methods, we will know that a value has been respectively added to or removed from the queue. Then, if multiple threads are simultaneously calling the `put` method, we can use the history to ensure that both elements have been added to the queue, however, we cannot ensure the order in which this was done. We can draw the same conclusion when multiple threads call the `take()` method, but instead of ensuring that elements were added, we can ensure that these exact elements were taken from the queue.

So, by added information about the history of the queue, we can draw more conclusions about the contents of a concurrent queue after an element has been added to or removed from that queue.

6 Work distribution

Because the bulk of the work revolved around figuring out a correct approach for the verification of the sequential and concurrent queues, the main activity for this assignment was discussing the possibilities. Therefore, nearly everything was done as a group and there was not actually much distribution of work.