This is a closed book test. You may not refer to any resources.

This is a 50 minute test.

Please write your answers in ink. Pencil answers will be marked, but will not be re-marked under any circumstances.

The test will be marked out of 50.

<table>
<thead>
<tr>
<th>Question</th>
<th>/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Question 1</td>
<td>12</td>
</tr>
<tr>
<td>Question 2</td>
<td>12</td>
</tr>
<tr>
<td>Question 3</td>
<td>12</td>
</tr>
<tr>
<td>Question 4</td>
<td>/2</td>
</tr>
<tr>
<td>Question 5</td>
<td>12</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>50</td>
</tr>
</tbody>
</table>
Question 1 (12 marks)

In Red-Black trees, one of the colouring rules is that a Red vertex cannot have a Red child. What is the purpose of this rule?

Solution:

*This rule has the result that on every path from the root to a leaf has at least as many Black vertices as Red vertices.* Combined with the rule that every path from the root to a leaf must have the same number of Black vertices, this guarantees that the longest path from a root to a leaf is no more than twice as long as the shortest path.

Marking:

*If the student understands that this rule means that every path must have at least as many Blacks as Reds (another way of saying this is that the Reds can at most alternate with the Blacks), they should get at least 6/12.*

*If the student understands that this is part of the method of guaranteeing that the longest path will be no more than twice the length of the shortest path, they should get at least 9/12.*

*If they mention the rule that requires an equal number of Black vertices on all paths, they should get full marks.*
Question 2 (12 marks):

The last step of the insertion algorithm for Red-Black Trees is “Colour the root Black”.

a) [6 marks] Explain why the colour of the root vertex of a Red-Black tree is actually irrelevant for ensuring that the “black-height” balance is maintained.

Solution:

Since the root is on every path from the root to a leaf (obviously!) it will contribute equally to the number of Black vertices on each path whether it is Red or Black.

Marking:

If the student states that all paths must have equal numbers of Black vertices, they should get at least 3/6, even if they don’t state that the root contributes equally to all the paths.

b) [6 marks] Explain why this step is included in the algorithm.

Solution:

During the balancing process we look at the parent of the current vertex, and if the parent is Red we also look at the grandparent. The fact that the root cannot be Red means that if the parent is Red, it cannot be the root ... so the grandparent must exist. This simplifies the coding of the algorithm because in guarantees we cannot have a rebalancing situation in which the grandparent does not exist.

Marking:
If the student says it simplifies the implementation, they should get at least 3/6. If they understand that it prevents any Red-Red situation in which the parent is the root, they should get full marks. If they can’t quite express why it simplifies the implementation, they should get 4 or 5.
Question 3 (12 marks):

a) [6 marks] Prove that if the root of a Red-Black tree has two Red children, we could colour them both Black without affecting the validity of the colouring.

Solution:

First, observe that colouring the vertices Black cannot violate the colouring rule, since no Red-Red situation could be created. Second, the number of Black vertices on all paths from the root down to the leaves will increase by 1, so they will all still be equal.

Marking:

If the student only addresses one of the two points (ie only the Red-Red rule, or only the balance rule) they should get 4/6

b) [6 marks] Show that if the root of a Red-Black tree has two Black children, it is not necessarily true that we can colour them both Red without affecting the validity of the colouring.

Solution:

One or both of the vertices could have a Red child, in which case colouring the vertex Red would introduce a Red-Red violation of the rules for Red-Black trees.

Marking:
If they can’t identify the reason but they at least understand the question, they should get 3/6

Question 4 (2 marks):

If a Red-Black tree is constructed using the insertion algorithm we have studied, and contains at least 2 values, then it always contains at least one Red vertex.

Choose one answer:

True
False

Solution: True

Marking:

True: 2 marks, False: 0 marks
Question 4 (12 marks):

Here is an insert function for a hash table:

```python
def insert(k):
    a = h(k)
    i = 0
    while i < m:
        p = (a + 2*i) % m
        if T[p] == 'empty' or T[p] == 'deleted':
            T[p] = k
            return "success"
        else:
            i += 1
    return "fail"
```

This is simply linear probing using a step size of 2 instead of 1. You may assume that m (the size of the hash table) is odd.

Does this reduce primary clustering, when compared to standard linear probing? (Hint: what do the probe sequences look like?)

Solution:

No, it doesn’t reduce primary clustering. Consider the probe sequence that starts at 0: 0,2,4,6,8,...,m-1,1,3,5,...,m

Now consider the probe sequence that starts at 2: 2,4,6,8,...,m-1,1,3,5,...,m,0

And consider the probe sequence that starts at 1: 1,3,5,...,m,0,2,4,6,...,m-1

And the probe sequence that starts at 3: 3,5,7,...,m,0,2,4,...,m-1,1

From these we can see that the probe sequences overlap in exactly the same way as the sequences in standard linear probing do: once two probe sequences collide, they match exactly from that point onward. So for example if addresses
2, 4, 6 and 8 are all filled, address 10 has a greatly increased probability of being filled next. Sequences of filled addresses will expand in just the same way as with standard linear probing, with the sole difference being that each address in the sequence will be 2 more than the previous one, rather than 1 more.

Marking:

If the student clearly understands what primary clustering is, they should get at least 7/12

If the student suggests that there are two probe sequences (one consisting of all even addresses, the other consisting of all odds) which would give a slight reduction in primary clustering, they should get at least 10/12

If they recognize that all the probe sequences are effectively identical, they should get full marks.
Student Number: __________________

Nearly Blank Page