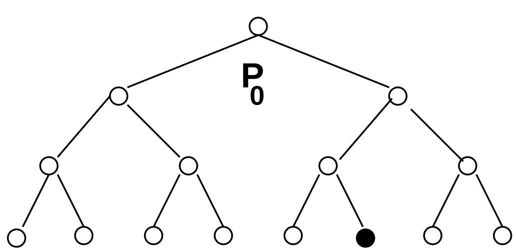
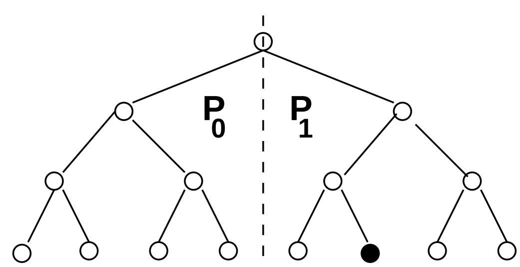
CSC 478 Winter 2010 exam 1. Due date: Thursday February 11, before class.

This exam is worth 10% of your course grade, graded on a 100-point scale (all questions are weighted equally). The exam should be submitted in a document through the file upload, emailed, or handed-in in person on the due date. Although a typed is preferred, (scanned) hand-written is also acceptable as long as it is legible.

1. Consider a parallel processor with 8 cores. You have a program that requires 20 minutes if it is executed with only one thread (i.e., sequentially). Otherwise, consider the following:
   (a) 50% of the program must remain sequential. What's the speedup and the efficiency if the remaining part is fully parallelizable on the 8-core machine?
   (b) 10% of the program must remain sequential. What's the speedup and the efficiency if the remaining part is fully parallelizable on the 8-core machine?
2. Identify and describe the four types of PRAM.
3. (Superlinear speedup) Consider the search tree shown in the figure below, in which the dark node represents the solution.
   (a) If a sequential search of the tree is performed using a standard depth-first search (DFS) algorithm (going left-to-right), then how much time does it take to find the solution if traversing each arc of the tree takes one unit of time? (Use the tree on the left with one processor.)
   (b) Assume that the tree is partitioned between two processing elements that are assigned to do the search job as shown in the figure on the right. If both processing elements perform a DFS on their respective halves of the tree, then how much time does it take for the solution to be found? What is the speedup?





4. Consider the following sequential code segment:

```
S1:  x = (b - e) (e + c)
S2:  a = 2d (d + c)
S3:  z = z (x + a)
S4:  c = e (f - b)
S5:  y = a + 2f - b
```

   (a) Identify all of the data dependencies and what type they are.

   (b) To the extent that it is possible, show how the code can be parallelized if an unlimited number of necessary arithmetic units (e.g. multiply, add, etc.) are available.

5. Assume Dr. Turner's machine's GPU can execute at a rate of 700 GFLOPS in single-precision, 70 GFLOPS in double-precision, and that the CPU can execute (single-core) at a rate of 16 GFLOPS.

   (a) What's the average execution rate of a program under the following conditions? The single-core CPU occupies 40% of the work. Single-precision GPU work is an additional 40%. Double precision is thus 20%.

   (b) Given the conditions of part (a), how much time will a program requiring 1 trillion operations take, if we assume they're all floating-point?

6. On a machine having no maximum operation, the computation of an N element vector C whose elements are the maxima of corresponding elements of vectors A and B can be done by comparisons:

```
for i := 0 step 1 until N-1
  if A[i] > B[i] then C[i] := A[i]
  else C[i] := B[i];
```

   Using vector pseudo code, show how to do this operation on an SIMD machine with N PEs.

7. Explain the meaning of the following terms related to thread programming:

   (a) Thread

   (b) race condition

   (c) mutex

   (d) lock contention

   (e) granularity

   (f) false sharing; additionally, what is a practical solution to this?

8. Explain how parallel summation (Figure 2-2) can be changed to find the maximum element of an array.