The Theory of Computation
Bernard M.E. Moret
Addison-Wesley-Longman, 1998

Solution Manual
Chapter 7, Problem 7.24

The decision problem is clearly in NP; we prove that it is NP-complete by transforming *X3C* to it. The identification problem will have an object for each element in the cover problem, call them $\{x_1, \ldots, x_n\}$, plus 3 extra elements, u, v, w; each subset of the cover problem will become a test in the identification problem; in addition, we'll have a test of the form $\{x_i\}$ for each x_i and tests $\{u\}$ and $\{v\}$. If the exact cover problem has a solution, we can construct an identification tree that successively uses all 3-sets of the solution to partition the objects into $n+1$ subsets of 3 objects each; the objects in each subset are then resolved by a simple subtree of 3 leaves (using, for instance, the $\{x_i\}$ tests). The subtree of 3 leaves has external path length of 5; the identification tree would have total external path length of $5 + 3 \cdot (n-1) + \sum_{i=1}^{n} 5 + 3i = 5n + 3 \cdot (n^2 + n - 3)/2 = 3 \cdot (n^2 + 4n - 3)/2$. Now we let the bound be this value (which we will claim is optimal). Clearly, if *X3C* has a solution, so does our decision problem; conversely, we claim that, if our decision problem has a solution, its cost is precisely equal to the bound and the tree constructed is of the exact description given above (in which case we have a solution for *X3C*). That our claim is true is intuitively obvious: no test can do better than distinguish a 3-subset from the remaining objects, and that can be done to the end only if all 3-subsets are disjoint; a formal proof can also be given, for which see the original paper by Hyafil and Rivest (*Inf. Proc. Letters* 5 (1976), 15–16).