Read Chapters 7.1–7.6, and 12.4 from the textbook, as well as the paper by Fakcharoenphol, Rao, and Talwar (I recommend the journal version), and optionally the two papers by Bartal.

Since this is the first time we have problems assigned from the textbook, and some of those problems point to papers in the literature where they are solved, I just want to clarify that it is of course not appropriate to seek out these references (or any summaries of them) as hints for solutions.

(1) Let \( R \) be a range space of VC-dimension \( d \). Define \( R' := \{ A \cap B \mid A, B \in R \} \) to be the set of all pairwise intersections of sets in \( R \). Prove that \( R' \) has VC-dimension \( O(d) \).

Hint: Use some ideas from our proof of the \( \epsilon \)-net theorem, including the Sauer-Shelah Lemma. Recall that the lemma states that if \( R \) is a range space of VC-dimension \( d \) on a ground set of size \( n \), then \( \| R \| \leq \sum_{k=0}^{d} \binom{n}{k} \). Then apply suitable bounds based on Stirling’s approximation. (See, e.g., Appendix B of the textbook.)

If you do not manage to prove an \( O(d) \) bound, you will get significant partial credit for proving an \( O(d \log d) \) bound.

(2) Problem 7.2 from the textbook.

(3) Problem 12.24 from the textbook. (Hint: It might help you to first solve the problem for forests instead of general graphs. Among other things, you will get significant partial credit for doing so.)

(4) Recall that the key step in the FRT construction was the (randomized) construction of clusters of fairly low diameter, such that node pairs that were close were unlikely to end up in separate clusters. Here, you will show stronger bounds for this subproblem when the points lie in a low-dimensional Euclidean space. You are given points in \( \mathbb{R}^d \), where \( d \) is a constant. Your goal is to randomly partition them into clusters of diameter at most \( \Delta \) such that the probability of \( x \) and \( y \) being separated is bounded by \( O(||x - y||_2/\Delta) \).

(a) For \( d = 1 \) (i.e., all points are on a line), give and analyze a randomized partitioning algorithm that always produces clusters of diameter at most \( \Delta \) such that for any pair \( x, y \), the probability of \( x \) and \( y \) being separated is at most \( |x - y|/\Delta \). (So here, the constant is 1.)

(b) For arbitrary constant \( d \), give and analyze a randomized partitioning algorithm that always produces clusters of diameter at most \( \Delta \), such that for any pair \( x, y \), the probability of \( x \) and \( y \) being separated is at most \( f(d) \cdot ||x - y||_2/\Delta \). Here, \( f(d) \) can be any function of your choice, but it can depend only on the dimension \( d \).