St. Francis Xavier University
Department of Computer Science

CSCI 435: Algorithms and Complexity
Week 10 Discussion Questions
Winter 2023

[Joh74] David S. Johnson. Approximation Algorithms for Combinatorial Problems. *Journal of Computer and System Sciences*, 9:256–278, 1974.

David Johnson was the head of the Algorithms and Optimization Department at AT&T Labs, and a visiting professor at Columbia University. He was well-known for his work in algorithm design and analysis, and he was one of the co-authors of the famous book *Computers and Intractability: A Guide to the Theory of NP-Completeness*. Johnson died in 2016.

Read Sections 1 and 2, then skim through Section 5. Then, consider the following questions.

1. What is the class of "polynomial complete" problems? How does this relate to the notion of NP-completeness?
2. Johnson gives two definitions of the "set covering" problem in this paper. How do these problems relate, and how do they differ? Which of the two definitions is closest to the problem we discussed in lecture?
3. In the definition of an optimization problem, why is the measure m_P needed? Later, following the definition of an approximation algorithm, what does the ratio $r_P(A, u)$ describe?
4. What is the approximation performance guarantee obtained in this paper for the set cover problem? How does it differ from the performance guarantees we obtained in lecture?
5. Johnson gives an example of how to reduce between two optimization problems by introducing the "node cover" problem. What is the definition of this problem, and how does it relate to the set cover problem?

[Chv79] Václav Chvátal. A Greedy Heuristic for the Set-Covering Problem. *Mathematics of Operations Research*, 4(3):233–235, 1979.

Václav Chvátal is a professor emeritus at Concordia University in Montréal. His work centers on combinatorics, combinatorial optimization, and graph theory. He is particularly well known for his work on approximation algorithms and linear programming.

Read the entire paper. Then, consider the following questions.

1. How does the definition of the set covering problem in this paper differ from that in the previous paper? How can we reduce the problem in this paper to Johnson's version of the problem?
2. What is the heuristic Chvátal proposes for selecting which indices j to add to our cover? Is the proposed algorithm truly "recursive", or is it more iterative?
3. Midway through the proof of the main theorem, there is a discussion on the values y_i, which denote the "price paid by the greedy heuristic for covering the point i". How does this relate to or differ from our discussion of the dual method in lecture?
4. Ultimately, the approximation performance guarantee of the given algorithm is $H(d) = \sum_{j=1}^{d} 1/j$. These are known as the *harmonic numbers*. What is the asymptotic growth rate of the harmonic numbers?