Relations

A "relation" is a fundamental mathematical notion expressing a relationship between elements of sets.

Definition 0.1. A *binary relation from a set A to a set B* is a subset $R \subseteq A \times B$.

So, R is a set of ordered pairs. We often write $a \sim_R b$ or aRb to mean that $(a, b) \in R$.

Functions, for example, are a type of relation. The abstract notion of a relation is useful both in mathematics and in practice for modeling many different sorts of relationships. It's the basis of the *relational database* model, the standard data model for practical data processing systems.

Many times we will talk about a "relation on the set A", which means that the relation is a subset of $A \times A$. We can also define a *ternary* relation on A as a subset $R \subseteq A^3$ or, in general, an n-ary relation as a subset $R \subseteq A^n$, or $R \subseteq A_1 \times A_2 \times \cdots \times A_n$ if the sets A_i are different. In this class, we will focus only on binary relations. Here are some examples:

1. The relation "is taking class" as a subset of $\{$students at MIT$\} \times \{$classes at MIT$\}$. A relation from students to classes.
2. The relation "has lecture in" as a subset of $\{$classes at MIT$\} \times \{$rooms at MIT$\}$. A relation from classes to rooms.
3. The relation "is living in the same room" as a subset of $\{$students at MIT$\} \times \{$students at MIT$\}$. A relation on students.
4. The relation "can drive from first to second city". (Not necessarily directly—just some way, on some roads.)
5. Relation on computers, "are connected (directly) by a wire"
6. "meet one another on a given day"
7. "likes"
8. Let $A = \mathbb{N}$ and define $a \sim_R b$ iff $a \leq b$.
9. Let $A = \mathcal{P}(\mathbb{N})$ and define $a \sim_R b$ iff $a \cap b$ is finite.
10. Let $A = \mathbb{R}^2$ and define $a \sim_R b$ iff $d(a, b) = 1$.
11. Let $A = \mathcal{P}(\{1, \ldots, n\})$ and define $a \sim_R b$ iff $a \subseteq b$.

1 Properties of Relations

Once we have modeled something abstractly as a relation, we can talk about its properties without referring to the original problem domain. For a relation on a set A there are several standard properties of relations that occur commonly. Later on we will use these properties to classify different types of relations.

Definition 1.1. A binary relation R on A is:

1. *reflexive* if for every $a \in A$, $a \sim_R a$.
2. *symmetric* if for every $a, b \in A$, $a \sim_R b$ implies $b \sim_R a$.
3. *antisymmetric* if for every $a, b \in A$, $a \sim_R b$ and $b \sim_R a$ implies $a = b$.
4. *asymmetric* if for every $a, b \in A$, $a \sim_R b$ implies $\neg(b \sim_R a)$.
5. *transitive* if for every $a, b, c \in A$, $a \sim_R b$ and $b \sim_R c$ implies $a \sim_R c$.

The difference between antisymmetric and asymmetric relations is that antisymmetric relations may contain pairs (a, a), i.e., elements can be in relations with themselves, while in an asymmetric relation this is not allowed. Clearly, any asymmetric relation is also antisymmetric, but not vice versa.

Among our relations from Example :

- Relation 3 is reflexive, symmetric, transitive.
- Relation 4 is reflexive, transitive. Not necessarily symmetric, since roads could be one-way (consider Boston), but in actuality . . . . But definitely not antisymmetric.
- Relation 5 is symmetric but not transitive. Whether it is reflexive is open to interpretation.
- Relation 6 likewise.
- Relation 7 is (unfortunately) not symmetric. Not antisymmetric. Not transitive. Not even reflexive!
- Relation 8 is reflexive, antisymmetric, transitive.
- Relation 9 is not reflexive. It is symmetric. It is not transitive. $\{\text{even naturals}\} \cap \{\text{odd naturals}\}$ is finite (empty), but not $\{\text{even naturals}\} \cap \{\text{even naturals}\}$.
- Relation 10 is only symmetric.
- Relation 11 is reflexive, antisymmetric and transitive.

2 Representation

There are many different ways of representing relations. One way is to describe them by properties, as we did above. For infinite sets, that's about all we can do. But for finite sets, we usually use some method that explicitly enumerates all the elements of the relation. Some alternatives are lists, matrices and graphs. Why do we have so many different ways to represent relations? Different representations may be more efficient for encoding different problems and also tend to highlight different properties of the relation.

2.0.1 Lists

A finite relation from set A to set B can be represented by a list of all the pairs.

Example 2.1. The relation from $\{0, 1, 2, 3\}$ to $\{a, b, c\}$ defined by the list:

$\{(0, a), (0, c), (1, c), (2, b), (1, a)\}$.

Example 2.2. The divisibility relation on natural numbers $\{1, \ldots, 12\}$ is represented by the list:

$$\{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(1,7),(1,8),(1,9),(1,10),(1,11),(1,12),(2,2),(2,4),\\ (2,6),(2,8),(2,10),(2,12),(3,3),(3,6),(3,9),(3,12),(4,4),(4,8),(4,12),(5,5),(6,6),\\ (6,12),(7,7),(8,8),(9,9),(10,10),(11,11),(12,12)\}.$$

We can recognize certain properties by examining this representation:

Reflexivity: Contains all pairs (a, a).

Symmetry: Contains (a, b) then contains (b, a).

Transitivity: Contains (a, b) and (b, c) then contains (a, c).

2.0.2 Boolean Matrices

Boolean matrices are a convenient representation for representing relations in computer programs. The rows are for elements of A, columns for B, and for every entry there is a 1 if the pair is in the relation, 0 otherwise.

Example 2.3. The relation from Example 2.1 is represented by the matrix

	a	b	c
0	1	0	1
1	1	0	1
2	0	1	0
3	0	0	0

Example 2.4. The divisibility relation over $\{1, 2, \ldots, 12\}$ is represented by the enormous matrix

	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	1	1	1	1	1	1	1	1	1
2	0	1	0	1	0	1	0	1	0	1	0	1
3	0	0	1	0	0	1	0	0	1	0	0	1
4	0	0	0	1	0	0	0	1	0	0	0	1
5	0	0	0	0	1	0	0	0	0	1	0	0
6	0	0	0	0	0	1	0	0	0	0	0	1
7	0	0	0	0	0	0	1	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	0	0	0
9	0	0	0	0	0	0	0	0	1	0	0	0
10	0	0	0	0	0	0	0	0	0	1	0	0
11	0	0	0	0	0	0	0	0	0	0	1	0
12	0	0	0	0	0	0	0	0	0	0	0	1

Again, properties can by recognized by examining the representation:

Reflexivity the major diagonal is all 1.

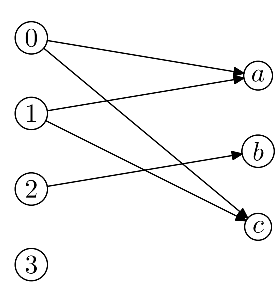
Symmetry: the matrix is clearly not symmetric across the major diagonal.

Transitivity: not so obvious . . .

2.0.3 Digraphs

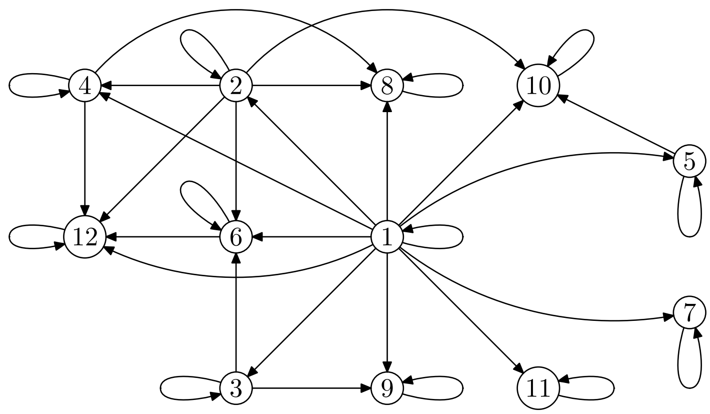
We can draw a picture of a relation $R \subseteq A \times B$ by drawing a dot for every element of A, a dot for every element of B, and an arrow from $a \in A$ to $b \in B$ iff aRb. Such a picture is called a *directed graph*, or *digraph* for short.

Example 2.5. The relation from Example 2.1 is represented by the digraph



Digraphs are mainly used for relations where $A = B$, i.e., for relations on a finite set A. To represent such a relation as a digraph we draw a dot (vertex) for each element of A, and draw an arrow from first element to second element of each pair in the relation. The digraph may contain self-loops, i.e. arrows from a dot to itself, associated to the elements a such that (a, a) is in the relation.

Example 2.6. The divisibility relation over $\{1, 2, \ldots, 12\}$ is represented by the digraph



Reflexivity: All nodes have self-loops.

Symmetry: all edges are bidirectional.

Transitivity: Short-circuits—for any sequence of consecutive arrows, there is a single arrow from the first to the last node.

3 Operations on Relations

3.1 Inverse

If R is a relation on $A \times B$, then R^{-1} is a relation on $B \times A$ given by $R^{-1} = \{(b, a) \mid (a, b) \in R\}$. It's just the relation "turned backwards."

Example 3.1. Inverse of "is taking class" (Relation 1 in Example) is the relation "has as a student" on the set $\{\text{classes at MIT}\} \times \{\text{students at MIT}\}$; a relation from classes to students.

We can translate the inverse operation on relation to operations on the various representations of a relation. Given the matrix for R, we can get the matrix for R^{-1} by transposing the matrix for R (note that the inverse of a relation is not the same thing as the inverse of the matrix representation). Given a digraph for R, we get the graph for R^{-1} by reversing every edge in the original digraph.

3.2 Composition

The *composition* of relations $R_1 \subseteq A \times B$ and $R_2 \subseteq B \times C$ is the relation

$$R_2 \circ R_1 = \{(a, c) \mid (\exists b)((a, b) \in R_1) \wedge ((b, c) \in R_2)\}.$$

In words, the pair (a, c) is in $R_2 \circ R_1$ if there exists an element b such that the pair (a, b) is in R_1 and the pair (b, c) is in R_2. Another way of thinking about this is that a "path" exists from element a to c via some element in the set B[1].

[1]Notice that $R_2 \circ R_1$ and $R_1 \circ R_2$ are different. The symbol $\circ$ is a source of eternal confusion in mathematics—if you read a book you should always check how the authors define $\circ$—some of them define composition the other way around.

Example 3.2. The composition of the relation "is taking class" (Relation 1 in Example) with the relation "has lecture in" (Relation 2 in Example) is the relation "should go to lecture in", a relation from {students at MIT} to {rooms at MIT}.

Example 3.3. Composition of the parent-of relation with itself gives grandparent-of. Composition of the child-of relation with the parent-of relation gives the sibling-of relation. (Here we relax the meaning of sibling to include that a person is the sibling of him/herself.) Does composition of parent-of with child-of give married-to/domestic partners? No, because that misses childless couples.

Example 3.4. Let B be the set of boys, G be the set of girls, $R_1 \subseteq B \times G$ consist of all pairs (b, g) such that b is madly in love with g, and $R_2 \subseteq G \times B$ consist of all pairs (g, b) such that g is madly in love with b. What are the relations $R_2 \circ R_1$ and $R_1 \circ R_2$, respectively?

3.2.1 Computing Composition and Path Lengths

If we represent the relations as matrices, then we can compute the composition by a form of "boolean" matrix multiplication, where $+$ is replaced by $\vee$ (Boolean OR) and $\times$ is replaced by $\wedge$ (Boolean AND).

Example 3.5. Let R_1 be the relation from Example 2.1:

	a	b	c
0	1	0	1
1	1	0	1
2	0	1	0
3	0	0	0

Let R_2 be the relation from $\{a, b, c\}$ to $\{d, e, f\}$ given by:

	d	e	f
a	1	1	1
b	0	1	0
c	0	0	1

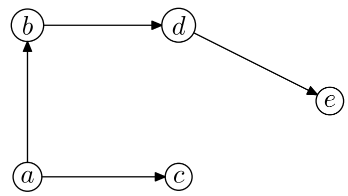
Then $R_2 \circ R_1 = \{(0, d), (0, e), (0, f), (1, d), (1, e), (1, f), (2, e)\}$, that is,

	d	e	f
0	1	1	1
1	1	1	1
2	0	1	0
3	0	0	0

A relation on a set A can be composed with itself. The composition $R \circ R$ of R with itself is written R^2. Similarly R^n denotes R composed with itself n times. R^n can be recursively defined: $R^1 = R$, $R^n = R \circ R^{n-1}$.

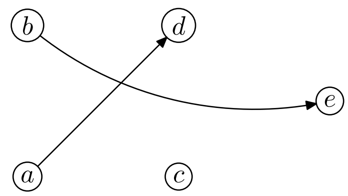
Example 3.6. Consider the relation $R = \{(a, b), (a, c), (b, d), (d, e)\}$ or:

	a	b	c	d	e
a	0	1	1	0	0
b	0	0	0	1	0
c	0	0	0	0	0
d	0	0	0	0	1
e	0	0	0	0	0



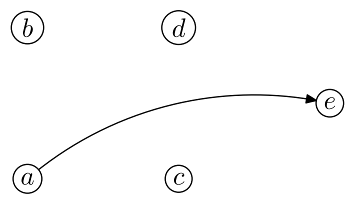
R^2 will be:

	a	b	c	d	e
a	0	0	0	1	0
b	0	0	0	0	1
c	0	0	0	0	0
d	0	0	0	0	0
e	0	0	0	0	0



R^3 will be:

	a	b	c	d	e
a	0	0	0	0	1
b	0	0	0	0	0
c	0	0	0	0	0
d	0	0	0	0	0
e	0	0	0	0	0



Definition 3.7. A *path* in a relation R is a sequence $a_0, \ldots, a_k$ with $k \geq 0$ such that $(a_i, a_{i+1}) \in R$ for every $i < k$. We call k the *length* of the path.

In the digraph model, a path is something you can trace out by following arrows from vertex to vertex, without lifting your pen. Note that a singleton vertex is a length 0 path (this is just for convenience). A *simple path* is a path with no repeated vertices.

Lemma 3.8. $R^n = \{(a, b) \mid$ *there is a length n path from a to b in R*$\}$

Proof. By induction. Let

$$P(n) ::= R^n = \{(a, b) \mid \text{there is a length } n \text{ path from } a \text{ to } b \text{ in } R\}.$$

The base case is clear. There is exactly one edge from a to b for every $(a, b) \in R$. And there is exactly one pair $(a, b) \in R$ for every edge from a to b, $P(1)$ is true. Note that since the induction hypothesis is an equality we have to prove both sides.

For the inductive step, suppose $P(n)$ is true.

First consider a path $a_0, \ldots, a_{n+1}$ in R. This is a path $a_0, \ldots, a_n$ in R followed by a pair (a_n, a_{n+1}) of R. By the inductive hypothesis, we can assume that $(a_0, a_n) \in R^n$. And we have already mentioned that $(a_n, a_{n+1}) \in R$. Therefore $(a_0, a_{n+1}) \in R^{n+1}$ by the definition of composition. Thus, every path of length $n+1$ corresponds to a relation in R^{n+1}.

Now consider a pair $(a, b) \in R^{n+1}$. By the definition of composition, there exists a c such that $(a, c) \in R^n$ and $(c, b) \in R$. By the inductive hypothesis, we can assume that (a, c) corresponds to a length n path from a to c, and since $(c, b) \in R$ there is an edge from c to b. Thus, there is a length $n+1$ path from a to b. To conclude, $P(n) \longrightarrow P(n+1)$. □

3.3 Closure

A closure "extends" a relation to satisfy some property. But extends it as little as possible.

Definition 3.9. The *closure* of relation R with respect to property P is the relation S that

(i) contains R,

(ii) has property P, and

(iii) is contained in *any* relation satisfying (i) and (ii).

That is, S is the "smallest" relation satisfying (i) and (ii).

As a general principle, there are two ways to construct a closure of R with respect to property P: We can either start with R and add as few pairs as possible until the new relation has property P; or we can start with the largest possible relation (which is $A \times A$ for a relation on A) and then remove as many not-in-R pairs as possible while preserving the property P.

3.3.1 The Reflexive Closure

Lemma 3.10. *Let R be a relation on the set A. The reflexive closure of R is $S = R \cup \{(a, a), \forall a \in A\}$.*

Proof. It contains R and is reflexive by design. Furthermore (by definition) any relation satisfying (i) must contain R, and any satisfying (ii) must contain the pairs (a, a), so any relation satisfying both (i) and (ii) must contain S. □

Example 3.11. Let $R = \{(a, b)(a, c)(b, d)(d, e)\}$, then the reflexive closure of R is

$$\{(a, b)(a, c)(b, d)(d, e)(a, a)(b, b)(c, c)(d, d)(e, e)\}.$$

3.3.2 The Symmetric Closure

Lemma 3.12. *Let R be a relation on the set A. The symmetric closure of R is $S = R \cup R^{-1}$.*

Proof. This relation is symmetric and contains R. It is also the smallest such. For suppose we have some symmetric relation T with $R \subseteq T$. Consider $(a, b) \in R$. Then $(a, b) \in T$ so by symmetry $(b, a) \in T$. It follows that $R^{-1} \subseteq T$. So $S = R \cup R^{-1} \subseteq T$. □

Example 3.13. Let $R = \{(a, b)(a, c)(b, d)(d, e)\}$, then the symmetric closure of R is

$$\{(a, b)(a, c)(b, d)(d, e)(b, a)(c, a)(d, b)(e, d)\}$$

3.3.3 The transitive closure

The transitive closure is a bit more complicated than the closures above.

Lemma 3.14. *Let R be a relation on the set A. The transitive closure of a relation R is the set*

$$S = \{(a, b) \in A^2 \text{given there is a path from } a \text{ to } b \text{ in } R\}.$$

Proof. Obviously, $R \subseteq S$. Next, we show that S is transitive. Suppose $(a, b) \in S$ and $(b, c) \in S$. This means that there is an (a, b) path and a (b, c) path in R. If we "concatenate" them (attach the end of the (a, b) path to the start of the (b, c) path, we get an (a, c) path. So $(a, c) \in S$. So S is transitive.

Finally, we need to show that S is the smallest transitive relation containing R. So consider any transitive relation T containing R. We have to show that $S \subseteq T$. Assume for contradiction that $S \nsubseteq T$. This means that some pair $(a, b) \in S$ but $(a, b) \notin T$. In other words, there is a path $a_0, \ldots, a_k = b$ in R where $k \geq 1$ and $(a_0, a_k) \notin T$. Call this a *missing path*. Now let M be the set of missing paths.

Now we use well-ordering. We have just claimed that the set M of missing paths is nonempty. So consider a shortest missing path $s_0, \ldots, s_m$. We will derive a contradiction to this being the shortest missing path.

Case 1: $m = 1$. Then s_0, s_1 is a path in R, so $(s_0, s_1) \in R$. But we know T contains R, so $(s_0, s_1) \in T$, a contradiction.

Case 2: $m > 1$. Then $s_1, \ldots, s_{m-1}$ is a path in R ($m - 1 > 0$). But it is shorter than our original shortest missing path, so cannot be missing. Thus $(s_1, s_{m-1}) \in T$. Also we have $(s_{m-1}, s_m) \in T$ since $R \subseteq T$. Thus by transitivity of T, $(s_1, s_m) \in T$, a contradiction.

We get a contradiction either way, so our assumption (that $S \nsubseteq T$) is false. This completes the proof. □

Wait a minute. Well-ordering is applied to sets of *numbers*; we applied it to a set of paths! How? Well, look at the set of "lengths of missing paths". It is nonempty, so has a smallest element. There is path that has this length—so it is a shortest path.