Introduktion til diskret matematik og algoritmer: Problem Set 3

Due: Wedneday March 12 at 12:59 CET.
Submission: Please submit your solutions via *Absalon* as a PDF file. State your name and e-mail address close to the top of the first page. Solutions should be written in LaTeX or some other math-aware typesetting system with reasonable margins on all sides (at least 2.5 cm). Please try to be precise and to the point in your solutions and refrain from vague statements. Never, ever just state the answer, but always make sure to explain your reasoning. *Write so that a fellow student of yours can read, understand, and verify your solutions.* In addition to what is stated below, the general rules for problem sets stated on *Absalon* always apply.
Collaboration: Discussions of ideas in groups of two to three people are allowed—and indeed, encouraged—but you should always write up your solutions completely on your own, from start to finish, and you should understand all aspects of them fully. It is not allowed to compose draft solutions together and then continue editing individually, or to share any text, formulas, or pseudocode. Also, no such material may be downloaded from or generated via the internet to be used in draft or final solutions. Submitted solutions will be checked for plagiarism.
Grading: A score of 120 points is guaranteed to be enough to pass this problem set.
Questions: Please do not hesitate to ask the instructor or TAs if any problem statement is unclear, but please make sure to send private messages—sometimes specific enough questions could give away the solution to your fellow students, and we want all of you to benefit from working on, and learning from, the problems. Good luck!

1 (40 p) Recall that a standard deck of cards has 52 cards partitioned into four *suits* (hearts, spades, clubs, and diamonds) with 13 ranks each (2–10 plus jack, queen, king, and ace). In this problem, we assume that you are dealt 5 cards from a perfectly shuffled deck of cards, and we wish to analyse the probability of getting flushes and/or straights. Note that there are two straights involving the ace: both ace–2–3–4–5 and 10–jack–queen–king–ace are valid straights.

1a What is the probability that you get a *flush*, i.e., 5 cards of the same suit but not all in sequence with respect to rank? (Because five cards of the same suit in sequential rank would be a *straight flush*.)

Solution: The number of flushes, straight or not, is obtained by first choosing a suit in 4 ways, and then choosing 5 cards in $\binom{13}{5}$ ways. After choosing the suit, the number of straight flushes is obtained by choosing the lowest card in the straight *ace*–10 in 10 possible ways. Hence, the total number of (non-straight) flushes is $4\left(\binom{13}{5} - 10\right)$. The total possible number of outcomes for a hand of 5 cards is $\binom{52}{5}$. Since all outcomes are equally likely, the probability of getting a (non-straight) flush is

$$\frac{4\left(\binom{13}{5} - 10\right)}{\binom{52}{5}} ,$$

i.e., the number of successful outcomes divided by the total number of outcomes.

1b What is the probability that you get a *straight*, i.e., 5 cards of sequential rank but not all of the same suit? (Because if the latter condition also held, we would again have a straight flush.)

Solution: The number of straights, flushes or not, is obtained by first choosing the rank of the lowest card 2–10 in 10 possible ways, and then for each of the 5 cards choosing the suit in a total of 4^5 ways. To get a straight flush, we can choose the suit of the first card in 4 ways, but then there is no freedom of choice for the rest of the cards. Hence, the total number of (non-flush) straights is $10\bigl(4^5 - 4\bigr)$. The total possible number of outcomes for a hand of 5 cards is still $\binom{52}{5}$, and so the probability of getting a (non-flush) straight is

$$\frac{10\bigl(4^5 - 4\bigr)}{\binom{52}{5}} \ .$$

2 (40 p) Prove mathematically that among all numbers on the form $11\ldots100\ldots0$, i.e., numbers consisting of m ones followed by n zeros for some $m, n \in \mathbb{N}^+$ (sometimes notation like 1^m0^n is used to describe text strings constructed in such a way), there is some number that is divisible by 2025.

Hint: Look at all numbers $1^m = 11\ldots1$ and consider what their remainders can be modulo 2025.

Solution: Using the notation introduced within parentheses above, among the first 2026 numbers on the form $11\ldots1$ there have to be two numbers 1^m and 1^n (i.e., with m and n digits 1, respectively) for $m > n$ with the same remainder modulo 2025. Here we are using the pigonhole principle with the numbers $11\ldots1$ serving as pigeons and the possible remainders modulo 2025 serving as holes. But that means that 2025 divides the difference of these two numbers, and subtracting the smaller number from the larger one we get a number on the form 1^m0^n that is divisible by 2025, as desired.

As a side note, if one would run computer experiments to find a large number on the form 1^m0^n that the computer claims is divisible by 2025, such a number cannot just be reported without any further evidence—there is no obvious reason why we should believe such a claim. In order to get a full score for such a solution, one would also need to state clearly the factorization and work out the details to show that it is correct. (Another matter is that it should be fairly clear that this is not how the problem was intended to be solved, and that problems on the exam will not be possible to solve in a similar way, so it is arguably a somewhat short-sighted approach to try to solve the problem in this way.)

3 (40 p) Let $a \in \mathbb{R}^+$ be any positive real number. Show that for any integer $n \geq 2$ there is a rational number $\frac{c}{d}$, $c, d \in \mathbb{Z}$, $d \leq n$, that approximates a to within error $\frac{1}{dn}$, i.e., $\bigl|a - \frac{c}{d}\bigr| \leq \frac{1}{dn}$.

Hint: Consider the numbers $a, 2a, \ldots, n \cdot a$ and show that one of these numbers is at distance at most $1/n$ from some integer.

Solution: Using the hint, for some $d = 1, \ldots, n$ we would like to find an integer c such that $|d \cdot a - c| \leq \frac{1}{n}$. To this end, we will use the pigeonhole principle.

For every number $d \cdot a$, $1 \leq d \leq n$, let $c_d = \lfloor d \cdot a \rfloor$ be the closest smallest integer. Clearly, by construction we have $0 \leq d \cdot a - c_d < 1$.

Divide the interval $[0, 1]$ between 0 and 1 into n evenly spaced subintervals $\bigl[\frac{i-1}{n}, \frac{i}{n}\bigr]$ for $1 \leq i \leq 1$. If for some d and c_d we have that $d \cdot a - c_d$ is in the subinterval $\bigl[0, \frac{1}{n}\bigr]$, then fix $c = c_d$. Otherwise, all other subintervals will serve as $n - 1$ pigeonholes, and the numbers $d \cdot a - c_d$ for $1 \leq d \leq n$ will serve as pigeons. By the pigeonhole principle, there are two numbers $c_{d'}$ and $c_{d''}$

for $d' > d''$ such that $d' \cdot a - c_{d'}$ and $d'' \cdot a - c_{d''}$ end up in the same subinterval/pigeonhole. But this means that $c = c_{d'} - c_{d''}$ and $d = d' - d''$ are integers such that $|d \cdot a - c| \leq \frac{1}{n}$.

We have now found integers c and d such that $|d \cdot a - c| \leq \frac{1}{n}$. Dividing this inequality by d, we get that $\left|a - \frac{c}{d}\right| \leq \frac{1}{dn}$ as desired. (Just in case we worry whether dividing like this is in order, we can see that the inequality with absolute sign is equivalent to the two inequalities $-\frac{1}{n} \leq d \cdot a - c \leq \frac{1}{n}$, and dividing by the positive integer d yields $-\frac{1}{dn} \leq a - \frac{c}{d} \leq \frac{1}{n}$, which is the same inequality with absolute sign as above.)

4 (70 p) In this problem we focus on relations. Suppose that $A = \{e_0, e_1, \ldots, e_5\}$ is a set of 6 elements and consider the relation R on A represented by the matrix

$$M_R = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(where element e_i corresponds to row and column $i + 1$).

4a Let us write S to denote the symmetric closure of the relation R. What is the matrix representation of S? Can you explain in words what the relation S is by describing how it can be interpreted?

Solution: Let us identify the elements $e_0, e_1, \ldots, e_5$ with the numbers $0, 1, \ldots, 5$ for the purposes of this discussion.

By studying the matrix M_R we can conclude that a and b are related by R—which we have learned to denote by aRb, or $(a, b) \in R$—precisely when $b \equiv a+2 \pmod 6$. Taking the symmetric closure means that a and b can switch places, so we get that $(a, b) \in S$ if $b \equiv a \pm 2 \pmod 6$, and

$$M_S = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}$$

is the matrix representation of this relation.

We can also compute directly on the matrix and obtain M_S as the coordinate-wise Boolean or of M_R and $(M_R)^\top$.

4b Now let T be the transitive closure of the relation S. What is the matrix representation of T? Can you explain in words what the relation T is by describing how it can be interpreted?

Solution: As just noted, we have $(a, b) \in S$ if $b \equiv a \pm 2 \pmod 6$. Now, if for a, b, c it holds that $b \equiv a \pm 2 \pmod 6$ and $c \equiv b \pm 2 \pmod 6$, then c and a differ by 0, 2, or 4 (mod 6). Or, in other words, a and c are both odd or both even. This shows that S^2 is the relation such that $(a, b) \in S^2$ if a and b are both odd or both even. If $(a, b) \in S^2$ and $(b, c) \in S^2$, then again we can conclude that a and c are both odd or both even, i.e., that $(a, c) \in S^2$ Hence, $T = S^2 = S^3$ is closed under transitivity, and so is the transitive closure of S.

Since we have concluded that $(a, b) \in T$ if a and b are both odd or both even, we see that this relation can be represented as

$$M_T = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

in matrix form.

4c Suppose that we instead let T' be the transitive closure of the relation R, and then let S' be the symmetric closure of T'. Are S' and T the same relation? If they are not the same, show some way in which they differ. If they are the same, is it true that S' and T constructed in this way from some relation R on a set A will always be the same? Please make sure to motivate your answers clearly.

Solution: Applying the same kind of reasoning as in Problem 4b, we get that if $b \equiv a + 2 \pmod 6$ and $c \equiv b + 2 \pmod 6$, then $c \equiv a \pm 2 \pmod 6$, and one more application of transitivity gives us the relation T. Since we already argued that T is closed under transitivity, we have $T' = T$. This relation is also symmetric, so taking the symmetric closure does not change it. Hence, we indeed have $S' = T' = T$.

This does not hold in general, however. Let $A = \{1, 2\}$ and let R be the relation containing only the pair $(1, 2)$. Then taking the transitive closure does not affect anything, so the relation stays the same, and taking the symmetric closure of the transitive closure yields the relation $\{(1, 2), (2, 1)\}$ with two pairs. However, if we take the symmetric closure first, so that we get $\{(1, 2), (2, 1)\}$, then the transitive closure will also relate 1 and 2 with themselves, so that the final relation after symmetric closure followed by transitive closure is $\{(1, 1), (1, 2)(2, 1), (2, 2)\}$. Expressed in in slightly more fancy language, we have now shown that the two operations of symmetric closure and transitive closure do not commute.

5 (120 p) Recall that an undirected graph $G = (V, E)$ consists of a set of vertices V connected by edges E, where every edge is an unordered pair of vertices. If there is an edge (u, v) between two vertices u and v, then we say that u and v are the *endpoints* of the edge, and the two vertices are said to be *neighbours*. We say that a sequence of edges (v_1, v_2), (v_2, v_3), (v_3, v_4), $\ldots$, (v_{k-1}, v_k), in E is a *path* from v_1 to v_k.

In this problem, we wish to express properties of graphs in both natural language and predicate logic, and to translate between the two forms. We do this as follows:

- The universe is the set of vertices V of G.
- The binary predicate $E(u, v)$ holds if and only if there is an edge between u and v in G.
- The unary predicate $S(v)$ is used to identify a subset of vertices $S = \{v \mid v \in V, S(v) \text{ is true}\}$ for which some property might or might not hold.

For example, we can write the natural language statement "S is a set containing exactly k distinct elements" as a formula

$$\begin{aligned} \textsf{setsize}\,(S,k) \;:=\; & \exists u_1 \cdots \exists u_k \left(\bigwedge_{i=1}^{k-1} \bigwedge_{j=i+1}^{k} (u_i \neq u_j) \wedge \bigwedge_{i=1}^{k} S(u_i) \right) \\ & \wedge \neg \exists u_1 \cdots \exists u_{k+1} \left(\bigwedge_{i=1}^{k} \bigwedge_{j=i+1}^{k+1} (u_i \neq u_j) \wedge \bigwedge_{i=1}^{k} S(u_i) \right) \end{aligned} \tag{1}$$

where $\bigwedge_{i=1}^{k} \phi_i$ is shorthand for $\phi_1 \wedge \phi_2 \wedge \cdots \wedge \phi_k$, and where we use standard notation $\neg$ for logical negation. In natural language, this formula can be read as: "There exist k elements u_1 to u_k such that (i) for every pair (u_i, u_j), $i \neq j$, the elements themselves are also distinct ($u_i \neq u_j$); and (ii) all the u_i for $i = 1, \ldots, k$ are members of S, but no such set of $k+1$ elements u_1 to u_{k+1} exists."

Below you find six graph properties defined in natural language and six graph properties written as predicate logic formulas. Most of the natural language definitions have equivalent predicate logic formulas, but not all.

Your task is to translate each predicate logic formula (a), …, (f) into a natural language description, and argue which—if any—of the natural language definitions (1), …, (6) it matches.

Natural Language Definitions:

(1) A *dominating set* of size k for a graph $G = (V, E)$ is a set S of k distinct vertices such that every vertex v in the graph either is in S or is a neighbour of a vertex in S.

(2) A *clique* S of size k in a graph $G = (V, E)$ is a set S of k distinct vertices such that all vertices in S are neighbours with each other.

(3) A *disconnected vertex set* of size k in a graph $G = (V, E)$ is a set S of k distinct vertices such that there are no edges from any $u \in S$ to any $v \in V \setminus S$. [Here $V \setminus S$ denotes set subtraction, so that $V \setminus S = \{v \mid v \in V \text{ and } v \notin S\}$.]

(4) A *vertex cover* of size k of a graph $G = (V, E)$ is a set S of k distinct vertices such that for every edge $(u, v) \in E$ it holds that at least one of the endpoints is in S.

(5) A graph $G = (V, E)$ is *bipartite*, with one of the parts in the bipartition having size k, if there a set S of k distinct vertices such that all edges in the graph go between S and $V \setminus S$.

(6) A *connected component* S of size k in a graph $G = (V, E)$ is a set S of k distinct vertices such that for every pair of distinct vertices u and v in S there is a path from u to v consisting of vertices in S.

Predicate Logic Formulas:

(a) $\textsf{setsize}\,(S,k) \wedge \forall v \forall w \big(E(v,w) \to (S(v) \vee S(w))\big)$

(b) $\textsf{setsize}\,(S,k) \wedge \forall v \big(S(v) \vee \exists w (S(w) \wedge E(v,w))\big)$

(c) $\textsf{setsize}\,(S,k) \wedge \forall u \forall w \big((u \neq w \wedge S(u) \wedge S(w)) \to \exists v (E(u,v) \wedge E(v,w)\big)$

(d) $\textsf{setsize}\,(S,k) \wedge \forall v \forall w \big(E(v,w) \to ((S(v) \wedge \neg S(w)) \vee (\neg S(v) \wedge S(w)))\big)$

(e) $\textsf{setsize}\,(S,k) \wedge \forall v \big(S(v) \to \exists w (\neg S(w) \wedge E(v,w))\big)$

(f) $\mathsf{setsize}\,(S,k) \wedge \forall v \forall w \big((v \neq w \wedge S(v) \wedge S(w)) \rightarrow E(v,w) \big)$

For every predicate logic formula that matches a natural language description, explain clearly why there is a match. For formulas that do not match a description, write a natural language definition along the lines of (1), ..., (6) that describes the property that the formula encodes.

Solution: Formula (a) matches the description (4) of vertex cover. It says that if there is an edge between two vertices v and w, then at least one of these vertices is in S. This is the definition of S being a vertex cover.

Formula (b) matches the description (1) of dominating set. It says that for every vertex v it holds that either v is in the set S or there is some vertex w in S that is connected to v via an edge. This precisely matches the given description of what a dominating set is.

Formula (c) is unmatched. It encodes that S is a vertex set of size k such that any two distinct vertices in S have a common neighbour in the graph.

Formula (d) matches the description (5) of bipartiteness. It says that if there is an edge between v and w, then either v is in S and w is not, or the other way round. That is, exactly one of the vertices is any edge is in S, which shows that $(S, V \setminus S)$ is indeed a bipartition of the graph.

Formula (e) is unmatched. It encodes that S is a vertex set of size k such that every vertex v in S has an edge to some vertex w outside of S.

Formula (f) matches the description (2) of clique. It says that if two distinct vertices v and w are both in S, then there has to be an edge between them. This is exactly the condition for vertices in a clique.