Mathematical Foundations of Computer Science
Lecture Outline
November 14, 2024

Operations on Relations

We can take a relation or a pair of relations and produce a new relation. Since a relation R from set A to set B is a subset of $A \times B$, operations that apply to sets apply to relations.

Example. Let $A = \{1, 2, 3\}$ and $B = \{a, b, c, d\}$. Let $R_1 = \{(1, a), (1, c), (2, c), (3, a)\}$. Let $R_2 = \{(1, b), (1, c), (1, d), (2, b)\}$. Then we have

$$\begin{aligned}
R_1 \cup R_2 &= \{(1, a), (1, b), (1, c), (1, d), (2, b), (2, c), (3, a)\} \\
R_1 \cap R_2 &= \{(1, c)\} \\
R_1 \setminus R_2 &= \{(1, a), (2, c), (3, a)\} \\
R_2 \setminus R_1 &= \{(1, b), (1, d), (2, b)\}
\end{aligned}$$

Example. Let A and B be the set of all students and the set of all courses at a school, respectively. Suppose R_1 consists of all ordered pairs (a, b), where a is a student who has taken course b, and R_2 consists of all ordered pairs (a, b), where a is a student who requires course b to graduate. What are the relations $R_1 \cup R_2$, $R_1 \cap R_2$, $R_1 \oplus R_2$, $R_1 \setminus R_2$, and $R_2 \setminus R_1$?

Solution. $R_1 \cup R_2$ consists of all ordered pairs (a, b), where a is a student who has taken course b or requires course b to graduate.
$R_1 \cap R_2$ consists of all ordered pairs (a, b), where a is a student who has taken course b and requires course b to graduate.
$R_1 \oplus R_2$ consists of all ordered pairs (a, b), where a is a student who has taken course b or requires course b to graduate, but not both.
$R_1 \setminus R_2$ consists of all ordered pairs (a, b), where a is a student who has taken course b but does not require it to graduate.
$R_2 \setminus R_1$ consists of all ordered pairs (a, b), where a is a student who required course b to graduate but has not taken it.

Inverse Relation

Let R be a relation from A to B. Then the *inverse* of R, written R^{-1}, is the relation from B to A defined by

$$R^{-1} = \{(b, a) \mid (a, b) \in R\}$$

Example. Let $A = \{a, b, c\}$ and let $R = \{(a, a), (a, b), (b, a), (c, a)\}$. Then

$$R^{-1} = \{(a, a), (b, a), (a, b), (a, c)\}$$

Note that R and R^{-1} are almost equal.

Example. A relation R on a set A is symmetric iff $R = R^{-1}$.

Solution. ($\implies$) Suppose R is symmetric on A. We will prove that $R = R^{-1}$ by showing that $R \subseteq R^{-1}$ and $R^{-1} \subseteq R$. We will prove $R \subseteq R^{-1}$ by showing that an arbitrary element $(a, b) \in R$ is also in R^{-1}. Since R is symmetric, $(b, a) \in R$. By definition of R^{-1}, since $(b, a) \in R$, it must be that $(a, b) \in R^{-1}$. To prove $R^{-1} \subseteq R$, we will show that an arbitrary element $(a, b) \in R^{-1}$ is also in R. By definition of R^{-1}, it must be that $(b, a) \in R$. Since R is symmetric, (a, b) must also be in R.
($\impliedby$) Suppose that $R = R^{-1}$. Let (a, b) be an arbitrary ordered pair in R. To prove that R is symmetric we need to show that $(b, a) \in R$. By definition of R^{-1}, $(b, a) \in R^{-1}$. Since $R = R^{-1}$, R must contain (b, a).

Composition of Relations

Let R be a relation from A to B and S be a relation from B to C. The *composition of S with R* is the relation from A to C:

$$S \circ R = \{(x, z) \mid \text{ there exists a } y \in B \text{ such that } x\,R\,y \text{ and } y\,S\,z\}$$

Example. Let $A = \{1, 2, 3, 4\}$, $B = \{3, 4, 5, 6\}$, and $C = \{a, b, c\}$. Let R and S be relations from A to B and from B to C, respectively, where

$$R = \{(1, 3), (3, 3), (3, 4), (4, 5), (4, 6)\}$$
$$S = \{(3, b), (4, a), (4, c), (5, a), (5, b), (6, c)\}$$

What is the composite of the relations R and S?

Solution. $S \circ R = \{(1, b), (3, a), (3, b), (3, c), (4, a), (4, b), (4, c)\}$

Let R be a relation on a set A. The powers $R^n, n = 1, 2, 3, \ldots,$ are defined recursively by

$$R^1 = R \qquad \text{and} \qquad R^{n+1} = R^n \circ R$$

Observe that $R^2 = R \circ R, R^3 = R^2 \circ R = (R \circ R) \circ R$, and so on.

Example. Let R be a relation on a set A. Then R is transitive iff $R^n \subseteq R$, for all $n \geq 1$.

Solution. We first show that if $R^n \subseteq R$, for all $n \geq 1$, then R is transitive. Note that if $(a, b) \in R$ and $(b, c) \in R$ then $(a, c) \in R^2$. Since $R^2 \subseteq R$, it must be that $(a, c) \in R$, which means that R is transitive.

We will prove R is transitive $\implies$ $R^n \subseteq R$, for all $n \geq 1$, using induction on n.
<u>Induction hypothesis</u>: Assume that if R is transitive then $R^k \subseteq R$, for some $k \geq 1$.
<u>Base Case</u>: The claim holds tivially when $n = 1$, since $R^1 = R$.
<u>Induction Step</u>: We want to prove the claim when $n = k + 1$. In other words, we want to prove that if R is transitive then $R^{k+1} \subseteq R$. We will prove this by showing that an arbitrary

but particular ordered pair (a, b) in R^{k+1} is also present in R. By definition, $R^{k+1} = R^k \circ R$. Since $(a, b) \in R^{k+1}$, there must be a c, such that $(a, c) \in R$ and $(c, b) \in R^k$. We know by induction hypothesis that $R^k \subseteq R$, which means that $(c, b) \in R$. Since R is transitive, and $(a, c) \in R$ and $(c, b) \in R$, we have $(a, b) \in R$. This completes the proof.

An *independent set* S in G is a subset of vertices such that no two vertices in S share an edge. The *independence number* of a graph G, denoted by $\alpha(G)$ is the size of the largest independent set in G.

Example. Let n be the number of vertices in G and m be the number of edges, and let $d = \frac{2m}{n} \geq 1$ be the average degree. Then

$$\alpha(G) \geq \frac{n}{2d}$$

This is a weaker version of the celebrated Turán's theorem.

Solution. Construct a random subset S of vertices by placing each vertex in S independently with probability p (to be determined later). Let X be the random variable denoting the number of vertices in S and let Y be the random variable denoting the number of edges whose both endpoints are in S. Let Y_e be an indicator random variable that is 1 iff both endpoints of e are in S. By the Linearity of Expectation we have

$$\mathbf{E}[X] = np \qquad \text{and} \qquad \mathbf{E}[Y] = \sum_e \mathbf{E}[Y_e] = \sum_e \Pr[Y_e = 1] = mp^2 = \frac{nd}{2}p^2$$

Note that the quantity $X - Y$ denotes the number of vertices in S minus the number of edges with both endpoints in S. By the Linearity of Expectation we get

$$\mathbf{E}[X - Y] \geq np - \frac{nd}{2}p^2 = np\left(1 - \frac{dp}{2}\right)$$

This means that there exists a set S such that the number of vertices in S exceeds the number of edges in S by the above quantity. We now modify set S by deleting an arbitrary endpoint of each edge. The resulting set S' has at least $np\left(1 - \frac{dp}{2}\right)$ vertices left and has no edges between any of its vertices. We want to maximize $|S'|$, so we set $p = 1/d$ (using $d \geq 1$), giving us $|S'| = \frac{n}{2d}$.

For any graph $G = (V, E)$, a set of vertices $D \subseteq V$ is called a *dominating set* if every vertex in $V \setminus D$ is adjacent to a vertex in D.

Example. Prove that any connected graph $G = (V, E)$ with $n \geq 2$ vertices and minimum degree $\delta(G) = \delta$, contains a dominating set of size at most $\frac{n(1+\ln(1+\delta))}{1+\delta}$.

Solution. For each vertex $v \in V$, add it to the set X independently with probability p. Let $Y \subseteq V \setminus X$ be the vertices that are not dominated by X, i.e., they are vertices in $V \setminus X$ that are not dominated by X. Then $X \cup Y$ is a dominating set for G. We will now show that $\mathbf{E}[X \cup Y]$ is not too large. Since X and Y are disjoint sets, we have

$$\mathbf{E}[X \cup Y] = \mathbf{E}[X] + \mathbf{E}[Y] \tag{1}$$

We consider the following random variables.
X_v: random variable that is 1 if vertex v is in X, 0, otherwise.
Y_v: random variable that is 1 if vertex v and all of its neighbors are not in X, 0, otherwise.

$$\begin{aligned} X &= \sum_v X_v \\ \therefore \mathbf{E}[X] &= \sum_v \Pr[X_v = 1] \\ &= np \end{aligned}$$

$$\begin{aligned} Y &= \sum_v Y_v \\ \therefore \mathbf{E}[Y] &= \sum_v \Pr[Y_v = 1] \\ &= \sum_v (1-p)^{\deg(v)+1} \\ &\leq \sum_v (1-p)^{\delta+1} \\ &= n(1-p)^{\delta+1} \end{aligned}$$

Plugging the values of $\mathbf{E}[X]$ and $\mathbf{E}[Y]$ in (1) we get

$$\mathbf{E}[X \cup Y] \leq np + n(1-p)^{\delta+1} \leq np + ne^{-p(\delta+1)},$$

The last expression is minimized when

$$p = \frac{\ln(1+\delta)}{1+\delta}$$

Thus, we can find a dominating set of size at most $\frac{n(1+\ln(1+\delta))}{1+\delta}$.