

## Problem Set 8

*Harvard SEAS - Fall 2025**Due: Wed Nov. 12, 2025 (11:59pm)*

Please see the syllabus for the full collaboration and generative AI policy, as well as information on grading, late days, and revisions.

All sources of ideas, including (but not restricted to) any collaborators, AI tools, people outside of the course, websites, ARC tutors, and textbooks other than Hesterberg–Vadhan must be listed on your submitted homework along with a brief description of how they influenced your work. You need not cite core course resources, which are lectures, the Hesterberg–Vadhan textbook, sections, SREs, problem sets and solutions sets from earlier in the semester. If you use any concepts, terminology, or problem-solving approaches not covered in the course material by that point in the semester, you must describe the source of that idea. If you credit an AI tool for a particular idea, then you should also provide a primary source that corroborates it. Github Copilot and similar tools should be turned off when working on programming assignments.

If you did not have any collaborators or external resources, please write 'none.' Please remember to select pages when you submit on Gradescope. A problem set on the border between two letter grades cannot be rounded up if pages are not selected.

**Your name:**

**Collaborators and External Resources:**

**No. of late days used on previous psets:**

**No. of late days used after including this pset:**

The purpose of this problem set is to develop skills in implementing graph algorithms, appreciate the impact of different kinds of worst-case exponential algorithms in practice, and practice reducing problems to SAT.

1. (Another coloring algorithm) In the Github repository for PS8, we have given you basic data structures for graphs (in adjacency list representation) and colorings, an implementation of the coloring algorithm from ps6, and a variety of test cases (graphs) for coloring algorithms.
  - (a) Implement the reduction from 3-coloring to SAT given in class in the function `sat_3_coloring`, producing an input that can be fed into the SAT Solver [Glucose](#), and verify its correctness by running `python3 -m ps8_tests 3`.
  - (b) Compare the efficiency of Exhaustive-Search 3-coloring, the  $O(1.45^n)$ -time independent-set and BFS algorithm for 3-coloring from problem set 6 (feel free to use the staff solution or your own implementations from problem set 6), and your implementation from Part [1a](#) using `ps8_experiments`. In the experiments file, we've provided code to generate two types of graphs (lines of rings and clusters of independent sets) and some new hard graph instances. For each of those types of graphs, how many of the given instances, if any, can each algorithm solve within 1 second? You should fill out the table and briefly discuss your findings, as well as why these algorithms perform the way they do given what we have discussed in lecture.

Algorithm	Exhaustive	ISet BFS	SAT Color
# Solvable Ring Instances			
# Solvable Cluster Instances			
# Solvable Hard Graphs			

2. (Resolution) Use the algorithm **ResolutionInOrder** (algorithm 18.2 from the textbook) to decide the satisfiability of the following formulas, and use the algorithm **ExtractAssignment** to obtain a satisfying assignment for any that are satisfiable. (Please make sure to follow both algorithms *exactly*, including the order in which the clauses are processed. A correct final solution that does not show all of the intermediate steps of both algorithms will not receive full score.)

(a)  $\varphi(x_0, x_1, x_2, x_3) = (x_0) \wedge (x_2 \vee x_3) \wedge (x_1 \vee x_2) \wedge (x_1 \vee x_0)$

(b)  $\varphi(x_0, x_1, x_2, x_3) = (\neg x_1) \wedge (x_3) \wedge (\neg x_2 \vee \neg x_3) \wedge (x_1 \vee x_2)$

(c)  $\varphi(x_0, x_1, x_2, x_3) = (\neg x_1 \vee \neg x_2) \wedge (x_3 \vee x_1) \wedge (\neg x_0 \vee x_3) \wedge (\neg x_3) \wedge (x_0 \vee \neg x_1)$

(d)  $\varphi(x_0, x_1, x_2, x_3, x_4) = (x_1 \vee \neg x_2 \vee x_3 \vee \neg x_4) \wedge (x_0 \vee x_4) \wedge (\neg x_2 \vee x_4) \wedge (x_3 \vee x_0) \wedge (\neg x_3 \vee x_4)$

3. (Reductions to SAT) In Section 15.3 of the textbook, we saw an efficient algorithm to solve the Maximum Matching problem in Bipartite Graphs. A variant of Maximum Matching is Maximum 3-D Matching, where we are given not a graph  $G = (V, E)$  but a “3-uniform hypergraph”  $H = (V, E)$ , where the *hyperedges*  $e \in E$  now are *triples* of vertices. Analogously to restricting to bipartite graphs, we restrict to *tripartite* hypergraphs where  $V$  is the union of 3 disjoint sets  $V_0 \cup V_1 \cup V_2$  and every hyperedge  $e \in E$  has exactly one vertex from each of  $V_0, V_1, V_2$ . For example, hyperedges may consist of compatible patient-donor-timeslot triples (if there are only certain timeslots in which a patient and donor are available for a surgery), or compatible surfer-surfboard-fins triples (if each surfer only likes to ride certain surfboards with certain fins installed). Now the goal is to find a maximum-sized set  $M \subseteq E$  such that every vertex  $v \in V$  is contained in at most one edge of  $M$ . Unfortunately, unlike matching in graphs, there is no polynomial-time algorithm known for Maximum 3-D Matching, even if we restrict to tripartite hypergraphs and even if we only want to find *complete matchings* — ones that match every vertex in  $V_0$  (i.e. we are looking for a matching of size  $|M| = |V_0|$ ):

**Input:** Three disjoint finite sets of vertices  $V_0, V_1, V_2$  and a set  $E$  of hyperedges such that each  $e \in E$  contains exactly one vertex from each of  $V_0, V_1$ , and  $V_2$ .

**Output:** A set  $M$  of hyperedges such that every vertex is in at most one hyperedge in  $M$ , and every vertex in  $V_0$  is in a hyperedge in  $M$  (if one exists).

**Computational Problem** 3dCompleteMatching()

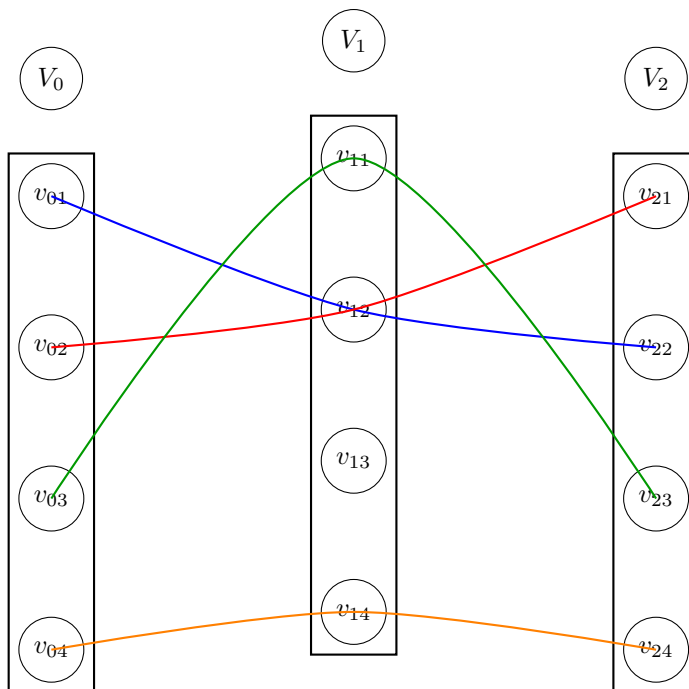


Figure 1: Example of hypergraph with sets of vertices

$V_0 = [v_{01}, v_{02}, v_{03}, v_{04}]$ ,  $V_1 = [v_{11}, v_{12}, v_{13}, v_{14}]$ ,  $V_2 = [v_{21}, v_{22}, v_{23}, v_{24}]$  and hyperedges

$$E = \{(v_{01}, v_{12}, v_{22}), (v_{02}, v_{12}, v_{21}), (v_{03}, v_{11}, v_{23}), (v_{04}, v_{14}, v_{24})\}$$

- (a) Show that there is a polynomial-time reduction of 3dCompleteMatching to SAT that, on hypergraphs with  $n = |V_0| + |V_1| + |V_2|$  vertices and  $m = |E|$  hyperedges, makes one oracle query on a formula with  $m$  variables and  $O(n + m^2)$  clauses, and runs in time  $O(n + m^2)$ . Be sure to prove the correctness of your algorithm and analyze its runtime. Optionally, justify why the reduction can be implemented in time  $O(m^2)$  (this can be part of an attempt at an R+ on this problem, but is not required for full credit). Thus, even though the fastest known algorithms for 3dCompleteMatching run in exponential time, we can use SAT Solvers to solve it much more efficiently on many instances that arise in practice.
  - (b) (optional<sup>1</sup>) Come up with a polynomial-time reduction for the version of 3dMatching where we are also given a number  $k \in \mathbb{N}$  as part of the input and want to find a matching of size  $k$  (rather than one of size  $|V_0|$ ). (Hint: you will probably want to use more than  $m$  boolean variables, at least  $k \cdot m$ , possibly more, depending on how you approach the problem.)
4. (reflection) **Active learning** involves engaging in activities that require you to think critically and apply your knowledge. Examples of active learning include reading the textbook carefully, asking questions, solving practice problems, and teaching concepts to others. In contrast, **passive learning** refers to taking in information without actively engaging with

<sup>1</sup>This problem won't make a difference between N, L, R-, and R grades. As this problem is purely extra credit, course staff will deprioritize questions about this problem at office hours and on Ed.

it—for instance, watching a lecture without participating, or listening to a collaborator or AI explain course concepts to you. Active learning has been shown to be more effective than passive learning for developing a deep understanding of course material.

In CS 1200, which active learning strategies do you use? Are there passive learning habits you tend to rely on? Reflect on your current approach: Are you satisfied with the amount of active learning you’re incorporating? Why or why not? If you think you could benefit from more active learning, what specific steps could you take to engage more actively in CS 1200?

Quick note on grading: Good responses are usually about a paragraph, with something like 7 or 8 sentences. Most importantly, please make sure your answer is specific to this class and your experiences in it. If your answer could have been edited lightly to apply to another class at Harvard, points will be taken off.

5. Once you’re done with this problem set, please fill out [this survey](#) so that we can gather students’ thoughts on the problem set, and the class in general. It’s not required, but we really appreciate all responses!