2 Relation theory

2.1 A brief overview of set theory

We collect in this section the basic facts about set theory which are needed to understand relations and functions. We will use quantifiers and first order formulas informally, deferring a proper treatment to the next chapter. The description of the foundational part of set theory will be cursory too.

□ **The language and axioms of set theory.** We use letters like x, lowercase or uppercase alike, as variables ranging over sets. If φ is a property relevant to sets, we write:

- $\varphi(a)$ if the set a has the property φ;
- $\exists x.\varphi(x)$ if there exists at least one set a such that $\varphi(a)$;
- $\forall x.\varphi(x)$ if $\varphi(a)$ for every set a.

The symbols $\exists$ and $\forall$ are the *existential* and *universal* quantifiers respectively. The expression $\exists x.\varphi(x)$ is read '*there exists an x such that $\varphi(x)$*' and $\forall x.\varphi(x)$ is read '*for all x, $\varphi(x)$*'. We write $x = y$ to indicate that the sets x and y are equal and $x \in y$ to indicate that the set x is an *element* of the set y; these are called *atomic formulas*. We write $x \notin y$ as an abbreviation of $\neg(x \in y)$, meaning that x is not an element of y. The admissible formulas φ are constructed from atomic formulas using propositional connectives and quantifiers. We can use this language to give simple definitions.

DEFINITION 2.1. A set x is a *subset* of a set y if every element of x is also an element of y. We write $x \subseteq y$ when this happens. Thus,

$$x \subseteq y := \forall z(z \in x \to z \in y)\,. \tag{2.1}$$

The main question is, of course, what a set is. In naïve set theory one constructs sets using the *comprehension* axiom:

$$\exists x(y \in x \leftrightarrow \varphi(y))\,. \tag{2.2}$$

This means that given any formula φ one can form the set x whose elements are all the sets y which have the property φ. It was soon discovered that this formulation leads to inconsistencies. To avoid these, rules have been introduced to explicitly describe which sets can be constructed, how they can be constructed and how they can be compared. We will not discuss these rules here; we will simply list the axioms of one of these theories, known as ZF from the names of the its first investigators, Zermelo and Fränkel.

$\forall xy(x = y \leftrightarrow \forall z(z \in x \leftrightarrow z \in y))$	Extensionality	(2.3)
$\exists x\forall y(y \notin x)$	Empty set	(2.4)
$\forall x\exists y\forall z(z \in y \leftrightarrow z \subseteq x)$	Power set	(2.5)
$\forall x\exists y\forall z(z \in y \leftrightarrow z \in x \wedge \varphi(z))$	Separation	(2.6)
$\forall x\exists y\forall z(z \in y \leftrightarrow \exists w(w \in x \wedge \varphi(w, z))$	Substitution	(2.7)
$\forall x\exists y\forall z(z \in y \leftrightarrow \exists w(z \in w \wedge w \in x)$	Union	(2.8)
$\forall x(x \neq \varnothing \to \exists y(y \in x \wedge \forall z(z \in y \to z \notin x)))$	Regularity	(2.9)
$\exists x(\varnothing \in x \wedge \forall y(y \in x \to y \cup \{y\} \in x))$	Infinity	(2.10)

We only provide a superficial explanation of some of these formulas. Extensionality means that two sets are equal exactly if they have the same elements. The empty set axiom states that there exists — i.e. one is allowed to construct — a set with no elements. The power set axiom says that if we have constructed a set x, then we can also construct the set y, called the powerset of x and denoted by Px, whose elements are the subsets of x. It can be shown that if x has *cardinality* n — i.e. n elements —,then Px has cardinality 2^n; starting with the empty set, we can therefore construct sets with an arbitrarily large albeit finite number of elements. Separation states that if we already have x and a formula φ, we can construct the set y whose elements are the elements of x satisfying φ; we write

$$y = \{z \in x : \varphi(z)\} \tag{2.11}$$

to denote this set. Unions states that given a set x we can construct the set y whose elements are the elements z of the sets w which are elements of x. To understand substitution, call a formula $\varphi(w, z)$ with two variables *functional* if for every set w there exists exactly one z such that w and z together satisfy φ. Then substitution states that given an existing set x we can construct a new set y by sustituting every element $w \in x$ with the corresponding $z \in y$ provided by φ. Regularity states that is a set x is non-empty, then it has an element which is disjoint from x, i.e. x and y have no common elements. An important consequence of regularity is that none of the sets we can construct in ZF is an element of itself. In particular, there is no set of all sets. Infinity states the existence of an infinite set: the set $sy := y \cup \{y\}$ is the *successor* of y and since $y \notin y$ by regularity, it has one more element than y.

□ **The Boolean algebra of subsets.** We now assume we have constructed a set U, called *the universe*, and examine the algebraic structure of its subsets.

Definition 2.2. Assume U is a fixed universe and $A, B \subseteq U$ are subsets.

- The *intersection* $A \cap B$ of A and B is the set whose elements are the elements of U belonging to both A and B.

$$A \cap B = \{x \in U : (x \in A) \wedge (x \in B)\} \tag{2.12}$$

- The *union* $A \cup B$ of A and B is the set whose elements are the elements of U belonging to either A or B (or both).

$$A \cup B = \{x \in U : (x \in A) \vee (x \in B)\} \tag{2.13}$$

- The *complement* A' of A is the set whose elements are the elements of U which do not belong to A.

$$A' = \{x \in U : x \notin A\} \tag{2.14}$$

There are other operation on PU which are less common, like the difference and symmetric difference of two subsets, defined by the formulas

$$A \setminus B = \{x \in U : (x \in A) \wedge (x \notin B)\}, \qquad A \triangle B = \{x \in U : (x \in A) \leftrightarrow (x \notin B)\} \tag{2.15}$$

Observe that all these operations are defined using propositional connectives. In fact, the properties of propositional connectives are exactly what is needed to prove the following

PROPOSITION 2.3. The powerset of a universe U is a Boolean algebra with respect to union, intersection and complement, i.e. the following formulas hold.

$(A \cap B) \cap C = A \cap (B \cap C)$	$(A \cup B) \cup C = A \cup (B \cup C)$	Associatività	(2.16)
$A \cap B = B \cap A$	$A \cup B = B \cup A$	Commutatività	(2.17)
$A \cap (A \cup B) = A$	$A \cup (A \cap B) = A$	Assorbimento	(2.18)
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Distributività	(2.19)
$A \cap U = A$	$A \cup \varnothing = A$	Elemento neutro	(2.20)
$A \cap A' = \varnothing$	$A \cup A' = U$	Complemento	(2.21)
$A \cap A = A$	$A \cup A = A$	Idempotenza	(2.22)
$(A \cap B)' = A' \cup B'$	$(A \cup B)' = A' \cap B'$	Dualità	(2.23)
$A \cap \varnothing = \varnothing$	$A \cup U = U$	Elemento assorbente	(2.24)
$(A')' = A$		Involuzione	(2.25)

Proof. The proof depends on the corresponding properties of propositional connectives. We prove associativity of intersection.

$$\begin{aligned} x \in (A \cap B) \cap C &\equiv (x \in A \cap B) \wedge (x \in C) && (2.26)\\ &\equiv ((x \in A) \wedge (x \in B)) \wedge (x \in C) && (2.27)\\ &\equiv (x \in A) \wedge ((x \in B) \wedge (x \in C)) && (2.28)\\ &\equiv (x \in A) \wedge (x \in B \cap C) && (2.29)\\ &\equiv x \in (A \cap (B \cap C)) && (2.30) \end{aligned}$$

All the equivalences follow from the definition of intersection, excep for the third which is a consequence of associativity of conkunction. Thus, $(A \cap B) \cap C$ e $A \cap (B \cap C)$ have the same elements and are therefore equal by the extensionality axiom. Proofs for the remaining formulas are similar. □

□ **Products.** We know from the extensionality axiom that the order of elements in a set is irrelevant: if we shuffle the elements of a set we still have the same set. There are, however, cases in which it is desirable to keep track of some ordering on the elements. One example is the case of cartesian coordinates of points in the plane: the point with first coordinate 1 and second coordinate 2, usually denoted by the symbol $(1, 2)$ is not the same as the point $(2, 1)$ with coordinates exchanged. This means that we can not use the set $\{1, 2\}$ to describe the coordinates of either point. What we need is a different construction.
Observe first that if a and b are sets, we can construct the set $\{a, b\}$ whose elements are precisely a and b. To see this, start from the set with two elements $2 = s^2(0) = \{0, 1\}$ and apply the substitution axiom with the formula

$$\varphi(x, y) := (x = 0 \to y = a) \wedge (x \neq 0 \to y = b)\,. \tag{2.31}$$

We can repeat this construction and, from a and the newly formed $\{a, b\}$, construct the set $\{a, \{a, b\}\}$. This is the definition we need.

DEFINITION 2.4. (KURATOWSKI) If a and b are sets, the set

$$(a, b) := \{a, \{a, b\}\} \tag{2.32}$$

is the *pair* with *first coordinate* a and *second coordinate* b.

This definition achieves the desired distinction:

PROPOSITION 2.5. Pairs (a, b) and (c, d) are equal if and only if their corresponding coordinates are equal:

$$(a, b) = (c, d) \Leftrightarrow (a = c) \wedge (b = d)\,. \tag{2.33}$$

Once pairs are defined we can describe binary products of set.

DEFINITION 2.6. If A and B are sets, their *cartesian product* is the set

$$A \times B = \{(a, b) \in P(A \cup P(A \cup B)) : (a \in A) \wedge (b \in B)\}\,. \tag{2.34}$$

The adjective *cartesian* is often dropped and we simply talk about the product of A and B.

Example 2.7 Given the sets

$$A = \{0, 1\}, \qquad B = \{0, 1, 2\} \tag{2.35}$$

their product is the set

$$A \times B = \{(0, 0),\ (0, 1),\ (0, 2),\ (1, 0),\ (1, 1),\ (1, 2)\}\,. \tag{2.36}$$

Observe that the cardinality of $A \times B$ is product of the cardinalities of A and B, in symbols $|A \times B| = |A| \cdot |B|$. This is true also in the infinite case, once the proper framewok for transinite arithmetic is in place. One can also compute the product of B and A in the opposite order,

$$B \times A = \{(0, 0),\ (0, 1),\ (1, 0),\ (1, 1),\ (2, 0),\ (2, 1)\}\,. \tag{2.37}$$

Note that $A \times B \neq B \times A$ because they do not have the same elements: for example, the pair $(0, 2)$ is an element of $A \times B$ but not of $B \times A$. In other words, the product of sets is not commutative. Once the notion is isomorphism is defined, though, it is possible to prove that it is commutative up to isomorphism. The product of a set with itself is the *cartesian square* of the set. For example,

$$A^2 := A \times A = \{(0, 0),\ (0, 1),\ (1, 0),\ (1, 1\}\,. \tag{2.38}$$

2.2 Binary relations

□ **Definition and examples.** Binary relations and functions provide a notion of becoming in set theory. If a set can regarded as a state, a binary relation between two sets can be regarded as a transformations between two states.

DEFINITION 2.8. Given sets A and B, a *relation* from A to B is a subset $R \subseteq A \times B$ of the cartesian product of the two sets. A and B are the *domain* and *codomain* of R and we will write

$$R : A \rightharpoonup B \tag{2.39}$$

to represent such a relation. If $a \in A$ and $b \in B$ we will write aRb instead of $(a, b) \in R$ to indicate that the pair (a, b) belongs to the relation.

We will also say that R is a *binary relation* or a relation of *arity 2* because it is contained in the product of two sets. Note that we consider the domain A and the codomain B as part of the definition of relation, in the sense that a relation is in fact a pair $((A, B), R)$ consisting of a a pair of sets A and B together with a subset R of the cartesian product $A \times B$. This implies that two relations R and S are equal if and only if they have the same domain A, the same codomain B and if $R = S$ as sets.

Example 2.9 To understand the importance of specifying domain and codomain, consider the sets

$$A = \{0\}, \qquad B = \{0, 1\} \tag{2.40}$$

and the relations $R : A \rightharpoonup B$ and $S : B \rightharpoonup A$ defined by the formulas

$$R = \{(0, 0)\}, \qquad S = \{(0, 0)\}. \tag{2.41}$$

Although $R = S$ as sets, they are different relations because their domains are different — and in fact so are also their codomains.

The sense in which $R : A \rightharpoonup B$ can be regarded as a transformation from A to B is the following: if aRb we can think of $a \in A$ as becoming $b \in B$. Observe that this is a very general notion of transformation: since we can have aRb for more than one b, we are allowing a to become many different things at the same time; and sice there could be no b such that aRb, we are also accounting for the fact that a disappears in the state B. On the other hand, $b \in B$ can be originated from more than one $a \in A$ in the transormation, or may be generated from nothing if there is no $a \in A$ such that aRb. This interpretation of relations should be kept in mind while reading the examples that follow.

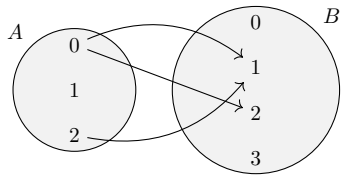
Example 2.10 (Finite relations) If $A = \{0, 1, 2\}$ and $B = \{0, 1, 2, 3\}$, the set

$$R = \{(0, 1),\ (0, 2),\ (2, 1)\} \tag{2.42}$$

is a binary relation $R : A \rightharpoonup B$ with domain A and codomain B because it is a subset of

$$A \times B = \{(0, 0),\ (0, 1),\ \ldots,\ (2, 2),\ (2, 3)\} \tag{2.43}$$

Observe that if both A and B are finite as in this case, so is $A \times B$; this implies that any relation $R : A \rightharpoonup B$ is finite, being a subset of a finite set. When this happens, we can represent R using a two-component graph as in the diagram below, with domain and codomain represented as Euler-Venn diagrams; we dra an arrow from $a \in A$ to $b \in B$ when aRb.



(2.44)

For computational purposes, however, a relation between finite sets is better represented by its *incidence matrix*. This is a matrix $M \in \mathrm{Mat}_{m\times n}(\Omega)$ where $m = |A|$ and $n = |B|$ are the cardinalities of the domain and codomain. We also fix an ordering on both A and B — which in our case is the usual order relation induced from $\mathbf{N}$— and then set $M_{ij} = 1 \Leftrightarrow a_i R b_j$, where a_i is the i-th element of A in the chosen order and likewise for b_j. In our case we have

$$M = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{2.45}$$

Here M has 3 rows because $|A| = 3$ and 4 columns because $|B| = 4$. Since $0R1$, we have that $0 \in A$, the first element of the domain, is in relation with $1 \in B$, the second element of the codomain; therefore we set $M_{12} = 1$. Likewise, the fact that $0R2$ means that 0, the first element of the domain, is in relation with 2, the third element of the codomain and therefore we set $M_{13} = 1$. Finally, $2R1$ implies that we must have $M_{32} = 1$. All other entries of M are then set to 0.

Example 2.11 (Extremal relations) Given any two sets A and B, among all possible relations $A \rightharpoonup B$ there are the *empty relation* $\varnothing \subseteq A \times B$ and the *total relation* $A \times B$. The first is denoteb by the symbol $\bot$, the second by $\top$. These two relations coincide only when $A \times B$ is empty, i.e. when either A or B are empty.

Example 2.12 (Subsets as reltions) If A is any set and $1 = \{0\}$ is a set with a single element, a relation $R : 1 \rightharpoonup A$ consists of pairs of type $(0, a)$ whose first coordinate is necessarily 0, the only element of 1 and whose second coordinate is an element of A. Thus, the elements of R are completely determined by the second coordinate and we can identify R with a subset of A, namely with

$$S = \{a \in A : (0, a) \in R\}. \tag{2.46}$$

Example 2.13 Let L be the propositional language generated by an alphabet A. Define a relation $R : L \rightharpoonup L$ setting $\varphi R \chi$ if φ and χ have a common subformula, i.e. if there exists a formula ψ which is a subformula of both φ and χ. If we assume that A has the standard set of connectives and variables $\{x, y\}$, then

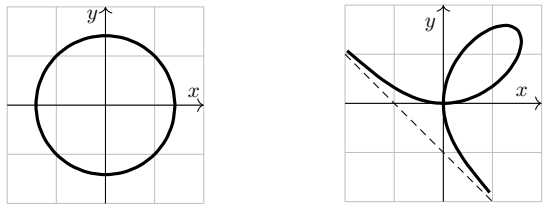
$$(\bot \to x, \bot \vee y) \in R, \qquad (x \to \top, y \to \bot) \notin R \tag{2.47}$$

because the first pair has $\bot$ as a common subformula and the second has none. Observe that if φ and χ have a common subformula ψ, they must also have a common atomic subformula, namely any of the atomic subformulas of ψ. The set of atomic subformulas of L is $\{\bot, \top, x, y\}$; thus $\varphi R \chi$ precisely when they have at least one of these subformulas in common.

Example 2.14 (Algebraic curves) Relations play an important role in geometry. If $\mathbf{R}$ is the field of real numbers and $p \in \mathbf{R}[x, y]$ is a polynomial in two variables with coeffcicients in $\mathbf{R}$, then p defines a relation $C : \mathbf{R} \rightharpoonup \mathbf{R}$, the *algebraic curve*

$$C = \big\{(x, y) \in \mathbf{R}^2 : p(x, y) = 0\big\}, \tag{2.48}$$

whose elements are the points of the cartesian plane $\mathbf{R}^2$ satisfying the equation $p(x, y) = 0$. The algebraic curve defined by the polynomials $p = x^2 + y^2 - 2$ and $p = x^3 + y^3 - 3xy$ are the unit circle and the folium of Descartes and are shown below.



(2.49)

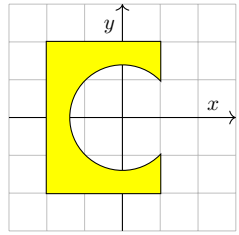
Example 2.15 (Semialgebraic sets) Inequalities over the real numbers define relations. Every set $\{p_1, p_2, ...\} \subseteq \mathbf{R}[x, y]$ of real polynomials in 2 variables defines a relation $R : \mathbf{R} \rightharpoonup \mathbf{R}$, the *semialgebraic variety*

$$R(p_1, p_2, ...) = \big\{(x, y) \in \mathbf{R}^2 : \forall i (p_i(x, y) \geq 0\big\}. \tag{2.50}$$

The weak inequalities can, of course, be replaced by strict inequalities. As an ezample, if we take

$$p_1 = 4 - x^2, \qquad p_2 = 4 - y^2, \qquad p_3 = 1 - x, \qquad p_4 = x^2 + y^2 - 2, \tag{2.51}$$

then $R(p_1, \dots, p_4) \subseteq \mathbf{R}^2$ is the set of points in the colored region below.



(2.52)

□ **The local structure.** Relations have a rich algebraic structure. To begin with, if we fix a domain A and a codomain B, the set of relations $A \rightharpoonup B$ is the powerset $P(A \times B)$ and therefore, from proposition 2.3, we immediately have

COROLLARY 2.16. The set of relations from A to B is a Boolean algebra with respect to union, intersection and complement, with minimum $\bot$ and maximum $\top$. □

Example 2.17 Given $A = \{1, 2, 3\}$ and $B = \{1, 2, 3, 4\}$, consider the relations $R, S : A \rightharpoonup B$, defined respectively by the incidence matrices

$$M(R) = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad M(S) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \tag{2.53}$$

We compute $(R \cap S')'$ by determining its incidence matrix as follows.

- The matrix of the intersection of two relations is the pointwise conjunction of the two matrices: $M(R \cap S) = M(R) \wedge M(S)$ where $[M(R) \wedge M(S)]_{ij} = M(R)_{ij} \wedge M(S)_{ij}$. This follows from the fact that $a_i (R \cap S) b_j \Leftrightarrow (a_i R b_j) \wedge (a_i S b_j)$.
- The matrix of the union is the pointwise disjunction of the two matrices: $M(R \cup S) = M(R) \vee M(S)$, where $[M(R) \vee M(S)]_{ij} = M(R)_{ij} \vee M(S)_{ij}$, because $a_i (R \cup S) b_j \Leftrightarrow (a_i R b_j) \vee (a_i S b_j)$.
- The matrix of the complement is the pointwise negation of the matrix: $M(R') = \neg M(R)$, where $[(\neg M(R)]_{ij} = \neg [M(R)_{ij}]$, because $a_i (R') b_j \Leftrightarrow \neg (a_i R b_j)$.

With these definitions we have

$$M((R \cap S')') = \neg M(R \cap S') = \neg (M(R) \wedge M(S')) = \neg (M(R) \wedge \neg M(S)) \tag{2.54}$$

$$= \neg \left[\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \wedge \neg \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \right] \tag{2.55}$$

$$= \neg \left[\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \wedge \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} \right] = \neg \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \tag{2.56}$$

$$= \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \tag{2.57}$$

The same result can be obtained more efficiently observing that, by corollary 2.16,

$$(R \cap S')' = R' \cup S'' = R' \cup S \tag{2.58}$$

and therefore

$$M((R \cap S')') = M(R' \cup S) = M(R') \vee M(S) = \neg M(R) \vee M(S) \tag{2.59}$$

$$= \neg \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \vee \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \vee \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \tag{2.60}$$

$$= \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}. \tag{2.61}$$

Example 2.18 Consider the relations $R, S : \mathbf{Z} \rightharpoonup \mathbf{Z}$ defined by the formulas

$$xRy := 6|(y - x), \qquad xSy := 10|(y - x), \tag{2.62}$$

where $x|y$ means that x is an integer divisor of y. We can not represent R and S using matrices because their domain and codomain are infinite. Instead, we have to argue abstractly as follows:

$$x(R \cap S)\, y \equiv (xRy) \wedge (xSy) \equiv (6|y - x) \wedge (10|y - x) \equiv 30|y - x \tag{2.63}$$

where the first equivalence follows from the definition of intersection of relations; the second from the definition of R and S; the last from the observation that a number is divisible by 6 and 10 when it is divisible by their least common multiple 30. Note, however, that $R \cup S$ is not so easily described

Example 2.19 Assume $R, S : A \rightharpoonup B$ are binary relations between finite sets with incidence matrices $M(R)$ and $M(S)$. Observe that

$$R \subseteq S \Leftrightarrow \forall ij(a_i R b_j \to a_i S b_j) \tag{2.64}$$
$$\Leftrightarrow \forall ij(M(R)_{ij} = 1 \to M(S)_{ij} = 1) \tag{2.65}$$
$$\Leftrightarrow \forall ij(M(R)_{ij} \leq M(S)_{ij}) \tag{2.66}$$

where the last condition follows from the fact that incidence matrices are only two valued, as they assume values in Ω, with $0 < 1$. When the last condition is satisfied we write $M(R) \leq M(S)$. For a concrete example, assume $A = \{0, 1, 2\}$ and $B = \{0, 1, 2, 3\}$. If the matrices of R and S are

$$M(R) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}, \qquad M(S) = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \tag{2.67}$$

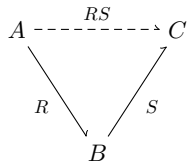
then $M(R) \leq M(S)$ and therefore $R \subseteq S$.

□ **The multiplicative structure.** The local structure is concerned with relations having fixed domain and codomain. However, ff relations describe becoming between states, we can consider successive transormations between different states and this leads to the definition of a product of relations.

Definition 2.20. A relation $R : A \rightharpoonup B$ is *composable* with a relation $S : B \rightharpoonup C$ if the codomain of R coincides with the domain of S. When this happens, the *product* of R and S is the relation $RS : A \rightharpoonup C$ defined by the formula

$$aRSc \leftrightarrow \exists b \in B.((aRb) \wedge (bRc)). \tag{2.68}$$

Explicitly, $aRSc$ if we can find at least one intermediate element $b \in B$ such that sRb and bSc. The situation is described in the diagram below.



(2.69)

The notation we have used to describe the product of relations is called *algebraic*. There is also an alternative *functional* notation $S \circ R$, read 'S *composed with* R', in which the position of the two relations is exchanged. The origin of this notation will become clear when we will discuss functions.

Example 2.21 (Matrix product) If R and S are composable relations between finite sets and can therefore be represented by matrices, we can compute the matrix of RS as follows. Using the notation from diagram 2.69 and assuming that $|B| = n$, we have

$$M(RS)_{ij} = 1 \Leftrightarrow a_i RSc_j \Leftrightarrow \exists b_k((a_i R b_k) \wedge (b_k S c_j)) \Leftrightarrow \exists k(M(R)_{ik} = 1) \wedge (M(S)_{kj} = 1) \tag{2.70}$$
$$\Leftrightarrow \exists k((M(R)_{ik} \wedge N(S)_{kj}) = 1) \Leftrightarrow \bigvee_{k=1}^{n} (M(R)_{ik} \wedge (N(S)_{kj}) = 1. \tag{2.71}$$

Since incidence matrices are only two-valued, this impies that

$$M(RS)_{ij} = \bigvee_{k=1}^{n} (M(R)_{ik} \wedge (N(S)_{kj}). \tag{2.72}$$

Note that this is the usual matrix multiplication from linear algebra, except that the field operations of multiplication and addition have been replaced respectively by conjunction and disjunction in Ω.

Example 2.22 For a concrete example of product in the finite case, assume $A = \{0, 1, 2\} = B$, $C = \{0, 1\}$ and the relations R and S are represented by the matrices

$$M(R) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad M(S) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{2.73}$$

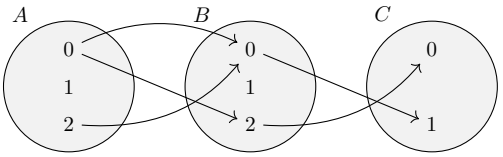
We compute explicitly a few elements of the product matrix using formula (2.72):

$$M(RS)_{11} = [M(R)_{11} \wedge M(S)_{11}] \vee [M(R)_{12} \wedge M(S)_{21}] \vee [M(R)_{13} \wedge M(S)_{31}] \tag{2.74}$$
$$= [1 \wedge 0] \vee [0 \wedge 0] \vee [1 \wedge 1] = 0 \vee 0 \vee 1 = 1 \tag{2.75}$$
$$M(RS)_{21} = [M(R)_{21} \wedge M(S)_{11}] \vee [M(R)_{22} \wedge M(S)_{21}] \vee [M(R)_{23} \wedge M(S)_{31}] \tag{2.76}$$
$$= [0 \wedge 0] \vee [0 \wedge 0] \vee [0 \wedge 1] = 0 \vee 0 \vee 0 = 0. \tag{2.77}$$

When all row by column products are expanded we find that the product matrix is

$$M(RS) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.78}$$

The graphs of R and S are shown below. Observe that if $a \in A$ and $c \in C$, then $aRSc$ precisely when there exists at least one path starting from a and ending in c that goes through some element $b \in B$.



(2.79)

When the relations are defined analitically, though, computing their product has to be done on ad hoc basis.

Example 2.23 Consider the relation $R, S : \mathbf{Z} \rightharpoonup \mathbf{Z}$ defined by the formulas

$$xRy := 2|(y - x), \qquad ySz := 3|(z - y)\,. \tag{2.80}$$

Observe that R and S are composable, because the codomain of R and the domain of S both coincide with $\mathbf{Z}$. We will prove that RS is the total relation on $\mathbf{Z}$, i.e. that $xRSz$ for all $x, z \in \mathbf{Z}$. To see this observe first that we ca rewrite R and S in a different form as follows:

$$xRy \Leftrightarrow 2|(y - x) \Leftrightarrow \exists m \in \mathbf{Z}.(y - x = 2m) \Leftrightarrow \exists m \in \mathbf{Z}.(y = x + 2m) \tag{2.81}$$
$$ySz \Leftrightarrow 3|(z - y) \Leftrightarrow \exists n \in \mathbf{Z}.(z - y = 3n) \Leftrightarrow \exists n \in \mathbf{Z}.(z = y + 3n) \tag{2.82}$$

Putting these two remarks together we find that

$$xRSz \Leftrightarrow \exists y(xRy \wedge ySz) \Leftrightarrow \exists ymn((x + 2m = y) \wedge (y + 3n = z)) \tag{2.83}$$
$$\Leftrightarrow \exists mn(x + (2m + 3n) = z) \Leftrightarrow \exists mn(z - x = (2m + 3n))\,. \tag{2.84}$$

Now observe that $2 \cdot (-1) + 3 \cdot 1 = 1$, so that for every $k \in \mathbf{Z}$ we have

$$2 \cdot (-k) + 3k = (2 \cdot (-1) + 3 \cdot 1)\, k = 1 \cdot k = k. \tag{2.85}$$

Thus, given any $x, z \in \mathbf{Z}$ set $k = z - x$; if $m = -k$ and $n = k$ we have

$$2m + 3n = 2 \cdot (-k) + 3k = k = z - x \tag{2.86}$$

so that every pair (x, z) belongs to RS.

DEFINITION 2.24. Given any set A, the *identity relation* on A is the relation $I : A \rightharpoonup A$ defined by the formula

$$I = \{(x, y) \in A^2 : x = y\}. \tag{2.87}$$

Because the identity relation is defined by equality, it is also called the *equality relation* on A; also, when A is finite the matrix representing I is the diagonal identity matrix; therefore the identity relation is also called the *diagonal relation* on A. Observe that every set A has its own identity relation; therefore, when confusion may arise we will write I_A to indicate the identity relation on A. Other symbols that we will use for the identity relation on A are A, $=$, Δ or 1, with a subscript A when necessary.

Example 2.25 Assume $A = \{a_0, \dots, a_n\}$ is a finite set. For the identity relation I on A we have

$$M(I)_{ij} = 1 \Leftrightarrow a_i I a_j \Leftrightarrow a_i = a_j \Leftrightarrow i = j \tag{2.88}$$

where the last equivalence assumes that we have given a reduced representation of A, with no repetitions of elements. Thus, the diagonal elements of $M(I)$ are all equal to 1 and all others are 0. For example, for $A = \{0, 1, 2\}$ we have

$$M(I) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.89}$$

PROPOSITION 2.26. Sets and relations form a category **Rel**. In other words, the following formulas hold whenever the relevant products are defined.

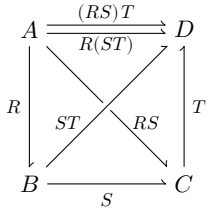
$(RS)T = R(ST)$	Associativity	(2.90)
$IR = R$	Left neutral element	(2.91)
$RI = R$	Right neutral element	(2.92)

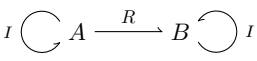
Proof. We prove associativity — see the diagram on the left in (2.95). Given $a \in A$ and $d \in D$, we have

$$a(RS)Td \equiv \exists c[(aRSc) \wedge (cTd)] \equiv \exists bc[((aRb) \wedge (bSc)) \wedge (cTd)] \tag{2.93}$$

$$\equiv \exists bc[(aRb) \wedge ((bSc) \wedge (cTd))] \equiv \exists b[(aRb) \wedge (bSTd)] \equiv aR(ST)\,d \tag{2.94}$$

where the first and second equivalences follow from the definition of product — here $b \in B$ and $c \in C$; the third is associativity of conjunction and the remaining follow again from the definition of product relation. Thus, $(RS)T$ and $R(ST)$ have the same elements and are equal by the extensionality axiom.





(2.95)

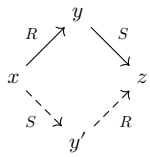
We prove the left neutral element formula – see the diagram on the right in (2.95). Note that in this case I is the identity relation on the domain A of R, whereas in (2.92) I is the identity on the codomain B. Given $a \in A$ and $b \in B$ we have

$$a(IRb) \equiv \exists a'((aIa') \wedge (a'Rb)) \equiv \exists a'((a = a') \wedge (a'Rb)) \equiv aRb \tag{2.96}$$

we the first equivalence follows from the definition of product relation, the second from the definition of identity relation and the third from the fact that a and a' are equal. The proof of right neutral element formula is similar and is omitted. □

Although proposition 2.26 exhibits similarities between the product of relations and products in numerical sets like **N**or **Z**, the analogy is only superficial. In fact other elementary properties of numerical sets do not hold for relations.

Example 2.27 The product of relations is not commutative: the formula $RS = SR$ does not hold in general, although it can be true for particular choices of R and S. The first problem is that if $R : A \rightharpoonup B$ and $S : B \rightharpoonup C$ so that $RS : A \rightharpoonup C$ is defined, we need $C = A$ in order to define $SR : B \rightharpoonup B$ and then $A = B = C$ in order to compare RS and SR. However, even if $R, S : A \rightharpoonup A$, commutativity may fail. For suppose $xRSz$, so that xRy and ySz for some $y \in A$ as in the upper part of diagram 2.97.



(2.97)

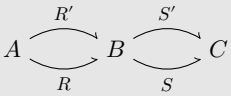
If $RS = SR$ then $xSRz$ and there must exists y' as in the lower part of the digarm such that xSy' and $y'Rz$. If this condition fails even for a single pair (x, z) then $RS \neq SR$. For a concrete case, take $A = \{0, 1\}$ and $R, S : A \rightharpoonup A$ defined by the incidence matrices

$$M(R) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad M(S) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.98}$$

This is an instance of diagram 2.97 with $x = 0$ and $y = z = 1$. Since that $0R1$ e $1S1$, we have a path as in the upper part of the diagram. However, 0 is not in S-relation with anything and therefore $RS \neq SR$. In fact,

$$RS = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad SR = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \tag{2.99}$$

Proposition 2.28. The product of relations is compatible with inclusions, in the sense that given relations as in the diagram below,



(2.100)

if $R \subseteq R'$ and $S \subseteq S'$ then $RS \subseteq R'S'$.

Proof. For $a \in A$ and $c \in C$ we have

$$aRSc \Rightarrow \exists b(aRb \wedge bSc) \Rightarrow \exists b(aR'b \wedge bS'c) \Rightarrow aR'S'c \tag{2.101}$$

where the first and third implications follow from the definition of product and the second from the assumptions $R \subseteq R'$ and $S \subseteq S'$. This proves that every element of RS is also an element of $R'S'$ and therefore that $RS \subseteq R'S'$. □

PROPOSITION 2.29. (COMPATIBILITY OF THE PRODUCT WITH THE BOOLEAN STRUCTURE) The product of relations is compatible with the Boolean structure of $A \times B$, in the sense that the following formulas hold when the operations are defined.

$$\bot R = \bot \qquad R\bot = \bot \tag{2.102}$$

$$\top R \subseteq \top \qquad R\top \subseteq \top \tag{2.103}$$

$$\left(\bigcup_{i\in I} R_i\right) S = \bigcup_{i\in I} R_i S \qquad R\left(\bigcup_{i\in I} S_i\right) = \bigcup_{i\in I} RS_i \tag{2.104}$$

$$\left(\bigcap_{i\in I} R_i\right) S \subseteq \bigcap_{i\in I} R_i S \qquad R\left(\bigcap_{i\in I} S_i\right) \subseteq \bigcap_{i\in I} RS_i \tag{2.105}$$

Proof. We only prove the formulas in the left column; proofs for the right columns are similar and therefore omitted. ① Bottom. Observe that

$$a(\bot R)\, c \equiv \exists b(a\bot b \wedge bRc) \equiv \bot \equiv a\bot c \tag{2.106}$$

where the first equivalence follows from the definition of product relation; the remaining two from the fact that $a\bot b$ and $a\bot c$ are always false for the empty relation — the symbol in the third formula is falsehood from PL. Thus $\bot R = \bot$ by extensionality. ② Top. Since $\top$ is the total relation it contains any relation and in particular $\top R$. ③ Union. We have

$$a\left[\left(\bigcup R_i\right) S\right] c \Leftrightarrow \exists b\left[a\left(\bigcup R_i\right) b \wedge bSc\right] \Leftrightarrow \exists b[(\exists i.aR_ib) \wedge bSc] \Leftrightarrow \exists b\exists i(aR_ib \wedge bSc) \tag{2.107}$$

$$\Leftrightarrow \exists i\exists b(aR_ib \wedge bSc) \Leftrightarrow \exists i(aR_iSc) \Leftrightarrow a\left[\bigcup(R_iS)\right] c \tag{2.108}$$

where the first equivalence follows from the definition of product; the second from the definition of union; the third from distributivity of conjunction over disjunction; the fourth from commutativity of existential quantifiers; the fifth from the definition of product; the last from the definition of union. ④ Intersection.

$$a\left[\left(\bigcap R_i\right) S\right] c \Leftrightarrow \exists b\left[a\left(\bigcap R_i\right) b \wedge bSc\right] \Leftrightarrow \exists b[\forall i(aR_ib) \wedge bSc] \Leftrightarrow \exists b[\forall i(aR_ib \wedge bSc)] \tag{2.109}$$

$$\Rightarrow \forall i(aR_iSc) \Leftrightarrow a\left[\bigcap(R_iS)\right] c \tag{2.110}$$

where the first equivalence follows from the definition of product; the second from the definition of intersection; the third from properties of the conjunction — in infinitary form. Note that the fourth step is only an implication: this is because if aR_iSc for every i, we might have different elements $b_i \in B$ such that aR_ib_i and b_iSc and not necessarily the same, as on the left hand side. The last equivalence follows from the definition of intersection. □

Example 2.30 The formulas in the two columns can be combined. Suppose we have sets of relations $R_i : A \rightharpoonup B$ and $S_j : B \rightharpoonup C$, where $i \in I$ and $j \in J$. Then

$$\left(\textstyle\bigcap_i R_i\right)\left(\bigcap_j S_j\right) \subseteq \bigcap_i\left[R_i\left(\bigcap_j S_j\right)\right] \subseteq \bigcap_i\left(\bigcap_j R_iS_j\right) = \bigcap_{ij} R_iS_j \tag{2.111}$$

where the two inclusions follow from the first and second formula respectively in (2.105). Likewise, from the formulas (2.104), we have

$$\left(\textstyle\bigcup_i R_i\right)\left(\bigcup_j S_j\right) = \bigcup_i\left[R_i\left(\bigcup_j S_j\right)\right] = \bigcup_i\left(\bigcup_j R_iS_j\right) = \bigcup_{ij} R_iS_j \tag{2.112}$$

The inclusions in formulas (2.103) and (2.105) can be proper.

Example 2.31 To see that the inclusion in formula (2.103) can be proper it suffices to take a nonempty set A and consider relations $\bot, \top : A \rightharpoonup A$. Then $\top \neq \bot$ and by (2.102) we have $\top\bot = \bot \subset \top$. On a more analytical side, observe that

$$\top R = \top \Leftrightarrow \forall ac(a\top Rc) \Leftrightarrow \forall ac\exists b(a\top b \wedge bRc) \Leftrightarrow \forall c\exists b(bRc) \tag{2.113}$$

where the first condition follows from extensionality, as every pair of elements belongs to $\top$; the second from the definition of product; the last from the fact that $a\top b$ is true for all $a \in A$ and $b \in B$ as $\top$ is the total relation. Thus $\top R = \top$ only when R satisfies the latter condition, in which case we say that R is op-serial.

Example 2.32 To see that the inclusion in formula (2.105) can be proper it suffices to exhibit a case in which the implication in (2.110) can not be inverted because the b_i's are different, and we can already achieve this when $|I| = 2$. In detail, let $A = \{0\} = C$, $B = \{0, 1\}$ and let $R1, R_2 : A \rightharpoonup B$ and $S : B \rightharpoonup C$ be defined by

$$R_1 = \{(0, 0)\}, \qquad R_2 = \{(0, 1)\}, \qquad S = \{(0, 0), (1, 0)\} = \top. \tag{2.114}$$

Then

$$(R_1 \cap R_2)\, S = \bot S = \bot, \qquad R_1 S \cap R_2 S = \top \cap \top = \top, \tag{2.115}$$

and the inclusion is proper.

The presence of a product also allows the definition of powers for a relation $R : A \rightharpoonup A$.

Definition 2.33. If $R : A \rightharpoonup A$ is a relation, its powers with natural exponent $n \in \mathbf{N}$ are recursively defined by the formulas

$$R^0 = 1, \qquad R^{n+1} = R^n R. \tag{2.116}$$

Thus, when $n > 0$ the power R^n is simply the product of n copies of R. We collect the basic properties of powers.

Proposition 2.34. If $R : A \rightharpoonup A$ is a relation and $m, n \in \mathbf{N}$ are natural numbers, the following formulas hold.

$$R^m R^n = R^{m+n} \tag{2.117}$$

$$(R^m)^n = R^{mn} \tag{2.118}$$

Proof. ① By induction on n. If $n = 0$ we have

$$R^m R^0 = R^m 1 = R^m = R^{m+0} \tag{2.119}$$

where the first equality follows from the recursive definition; the second from the fact that the identity on A is neutral for the product; the last from the fact that 0 is the neutral element for addition in $\mathbf{N}$. Assume the formula holds for some given n, i.e. that $R^m R^n = R^{m+m}$ for that particular n. Then

$$R^m R^{n+1} = R^m(R^n R) = (R^m R^n)\, R = R^{m+n} R = R^{m+n+1} \tag{2.120}$$

where the first equality follows from the recursive definition of powers; the second from associativity of the product;; the third from the inductive hypothesis; the last from the definition of powers. ② Again, by induction on n. For $n = 0$ we have

$$(R^m)^0 = 1 = R^0 = R^{m0} \tag{2.121}$$

where the first and second equalities follow from the definition of powers; the last from the fact that 0 is absorbing for the product in $\mathbf{N}$. For the inductive step, assume the formula holds for a specific n. Then

$$(R^m)^{n+1} = (R^m)^n R^m = R^{mn} R^m = R^{mn+m} = R^{m(n+1)} \tag{2.122}$$

where the first formula follows from the definition of powers; the second from the inductive hypothesis; the third from formula (2.117); the last from arithmetical properties of $\mathbf{N}$. □

□ The involutory structure.

DEFINITION 2.35. Given a relation $R : A \rightharpoonup B$, its *opposite* relation is the relation $R^{\mathrm{op}} : B \rightharpoonup A$ defined by the formula

$$bR^{\mathrm{op}}a \Leftrightarrow aRb. \tag{2.123}$$

Example 2.36 (THE MATRIX OF THE OPPOSITE RELATION) If $|A| = m$ and $|B| = n$ are finite then $R : A \rightharpoonup B$ is determined by its incidence matrix $M(R) \in \mathrm{Mat}_{m\times n}(\Omega)$. The opposite relation $R^{\mathrm{op}} : B \rightharpoonup A$ is then represented by a matrix $M(R^{\mathrm{op}}) \in \mathrm{Mat}_{n\times m}(\Omega)$. Furthermore,

$$M(R^{\mathrm{op}})_{ij} = 1 \Leftrightarrow b_i R^{\mathrm{op}} a_j \Leftrightarrow a_j R b_i \Leftrightarrow M(R)_{ji} = 1. \tag{2.124}$$

Since these matrices are only ywo-valued, this implies that

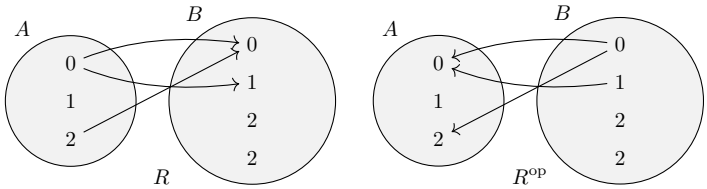
$$M(R^{\mathrm{op}})_{ij} = M(R)_{ji} \tag{2.125}$$

which, in linear algebra terminology, means that $M(R^{\mathrm{op}})$ is the transposed matrix of $M(R)$; in symbols $M(R^{\mathrm{op}}) = M(R)^{\mathrm{T}}$.

Example 2.37 If $A = \{0, 1, 2\}$, $B = \{0, 1, 2, 3\}$ and $R = \{(0, 0), (0, 1), (2, 0)\}$, then $R^{\mathrm{op}} = \{(0, 0), (1, 0), (0, 2)\}$, with swapped coordinates. The matrices of R and R^{op} are, respectively,

$$M(R) = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad M(R^{\mathrm{op}}) = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2.126}$$

As indicated in example 2.36, $M(R^{\mathrm{op}})$ is obtained by transposing $M(R)$, i.e. the rows of $M(R)$ become the columns of $M(R^{\mathrm{op}})$. It is also very suggestive to display the graph of both relations: since $bR^{\mathrm{op}}a \Leftrightarrow aRb$, the arrows in as the graph of R^{op} are obtained by reversing the direction of the arrows of R as shown in diagram 2.127.



(2.127)

The reason for which the opposite operation is called involutory is the following proposition, whose proof is straightformard and is omitted:

PROPOSITION 2.38. For every relation $R : A \rightharpoonup B$ the following formula holds:

$$(R^{\mathrm{op}})^{\mathrm{op}} = R. \tag{2.128}$$

We will often omit the parenthesis and write $R^{\mathrm{op}\,\mathrm{op}}$ for the double opposite.

PROPOSITION 2.39. The opposite operation is compatible with inclusion, in the sense that the following formula holds for relations $R, S : A \rightharpoonup B$,

$$R \subseteq S \Rightarrow R^{\mathrm{op}} \subseteq S^{\mathrm{op}}. \tag{2.129}$$

Proof. It suffices to observe that

$$bR^{\mathrm{op}}a \Rightarrow aRb \Rightarrow aSb \Rightarrow bS^{\mathrm{op}}a \tag{2.130}$$

where the first implication follows from the definition of opposite relation; the second from the assumption that R is contained in S; the last from the definition of opposite relation. Thus, every element of R^{op} belongs to S^{op} and $R^{\mathrm{op}} \subseteq S^{\mathrm{op}}$. □

PROPOSITION 2.40. The opposite operation is compatible with the Boolean structure of relations $A \rightharpoonup B$ in the sense that the following formulas hold.

$$\bot^{\mathrm{op}} = \bot \qquad \top^{\mathrm{op}} = \top \tag{2.131}$$

$$\left(\bigcup_{i\in I} R_i\right)^{\mathrm{op}} = \bigcup_{i\in I} R_i^{\mathrm{op}} \qquad \left(\bigcap_{i\in I} R_i\right)^{\mathrm{op}} = \bigcap_{i\in I} R_i^{\mathrm{op}} \tag{2.132}$$

$$(R')^{\mathrm{op}} = (R^{\mathrm{op}})' \tag{2.133}$$

Proof. ① For the empty relation we have

$$b\bot^{\mathrm{op}}a \Leftrightarrow a\bot b \Leftrightarrow \bot \Leftrightarrow b\bot a \tag{2.134}$$

where the first equivalence follows from the definition of opposite relation; the second and third from the definition of empty relation — the third formula is falsehood from PL. By extensionality, $\bot^{\mathrm{op}} = \bot$. ② For the total relation we have

$$b\top^{\mathrm{op}}a \Leftrightarrow a\top b \Leftrightarrow \top \leftrightarrow b\top a \tag{2.135}$$

where the first equivalence follows from the definition of opposite; the second and third from the fact that every pair belongs to the total relation — the third formula is truth of PL. ③ We have

$$b\Big(\bigcup R_i\Big)^{\mathrm{op}} a \Leftrightarrow a\Big(\bigcup R_i\Big) b \Leftrightarrow \exists i.(aR_ib) \Leftrightarrow \exists i.(bR_i^{\mathrm{op}}a) \Leftrightarrow b\bigcup(R_i^{\mathrm{op}})\, a \tag{2.136}$$

where the first equivalence follows from the definition of opposite relation; the second from the definition of union; the third from the definition of opposite relation; the last again from the definition of union. By extensionality, $(\bigcup R_i)^{\mathrm{op}} = \bigcup R_i^{\mathrm{op}}$. ④ Similarly,

$$b\Big(\bigcap R_i\Big)^{\mathrm{op}} a \Leftrightarrow a\Big(\bigcap R_i\Big) b \Leftrightarrow \forall i.(aR_ib) \Leftrightarrow \forall i.(bR_i^{\mathrm{op}}a) \Leftrightarrow b\Big(\bigcap R_i^{\mathrm{op}}\Big) a \tag{2.137}$$

where the first equivalence follows from the definition of opposite relation; the second from the definition of intersection; the third from the definition of opposite relation; the last again from the definition of intersection. By extensionality, $(\bigcap R_i)^{\mathrm{op}} = \bigcap R_i^{\mathrm{op}}$. ⑤ We have

$$b(R')^{\mathrm{op}}\, a \Leftrightarrow aR'b \Leftrightarrow \neg(aRb) \Leftrightarrow \neg(bR^{\mathrm{op}}a) \Leftrightarrow b(R^{\mathrm{op}})'\, a \tag{2.138}$$

where the first equivalence follows from the definition of opposite relation; the second from that of complement; the third again from the definition of opposite relation; the last from the definition of opposite. By extensionality, (2.133) holds. □

PROPOSITION 2.41. The opposite relation operation is compatible with the multiplicative structure, in the sense that the following formulas hold.

$$1^{\mathrm{op}} = 1, \qquad (RS)^{\mathrm{op}} = S^{\mathrm{op}} R^{\mathrm{op}}. \tag{2.139}$$

Proof. ① Assume $1 : A \rightharpoonup A$ is the identity relation on A. We have

$$a_1 1^{\mathrm{op}} a_2 \Leftrightarrow a_2 1 a_1 \Leftrightarrow a_2 = a_1 \Leftrightarrow a_1 = a_2 \Leftrightarrow a_1 1 a_2 \tag{2.140}$$

where the first equivalence follows from the definition of opposite; the second from the definition of identity; the third from properties of equality; the last again from the definition of identity. ② Assume $R : A \rightharpoonup B$ and $S : B \rightharpoonup C$. Then

$$c(RS)^{\mathrm{op}}\, a \Leftrightarrow aRSc \Leftrightarrow \exists b.(aRb \wedge bSc) \Leftrightarrow \exists b.(bSc \wedge aRb) \Leftrightarrow \exists b.(cS^{\mathrm{op}} b \wedge bR^{\mathrm{op}} a) \tag{2.141}$$

$$\Leftrightarrow cS^{\mathrm{op}} R^{\mathrm{op}} a \tag{2.142}$$

where the first equivalence follows from the definition of opposite; the second from the definition of product; the third from commutativity of conjunction; the fourth from the definition of opposite; the last from definition of product. □

It might be tempting to think of the opposite relation R^{op} as some sort of inverse process to R, in the sense that if we regard $R : A \rightharpoonup B$ as a tranformation from the state A to the state B, then $R^{\mathrm{op}} : B \rightharpoonup A$ ought to be the reverse transformation that restores the initial state when applied after R. This is not the case, however. To start with, the transformation on a state A that does not alter the state is the identity relation A; thus the claim that R^{op} restores the initial state would transtale into the formula $RR^{\mathrm{op}} = A$ And, in the opposite direction, $R^{\mathrm{op}} R = B$. This, however, does not happen even in simple cases.

Example 2.42 Let $A = \{0, 1\}$ and $R : A \rightharpoonup A$ be the relation with incidence matrix

$$M(R) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{2.143}$$

Observe that $M(R^{\mathrm{op}}) = M(R)^{\mathrm{T}} = M(R)$, so that $R^{\mathrm{op}} = R$. Moreover,

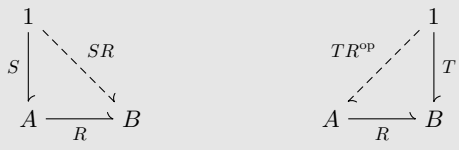
$$M(R^2) = M(R)\, M(R) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = M(R), \tag{2.144}$$

so that $R^2 = R$. Therefore,

$$RR^{\mathrm{op}} = RR = R \neq A, \qquad R^{\mathrm{op}} R = RR = R \neq B. \tag{2.145}$$

□ **Direct and inverse images.** The constructions of product and opposite relations provide a useful tool to transfer structure from one set to another. We have seen in example 2.12 how a subset $S \subseteq A$ can be identified with a relation $S : 1 \rightharpoonup A$; we will use this identification in the following.

DEFINITION 2.43. Assume $R : A \rightharpoonup B$ is a binary relation. The *direct image* of a subset $S \subseteq A$ is the subset $R_*(S) := SR \subseteq B$. In the opposite direction, the *inverse image* of a subset $T \subseteq B$ is the subset $R^*(T) := TR^{\mathrm{op}} \subseteq A$.



(2.146)

We can provide a more explicit description of the direct image observing that

$$b \in R_*(S) \Leftrightarrow 0SRb \Leftrightarrow \exists a(0Sa \wedge aRb) \Leftrightarrow \exists a \in S(aRb)\,. \tag{2.147}$$

In other words, the direct image of S consits of the elements of B coming from elements of S via R. Likewise, for the inverse image we have

$$a \in R^*(T) \Leftrightarrow 0TR^{\mathrm{op}}A \Leftrightarrow \exists b(0Tb \wedge bR^{\mathrm{op}}a) \Leftrightarrow \exists b \in T(aRb) \tag{2.148}$$

so that the inverse image consists of those lements of A that R senns to elements of T. When $S = \{a\}$ consists of a single element, it is customary to write $R_*(a)$ instead of $R_*(\{a\})$ and call this set the direct image of a. Likewise, if $T = \{b\}$ consists of a single element, we write $R^*(b)$ instead of $R^*(\{b\})$ and call this the inverse image of b along R.

Example 2.44 (Divisibility on $\mathbf{N}$) The divisibility relation $| : \mathbf{N} \rightharpoonup \mathbf{N}$ is defined by the formula

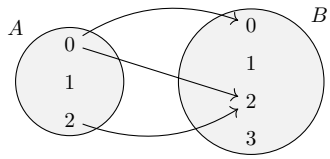
$$x|y \Leftrightarrow \exists z(xz = y)\,. \tag{2.149}$$

The formula $x|y$ is read '*x divides y*'. Since both the domain and codomain are infinite, the relation can not be represented by a graph or an incidence matrix. We can only describe the relation anlytically. Observe that x divides y precisely when y is a multiple of x. Thus the direct image of 0 is the set $\{0\}$ consisting of zero alone. The direct image of 1 is the whole set $\mathbf{N}$ of natural numbers. The direct image of 2 is the set $\{0, 2, 4, ...\}$ of even numbers and likewise for the other elements. In the opposite direction, the inverse image of 0 is $\mathbf{N}$, because $x0 = 0$ and therefore $x|0$ for every x. The inverse image of 1 is $\{1\}$, because $xz = 1$ implies $x = 1$. The inverse image of 2 is $\{1, 2\}$ and in general the inverse image of y is the set of natural divisors of y.

Example 2.45 Let $A = \{0, 1, 2\}$, $B = \{0, 1, 2, 3\}$ and let $R : A \rightharpoonup B$ be the relation defined by the matrix

$$M(R) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.150}$$

Since domain and codomain are finite, we can represent R using its graph.



(2.151)

The direct images of the elements of the domain are

$$R_*(0) = \{0, 2\}, \qquad R_*(1) = \varnothing, \qquad R_*(2) = \{2\} \tag{2.152}$$

and the inverse images of the elements of the codomain are

$$R^*(0) = \{0\}, \qquad R^*(1) = \varnothing, \qquad R^*(2) = \{0, 2\}, \qquad R^*(3) = \varnothing. \tag{2.153}$$

Since $a_i Rb_j \Leftrightarrow M(R)_{ij} = 1$, the direct image of $a_i \in A$ can be read from the i-th row of the matrix: it consists of those $b_j \in B$ which have a 1 in the row. Dually, the inverse image of b_j can be read from the j-th column of the matrix by picking the a_i's which have a 1 in that column.

Direct and inverse images play a fundemantal role in isolating the concept of function and in examining its properties.

2.3 Functions

□ **Definition and examples.** Although functions are a special class of relations, they are of particular interest for at least two reasons: they have been studied before and more extensively than general relations and they can be used to examne in more depth the structure of relations. The definition of function and many of its properties depend on the comcept of direct and inverse images along a relation.

DEFINITION 2.46. A relation $f : A \rightharpoonup B$ is a *function* if the direct image of every element $a \in A$ has a single element.

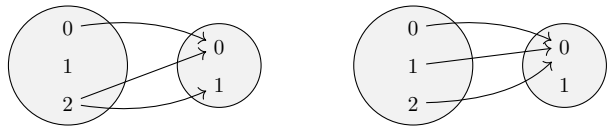
$$\forall a \in A(|f_*(a)| = 1)\,. \tag{2.154}$$

Traditionally, there is a special notation for functions: if f is a function with domain A and codomain B we write $f : A \to B$ with a full arrow, instead of $f : A \rightharpoonup B$ with half an arrow; if $a \in A$ and $b \in B$ is the only element in the direct image of a along f, we write $f(a) = b$ instead of afb. It should also be noted that whereas we tend to use capital letters R, S, ... for relations, lowercase letters like f, g, ... are more common for functions. We will write A^B or $[A, B]$ for the set of all functions from A to B.

Example 2.47 Let $A = \{0, 1, 2\}$ and $B = \{0, 1\}$. The matrices below describe two relations $A \rightharpoonup B$.

$$M(R) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix} \qquad M(S) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \tag{2.155}$$

Recall from example 2.45 that the direct image of $a_i \in A$ can be read off the i-th row of the incidence matrix; in particular, the cardinality of the direct image of a_i is the number of entries equal to 1 in the i-th row. Thus, for a function every row of the incidence matrix should have a single element equal to 1. Only $M(S)$ satisfies this condition. $M(R)$ fails on two counts: the second row has no element equal to 1 and the third row has two. Thus only S is a function. The graphs of the relations are shown below.



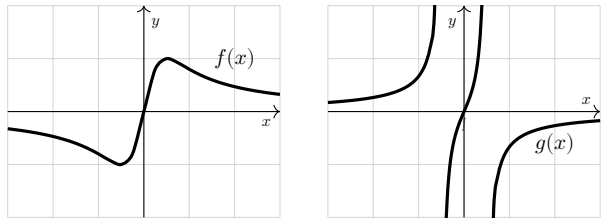
(2.156)

Observe that, from the point of view of the graph, the fact that the direct image of every element of the domain has a single element means that every element of the domain should be the source of a single arrow in the graph.

Example 2.48 Functions are often assigned in mathematics by their analitical expression. For example, consider the formulas

$$f(x) = \frac{x}{1 + x^2}, \qquad g(x) = \frac{x}{1 - x^2}. \tag{2.157}$$

The first defines a function $f : \mathbf{R} \to \mathbf{R}$. In fact what the formula does is to indicate that given any $x \in \mathbf{R}$ the only element in its direct image is $y := x/(1 + x^2)$ which can uniquely and unambiguously computed. The second formula does not define a function $g : \mathbf{R} \to \mathbf{R}$ because we can not compute $g(1) = 1/0$ and $g(-1) = -1/0$ and therefore the direct images of 1 and -1 are empty. Note, however, that we could change the domain of g to the set $D := \mathbf{R} \setminus \{\pm 1\}$ and in this case the second formula defines a function $g : D \to \mathbf{R}$.



(2.158)

Continuing with the idea that a binary relation can be regarded as a transformation from the state represented by the domain to the state of the codomain, we can think of functions as transformations subject to the condition that no element of the domain can either vanish or split; it is tranformed, possibly with contributions from other elements, into a single, well determined entity.
Functions can be characterized using the product of relations.

PROPOSITION 2.49. If $R : A \rightharpoonup B$ is a binary relation, then

1. $R_*(a)$ has at least one element for every $a \in A$ if and only if $A \subseteq RR^{\mathrm{op}}$;
2. $R_*(a)$ has at most one element for every $a \in A$ if and only if $R^{\mathrm{op}}R \subseteq B$.

Proof. ① Observe that

$$A \subseteq RR^{\mathrm{op}} \Leftrightarrow \forall a(aRR^{\mathrm{op}}a) \Leftrightarrow \forall a \exists b(ARb \wedge bR^{\mathrm{op}}a) \Leftrightarrow \forall a \exists b(ARb) \tag{2.159}$$

where the first equivalence follows from the definition of the irentity relation A; the second from the definition of product; the third from the fact that $bR^{\mathrm{op}}a$ is equivalent to aRb by the definition of opposite relation and from idempotency of conjunction. The last formula states precisely the fact that every $a \in A$ has at least a $b \in B$ in its direct image along R. ② In this case

$$R^{\mathrm{op}}R \subseteq B \Leftrightarrow \forall bb'(bR^{\mathrm{op}}Rb' \to b = b') \Leftrightarrow \forall bb'(\exists a(bR^{\mathrm{op}}a \wedge aRb') \to b = b') \tag{2.160}$$

$$\Leftrightarrow \forall abb'((aRb \wedge aRb') \to b = b') \tag{2.161}$$

where the furst equivalence follows from the definition of the identity relation B; the second from the definition of product; the third from properties of quantifiers and the definition of opposite relation. The last formula states that if b and b' belong to the direct image of some element a, they must be equal, i.e. that every direct image has at most one element. □

Since a function is a relation for which all elements of the domain have exactly one element in their direct image, we immediately obtain

COROLLARY 2.50. A relation $R : A \rightharpoonup B$ is a function if and only if

$$A \subseteq RR^{\mathrm{op}}, \qquad R^{\mathrm{op}}R \subseteq B. \tag{2.162}$$

□

Reflection and coreflection of relations into functions

□ **The category of functions.** Functions do not inherit all the structure of relations. If we fix two sets A and B, the set of all relations $A \rightharpoonup B$ is a Boolean algebra (corollary 2.16); however, the subset of all functions $A \to B$ is not.

Example 2.51 Let $A = \{0, 1\}$ and $B = \{0, 1, 2\}$ and fonsider the functions $f, g : A \to B$ represented by the matrices

$$M(f) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad M(g) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}. \tag{2.163}$$

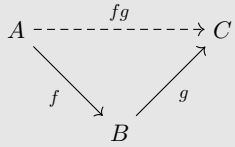
Since every function is, in particular, a relation, we can form the relations $f \cup g$, $f \cap g$ and f', whose matrices are

$$M(f \cup g) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \quad M(f \cap g) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad M(f') = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \tag{2.164}$$

and none of them is the matrix of a function.

What is preserved is part of the multiplicative structure of relations.

Proposition 2.52. Suppose f and g in the diagram below are functions. Then the product relation fg is also a function.



Proof. We use corollary 2.50. First observe that

$$A \subseteq ff^{\mathrm{op}} = fBf^{\mathrm{op}} \subseteq fgg^{\mathrm{op}}f^{\mathrm{op}} = (fg)\,(fg)^{\mathrm{op}} \tag{2.165}$$

where the firs step follows from the fact that f is a function and from the corollary; the second from the fact that the identity on B is neutral for the product of relations (proposition 2.26); the third from the fact that g is a function and the corollary, plus compatibility of the product with inclusion (proposition 2.28); the last from compatibility of the opposite involution with the product (proposition 2.41). In the opposite direction we also have

$$(fg)^{\mathrm{op}}\,(fg) = g^{\mathrm{op}}f^{\mathrm{op}}fg \subseteq g^{\mathrm{op}}Bg = g^{\mathrm{op}}g \subseteq C \tag{2.166}$$

where the first step follows from compatibility of the opposite with products; the second from the lemma, the fact that f is a function and compatibility of the product with inclusion; the third from the fact that the identity on N is neutral for the product; the last again from the lemma and the fact that g is a function. The inclusions (2.165) and (2.166) together with lemma 2.50 prove that fg is a function. □

Products of functions use a special notation too. When f and g are composable, one can write $g \circ f$ instead of fg — note the inversion of the order; this is called the *composition* of f and g and is read '*g composed with f*'. The origin of this notation is in the fact that we write $y = f(x)$ instead of xfy for functions; if one applies successively first f and then g to x one gets $g(f(x))$ which can be written as $(g \circ f)\,(x)$ keeping the relative position of f and g.
Since the identity relation is a function and the product of relations is associative and has identity functions both left and neutral elements, from proposition 2.26 and proposition 2.52 we immediately have

Corollary 2.53. Functions form a subcategory **Fun** of relations, i.e. the following formulas hold for functions f, g and h whenever the relevant products are defined.

$(fg)\,h = f(gh)$	associativity	(2.167)
$If = f$	left neutral element	(2.168)
$fI = f$	right neutral element	(2.169)

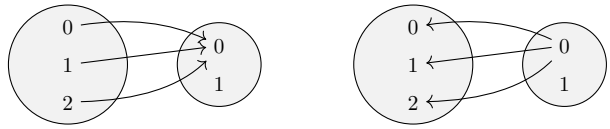
□

□ **Isomorphisms, monomorphisms, epimorphisms.** The relation of the involution with functions turns out to have more interesting consequences than one may expect. To start with, the opposite relation of a function is not, in general a function.

Example 2.54 Lat $A = \{0, 1, 2\}$, $B = \{0, 1\}$ and $f : A \to B$ be the function whose incidence matrix is given on the left below.

$$M(f) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \qquad M(f^{\mathrm{op}}) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \tag{2.170}$$

The matrix of the opposite relation is $M(f^{\mathrm{op}}) = M(f)^{\mathrm{T}}$ and is shown on the right. Inspection of the rows of $M(f^{\mathrm{op}})$ immediately shows that f^{op} is not a function, as it fails to have a single element equal to 1 in every row. A hint of the problem comes from looking at the graphs of f and f^{op}: the fact that the direct images of elements of B in f^{op} fsil to have a single element depends on the fact that the inverse images of elements of B along f do not have this property.



(2.171)

Example 2.55 For an analytic example, consider the function $f : \mathbf{Z} \to \mathbf{Z}$ defined by the formula $f(x) = x^2$. The fact that this is a function follows from the observation that given any integer $x \in \mathbf{Z}$, there is always one and exactly one way to compute its square $x^2 \in \mathbf{Z}$. The opposite relation, however, is defined by

$$yf^{\mathrm{op}}x \Leftrightarrow xfy \Leftrightarrow f(x) = y \Leftrightarrow x^2 = y. \tag{2.172}$$

Thus, the direct image of y under f^{op} is the set of square roots of y in $\mathbf{Z}$. It so happens that the only integer that has a single square root is 0. The integer 1 has two square roots in $\mathbf{Z}$, namely ± 1, and so do infinitely many integers: $4, 9, 16, \ldots$. There are also infinitely many integers that do not have any square root, including all the negatives. Thus, f^{op} is quite far from being a function.

The examples suggest that we take a more detailed look at inverse images of elements in a function.

Definition 2.56. Assume $f : A \to B$ is a function. We say that f is

- *surjective* or an *epimorphism* if the inverse image of every $b \in B$ has at least one element;
- *injective* or a *monomorphism* if the inverse image of every $b \in B$ has at most one element;
- *bijective* or an *isomorphism* if the inverse image of every $b \in B$ has exactly one element.

Note that being bijective is equivalent to being both injective and surjective. Surjectivity means that every element of the codomain is in the image of at least one element of the domain. Injectivity means that elements that are distinct in the domain are transformed into distinct elements of the codomain; it can not happen that different elements of A are sento to the same element of B.

Example 2.57 (Functions between finite sets) Assume A and B are finite and $f : A \to B$ is represented by the matrix M. If we fix $b_j \in B$, we have

$$a_i \in f^*(b_j) \leftrightarrow f(a_i) = b_j \Leftrightarrow M_{ij} = 1 \tag{2.173}$$

so that the cardinality of the inverse image of b_j is the number of elements equal to 1 in the j-th column. Thus f is surjective precisely when every column of M has a least a 1; it is injective when every column contains at most a 1; it is bijective when every column of M has a single 1.

Example 2.58 Let $A = \{0, 1, 2\}$ and $B = \{0, 1\}$. Consider the functions $f : A \to B$ and $g : B \to A$ defined by the matrices

$$M(f) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad M(g) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{2.174}$$

By example 2.57, f is surective but not injective and g is injective but not surjective. On the other hand, if we consider the functions $h, k : A \to A$ defined by the matrices

$$M(h) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad M(k) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.175}$$

we have that h is neither injective nor surjective and k is bijective.

Example 2.59 Consider the functions $\mathbf{N} \to \mathbf{N}$ defined by the following formulas:

$$f(x) = x + 1, \quad g(x) = \begin{cases} 0 & \text{if } x = 0 \\ x - 1 & \text{if } x \neq 0, \end{cases} \quad h(x) = \begin{cases} x + 1 & \text{if } x \text{ is even} \\ x - 1 & \text{if } x \text{ is odd.} \end{cases} \tag{2.176}$$

The function f, known as the *right shift* sends every natural number into its successor; since 0 is not the successor of any natural number, 0 does not belong to the direct image of any element and therefore f is not surjective. On the other hand, if two natural numbers x and y have the same successor $x + 1 = y + 1$, they must be equal; thus every elementof the codomain can have at most one element in the inverse image and f is injective.
The function g, known as the *left shift* sends every element to its predecessor, except for 0 which, not having a predecessor in $\mathbf{N}$, is sent to itself. This implies that g is not injective, because both 0 and 1 are sent to 0, so that 0 has two elements in the inverse image. It is surjective, though, because every $y \in \mathbf{N}$ belongs at least to the direct image of its successor.
The function h sends even numbers to their successors, which are odd, and odd numbers to their predecessors, which are even; since the first odd number is 1, $h(x)$ is always and uniquely defined, so that h is indeed a function. If now $y \in \mathbf{N}$ is even, its inverse image can only contain odd numbers and there is only one which is a successor of y; thus the inverse image of y has a single element. Similarly, if y is odd, its inverse image must consists of odd numbers, and there is only one which is a predecessor of y; in this case too, the inverse image of y has a single element. Therefore, h is biective.

We provide an equivalent algebraic-type description of injectivity and surjectivity.

LEMMA 2.60. Assume $R : A \to B$ is any relation. Then, for every $a \in A$ and every $b \in B$, the direct image of b along R^{op} is the inverse image of b along R and dually, the inverse image of a along R^{op} is the direct image of a along R.

$$(R^{\mathrm{op}})_*(b) = R^*(b), \qquad (R^{\mathrm{op}})^*(a) = R_*(a). \tag{2.177}$$

Proof. We have

$$a \in (R^{\mathrm{op}})_*(b) \Leftrightarrow bR^{\mathrm{op}}a \Leftrightarrow aRb \Leftrightarrow a \in R^*(b) \tag{2.178}$$

where the first equivalence follows from the definition of direct image; the second from that of opposite relation; the last from the definition of inverse image. This proves the first formula in (2.177) by extensionality. Likewise,

$$b \in (R^{\mathrm{op}})^*(a) \Leftrightarrow bR^{\mathrm{op}}a \Leftrightarrow aRb \Leftrightarrow b \in R_*(a) \tag{2.179}$$

which proves the second formula. □

PROPOSITION 2.61. Let $f : A \to B$ be a function. Then

1. f is surjective if and only if $f^{\mathrm{op}} f = B$,
2. f is injective if and only if $A = f f^{\mathrm{op}}$.

Proof. ① We have

$$f \text{ surjective} \Leftrightarrow \forall b \in B(|f^*(b)| \geq 1) \tag{2.180}$$
$$\Leftrightarrow \forall b \in B(|f^{\mathrm{op}}_*(b)| \geq 1) \tag{2.181}$$
$$\Leftrightarrow B \subseteq f^{\mathrm{op}} f \tag{2.182}$$
$$\Leftrightarrow B = f^{\mathrm{op}} f \tag{2.183}$$

where the first equivalence follows from the definition of injective function; the second from lemma 2.60; the third from proposition 2.49; the last from corollary 2.50. ② is similar and is omitted. □

We can now determine when the opposite relation of a function is a function.

COROLLARY 2.62. If $f : A \to B$ is a function, then $f^{\mathrm{op}} : B \rightharpoonup A$ isa function if and only if f is an isomorphism. In this case f^{op} is also an isomorphism and

$$f f^{\mathrm{op}} = A, \qquad f^{\mathrm{op}} f = B. \tag{2.184}$$

Proof. f is an isomorphism if and only if $f f^{\mathrm{op}} = A$ and $f^{\mathrm{op}} f = B$ (proposition 2.61) which in turn is equivalent to the fact that f^{op} is a function (corollary 2.50). The formulas in (2.184) prove that f^{op} is an isomorphism (proposition 2.61). □

Thus, when f is an aisomorphism it is a reversible transformation and f^{op} is the transformation that restores the initial process, in either direction. This raises the question of whether there can be other transformations beside f^{op} with this property. We generalize the situation a bit.

DEFINITION 2.63. Let $f : A \to B$ be a function. A function $g : B \to A$ is:

- a *right inverse* for f if $fg = A$;
- a *left inverse* for f if $gf = B$;
- a *(two sided) inverse* for f if $fg = A$ and $gf = B$.

When f has a left inverse we say that it is *left invertible*; we also say that it is *right invertible* and *invertible* when it has a right or a two sided inverse respectively. Note that when compositional notation is used we have $fg = g \circ f$ so that a right inverse should be called a left inverse and conversely. We first connect invertibility with isomorphisms.

PROPOSITION 2.64. Let $f : A \to B$ be a function.

1. f is an epimorphism if and only if it has left inverse.
2. f is a monomorphism if and only if it has a right inverse, provided $A \neq \varnothing$.
3. f is an isomorphism if and only if it has an inverse.

Proof. ① Necessity. If f is surjective, for every $b \in B$ we can choose an $a \in f^*(b)$ and set $g(b) = a$. By construction $g : B \to A$ is a function and $f(g(b)) = f(a) = b$ so that $gf = B$. Sufficiency. If $gf = B$, then for every $b \in B$ we have $f(g(b)) = b$ so that $a := b(b) \in f^*(b)$, $f^*(b) \neq \varnothing$ and f is surjective. A more interesting proof of sufficiency is the following: observe that $g \subseteq f^{\mathrm{op}}$, because

$$g = gA \subseteq gff^{\mathrm{op}} = Bf^{\mathrm{op}} = f^{\mathrm{op}} \tag{2.185}$$

where the first equality follows from the fact that the identity on A is neutral for the product; the second inclusion follows from the fact that f is a function, from corollary 2.50 and from caompatibility of the product with the order relation; the third equality follows from the assumption that $gf = B$; the last equality from the fact that the identity on B is neitral for the product. Bu then

$$B = gf \subseteq f^{\mathrm{op}}f \subseteq B \tag{2.186}$$

where the first step follows from the assumption that g is a left inverse of f; the second from the fact, just proved, that $g \subseteq f^{\mathrm{op}}$ and from compatibility of the product with the order relation; the last from the fact that f is a function and from corollary 2.50. Thus, $f^{\mathrm{op}}f = B$ and f is an epimorphism (proposition 2.61). ② Necessity. Since $A \neq \varnothing$ we can choose an element $c \in A$. Define $g : B \to A$ as follows: given $b \in B$, either $f^*(b) = \{a\}$ has a single element, in which case we set $g(b) = a$, or $f^*(b) = \varnothing$ in which case we set $g(b) = c$. For every $a \in A$ we have $g(f(a)) = g(b) = a$ because $f^*(f(a)) = \{a\}$, so that $fg = A$. Sufficiency. If $fg = A$ and $a, a' \in f^*(b)$, then $f(a) = b = f(a')$ and therefore $a = g(f(a)) = g(b) = g(f(a')) = a'$, so that every inverse image has a single element. We can provide a different proof of sufficiency as follows: if $fg = A$ then

$$f^{\mathrm{op}} = f^{\mathrm{op}}A = f^{\mathrm{op}}fg \subseteq Bg = g \tag{2.187}$$

so that $f^{\mathrm{op}} \subseteq g$ and therefore

$$A \subseteq ff^{\mathrm{op}} \subseteq fg = A, \tag{2.188}$$

$ff^{\mathrm{op}} = A$ and f is injective. ③ This follows from 1 and 2. □

Observe that the nonemptyness assumption on A has only been used to prove necessity of condition 2 and this does not affect the proof of 3.
We can now address the possible existence of over transformation that reverse the action of a function f.

PROPOSITION 2.65. If a function $f : A \to B$ has a left inverse g and a right inverse h, then $g = h$ is a two sided inverse. In particular, if f is invertible the inverse is unique and must coincide with f^{op}.

Proof. ① It suffices to observe that

$$g = gA = g(fh) = (gf)h = Bh = h \tag{2.189}$$

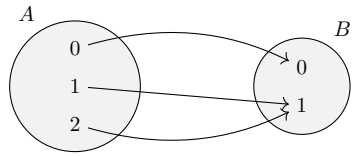
where the first equality follows from the fact that the identity on A is neutral for the product; the second from the fact that h is a right inverse of f; the third from associativity of the product; the fourth from the fact that g is a left inverse of f; the last from the fact that the identity on B is neutral for the product. ② Suppose now that f is invertible. If g and h are two inverses of f, then, in particular, g is a left inverse and h a right inverse and therefore $g = h$ by the first part. ③ Finally, if f is invertible it is an isomorphism (proposition 2.64) and therefore f^{op} is an inverse of f (corollary 2.62); therefore it is the only inverse. □

The only inverse of f, when it exists, is more frequently written f^{-1}. Note, however, that if f is only left invertible then it can have more than one left inverse and likewise, if it is only right invertible it can have more than one right inverse.

Example 2.66 Let $A = \{0, 1, 2\}$, $B = \{0, 1\}$ and $f : A \to B$ the function represented by the matrix

$$M(f) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}. \tag{2.190}$$

Observe that f is surjective because every column of the matrix contains at least a 1. To construct a left inverse $g : B \to A$ we have to define $g(0)$ and $g(1)$. However, sice we want g to be a left inverse of f, we must have $f(g(y)) = y$ for every $y \in B$. This means that $g(y)$ must be in the in the inverse image of y along f.



(2.191)

Since $f^*(0) = \{0\}$ we have no choice other tahn to set $g(0) = 0$. However, $f^*(1) = \{1, 2\}$ and we can set $g(1)$ to be either 1 or 2. In conclusion we have two left inverses of f, represented by the matrices

$$M(g_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad M(g_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.192}$$

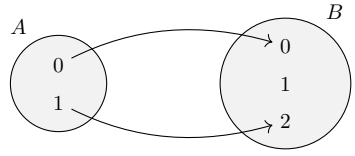
The argument we have just given shows that, in general, the number of left inverses of a surjective function $f : A \to B$ is

$$n = \prod_{b_i \in B} |f^*(b_i)|. \tag{2.193}$$

Example 2.67 Let $A = \{0, 1\}$, $B = \{0, 1, 2\}$ and $f : A \to B$ the function represented by the matrix

$$M(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.194}$$

The injectivity of f is easily verified because every column of the matrix has at most one element equal to 1. To construct a right inverse $g : B \to A$ we must have $g(f(x)) = x$ for ever $x \in A$. This formula has two consequences: if $y \in B$ is of the form $y = f(x)$, i.e. if the inverse image $y \in B$ along f has an element $x \in A$ — which is unique by injectivity — then we must set $g(y) = x$; on the other hand, if $f^*(y)$ is empty, then we have no conditions on $g(y)$ and we can assign to y any $x \in A$; since the assignment is necessary, however, we need at least one element in A and this is the reason for the condition $A \neq \varnothing$.



(2.195)

Thus, the image along g of elements $y \in B$ with nonempty inverse image is uniquely determined. Instead, if y has empty inverse image, we can choose any element of A. Therefore, we have two possible chices for a right inverse:

$$M(g_1) = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad M(g_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}. \tag{2.196}$$

Thus, if $S \subseteq B$ is the subset of elements with empty inverse image, the argument above shows that the number of right inverses is $n = |A|^{|S|}$.

Example 2.68 In proposition 2.156, the claim is not completely invertible: if h is surjective, then f need not be and likewise, if h is injective, g need not be. To see this, let $A = C = \{0\}$ and $B = \{0, 1\}$. Consider the functions $f : A \to B$ and $g : B \to C$ defined by

$$f = \{(0, 0)\}, \qquad g = \{(0, 0), (0, 1)\}. \tag{2.197}$$

Then $h = fg = \{(0, 0)\}$ is bijective, hence both injective and surjective; however neithr f is surjective nor g is injective.

We provide a final characterization of the different types of morphisms.

Definition 2.69. A function $f : A \to B$ is

- *left cancellable* if $fu = fv \Rightarrow u = v$ for all functions u and v with domain B;
- *right cancellable* if $uf = vf \Rightarrow u = v$ for all functions u and v with codomain A;
- *cancellable* if it is both left and right cancellable.

Proposition 2.70. A function $f : A \to B$ is

1. an epimorphism if and only if it is left cancellable;
2. a monomorphism if and only if it is right cancellable;
3. an isomorphism if and only if it is cancellable.

Proof. ① Assume f is an epimorphism. Then f has a left inverse g (proposition 2.64) and if $fu = fv$ we have

$$u = Bu = (gf)\,u = g(fu) = g(fv) = (gf)\,v = Bv = v. \tag{2.198}$$

Conversely, assume f is left cancellable. Consider the functions $u, v : B \to 2 = \{0, 1\}$ defined as follows. $u(y) = 0$ if the inverse image of y along f is empty and $u(y) = 1$ otherwise; instead $v(y) = 1$ for every $y \in B$. Observe that $u(f(x)) = v(f(x)) = 1$ for every $x \in A$ because $f(x)$ has non-empty inverse image. Thus, $fu = fv$ and by cancellability $u = v$. This means that $u(y) = v(y) = 1$ for every y which menas that every y has nonempty inverse image. ② Assume f is a monomorphism. Then f has a right inverse g and if $uf = vf$, then

$$u = uA = u(fg) = (uf)\,g = (vf)\,g = v(fg) = vA = v. \tag{2.199}$$

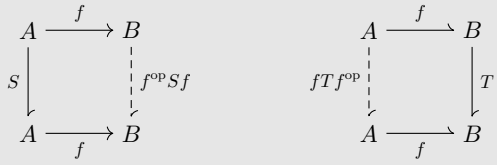
Conversely, observe that every $x \in A$ defines a function $\hat{x} : 1 = \{0\} \to A$ defined by the formula $\hat{x}(0) = x$. Thus, if f is right cancellable and $x_1, x_2 \in f^*(y)$, then

$$f(\hat{x}_1(0)) = f(x_1) = y = f(x_2) = f(\hat{x}_2(0)) \tag{2.200}$$

which means that $\hat{x}_1 f = \hat{x}_2 f$ and by right cancellability $\hat{x}_1 = \hat{x}_2$ so that $x_1 = \hat{x}_1(0) = \hat{x}_2(0) = x_2$ and f is injective. ③ Follows from 1 and 2. □

□ **Direct and inverse images of relations along a function.** We now examine how binary relations can be '*transported*' along functions.

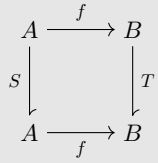
DEFINITION 2.71. Let $f : A \to B$ be a function. The *direct image* of a binary relation S on A along f is the relation $f_*(S) = f^{\mathrm{op}} S f$ on B. In the opposite direction, the *inverse image* along f of a binary relation T on B is the binary relation $f^*(T) := f T f^{\mathrm{op}}$ on A.



(2.201)

Observe that the two diagrams in (2.201) do not commute, in general; the importan property that direct and inverse images have is the following.

PROPOSITION 2.72. (IMAGES ADJOINTNESS) Given a function $f : A \to B$ and binary relations S on A and T on B as in the following diagram,



(2.202)

then the following adjointness relation holds:

$$f_*(S) \subseteq T \Leftrightarrow Sf \subseteq fT \Leftrightarrow S \subseteq f^*(T)\,. \tag{2.203}$$

Proof. We have

$$f_*(S) \subseteq T \Leftrightarrow f^{\mathrm{op}} S f \subseteq T \Leftrightarrow Sf \subseteq fT \Leftrightarrow S \subseteq fTf^{\mathrm{op}} \Leftrightarrow S \subseteq f^*(T)\,. \tag{2.204}$$

The first equivalence follows from the definition of direct image. In the second equivalence, left to right, we multiply on the left by f and use the fact that $1 \subseteq ff^{\mathrm{op}}$ because f is a function; conversely, in the right to left implication we multiply by f^{op} on the left and use the fact that $f^{\mathrm{op}} f \subseteq 1$ because f is a function. In the third equivalence, for the left to right implication we multiply on the right by f^{op} and use $1 \subseteq ff^{\mathrm{op}}$; for the right to left implication we multiply on the right by f and use $f^{\mathrm{op}} f \subseteq 1$. The last implication follows from the definition of inverse image. □

In definition 2.71 we could replace the function f with a binary relation R and define direct and inverse images along a relation. However, the properties of proposition 2.72 would not hold, as the proof depends on the fact that f is a function.
A first consequence of formula (2.203) is that the diagrams in (2.201) commute *"up to inclusion"*: setting $T = f_*(S)$ in the left inclusion and $S = f^*(T)$ in the second, the central inclusion becomes respectively

$$Sf \subseteq f f_*(S), \qquad\qquad f^*(T)\, f \subseteq fT. \tag{2.205}$$

Another consequence of setting $T = f_*(S)$ first and $S = f^*(T)$ in formula (2.203) is that

$$S \subseteq f^* f_*(S), \qquad\qquad f_* f^*(T) \subseteq T. \tag{2.206}$$

Here is another important consequence.

Corollary 2.73. Given $f : A \to B$, relations S_i on A and T_i on B, the following formulas hold

$$f_*\Big(\bigcup S_i\Big) = \bigcup f_*(S_i), \qquad f^*\Big(\bigcap T_i\Big) = \bigcap f^*(T_i)\,. \tag{2.207}$$

Proof. ① Given any binary relation T on B we have

$$\begin{aligned} f_*\Big(\bigcup S_i\Big) \subseteq T &\Leftrightarrow \bigcup S_i \subseteq f^*(T) && (2.208)\\ &\Leftrightarrow \forall i(S_i \subseteq f^*(T)) && (2.209)\\ &\Leftrightarrow \forall i(f_*(S_1) \subseteq T) && (2.210)\\ &\Leftrightarrow \bigcup f_*(S_i) \subseteq T && (2.211) \end{aligned}$$

where the first equivalence follows from the adjointness property; the second from properties of the union; the third again from adjointness; the last from the properties of union. Thus, $f_*(\bigcup S_i)$ and $\bigcup f_*(S_i)$ are contained in the same subsets of B^2 and therefore are equal. ② In the same way one proves the second equality by showing that both relations contain the same binary relations on A. □

2.4 Equivalence relations

□ Definition and examples.

Definition 2.74. An *equivalence relation* on a set A is a binary relation $E : A \rightharpoonup A$ which is reflexive, symmetric and transitive, i.e. which satisfies the following formulas.

$$\begin{aligned} &\forall x.xEx && \text{Reflexivity} && (2.212)\\ &\forall xy(xEy \to yEx) && \text{Symmetry} && (2.213)\\ &\forall xyz((xEy) \wedge (yEz) \to (xEz)) && \text{Transitivity} && (2.214) \end{aligned}$$

The definition can be reformulated in a more algebraic form which is often more useful for calculations.

Proposition 2.75. A binary relation $E : A \rightharpoonup A$ on a set A is an equivalence relation if and only if it contains the identity relation and satisfies the following formulas.

$$\begin{aligned} &E^{\mathrm{op}} = E && \text{Involution} && (2.215)\\ &E^2 = E && \text{Idempotency} && (2.216) \end{aligned}$$

Proof. Formulas (2.212)–(2.214) in the definition of equivalence relation amount respectively to $A \subseteq E$ for reflexivity, $E \subseteq E^{\mathrm{op}}$ for symmetry and $E^2 \subseteq E$ for transitivity. Thus, the conditions are sufficient. For necessity, assume E is an equivalence. Then

$$E \subseteq E^{\mathrm{op}} \Rightarrow E^{\mathrm{op}} \subseteq E^{\mathrm{op}\,\mathrm{op}} \Rightarrow E^{\mathrm{op}} \subseteq E \tag{2.217}$$

where the first implication follows from compatibility of opposites with inclusions; the second from the fact that opposites are involutory. From the double inclusion we conclude that $E^{\mathrm{op}} = E$. Also,

$$E = EI \subseteq EE \subseteq E \tag{2.218}$$

where the first equality follows from the fact that the identity relation is neutral for the product; the second inclusion from the fact that $I \subseteq E$ and from compatibility of products with inclusions; the last from transitivity. Again, from the double inclusion we have $E^2 = E$. □

Observe that the idempotency condition $E^2 = E$ implies, by induction, that$E^n = E$ for every $n > 0$. Equivalence relations are ubiquitous in mathematics. We give here a few, elementary examples.

Example 2.76 (Equality) The most important example of equivalence relation on any set A is equality. The relation is reflexive because $a = a$ for every $a \in A$. It is symmetric because if $a = b$ then $b = a$. It is transitive because if $a = b$ and $b = c$ then $a = c$.

Example 2.77 (Extremal equivalences) On every set A the identity relation I and the total relation $\top$ are equivalence relations.

- For the identity relation this follows from proposition 2.75 and from the observation that $I^{\mathrm{op}} = I$ by compatibility of the opposite with the identity and the fact that $I^2 = I$ because the identity is neutral for the product.
- As to the total relation, since every binary relation on A is conatined in $\top$, we have in particular that $I, \top, \top^2 \subseteq \top$, which are precisely the conditions for an equivalence relation.

Observe that if $E : A \rightharpoonup A$ is any equivalence relation on A, then $I \subseteq E \subseteq \top$, where the first inclusion follows from reflexivity of E and the second by the definition of $\top$. Thus, I and $\top$ are, respectively, the smallest and the largest equivalence relations on A.

Example 2.78 Let $A = \{0, 1, 2, 3\}$. The binary relation $E : A \rightharpoonup A$ whose incidence matrix is

$$M = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \tag{2.219}$$

is an equivalence relation by proposition 2.75. In fact, if $I = M(I)$ is the matrix of the identity relation on A and recalling that $M(E^{\mathrm{op}}) = M^{\mathrm{T}}$ and $M(E^2) = M^2$, the conditions of the proposition translate into the conditions

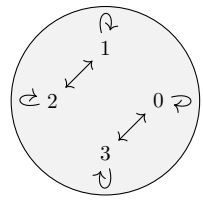
$$I \leq M, \qquad M^{\mathrm{T}} = M, \qquad M^2 = M \tag{2.220}$$

on matrices. And these conditions are satisfied because

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \leq \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \tag{2.221}$$

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \tag{2.222}$$

The graph of E is represented below. Observe that since both the domain and codomain are A, we have represented it only once; this compact representation is often more convenient for equivalence relations, or even any relation $A \rightharpoonup A$ when A is finite.



(2.223)

Note that the representation with a single set has some consequences on the way arrows are represented on the graph. Since E is reflexive, we have xEx for every $x \in A$; this is represented by drawing a loop on every element x. Symmetry of E implies that if xEy then yEx; this means that whenever we have an arrow from x to y there is also an arrow in the opposite direction , from y to x; rather than drawing two distinct arrows in opposite direction, we rather draw a single, two headed arrow between x and y. Transitivity is not easily verified on the graph: the condition $xEy \wedge yEz \to xEz$ means that if we have arrows $x \to y \to z$ in the graph, there must also be an arrow $x \to z$. Repeated application of this condition implies actually that every time there is a path

$$x_0 \to x_1 \to \cdots \to x_n \tag{2.224}$$

of length n in the graph, there must be an arrow $x_0 \to x_n$ between the endpoints of the path.

Example 2.79 (Semantical equivalence) Let L be a propositional language. Recall that two formulas $\varphi, \psi \in L$ are semantically equivalent if they have the same truth table, i.e. if $v(\varphi) = v(\psi)$ for every valuation v. This is an equivalence relation.

1. Reflexivity. Given $\varphi \in L$, we have $v(\varphi) = v(\varphi)$ for every valuation v and therefore $\varphi \equiv \varphi$.
2. Symmetry. Assume $\varphi \equiv \psi$. Then $v(\varphi) = v(\psi)$ for every valuation v. By symmetry of the equality relation on Ω, we have $v(\psi) = v(\varphi)$ for every v and therefore $\psi \equiv \varphi$.
3. Transitivity. Assume $\varphi \equiv \chi$ and $\chi \equiv \psi$. Given any valuation v we have $v(\varphi) = v(\chi)$ by the first assumption and $v(\chi) = v(\psi)$ by the second. Since equality in Ω is an equivalence relation and hence transitive, we have $v(\varphi) = v(\psi)$. Therefore $\varphi \equiv \psi$.

Now recall that we can reformulate semantical equivalence internally in the form $\vDash \varphi \leftrightarrow \psi$. This allows the following generalization: if $T \subseteq L$ is any theory, we can define *equivalence modulo* T by the formula

$$\varphi \underset{T}{\equiv} \psi \Leftrightarrow T \vDash \varphi \leftrightarrow \psi. \tag{2.225}$$

The sense of the definition is that assuming T, the formulas φ and ψ behave in the same way. More precisely, $v(\varphi) = v(\psi)$ for all valuations v such that $v \vDash T$. Equivalence modulo T is an equivalence relation: copy the proof given above, restricting only to valuations satisfying T. In fact, we can further generalize as follows: consider any set of valuations $V \subseteq \Omega^X$; then *equivalence modulo* V, defined by the formula

$$\varphi \underset{V}{\equiv} \psi \Leftrightarrow \forall v \in V(v(\varphi) = v(\psi)) \tag{2.226}$$

is an equivalence relation on L, exactly with the same proof given above. The case of equivalence modulo T can be recovered taking $V = \left\{v \in \Omega^X : v \vDash T\right\}$ and ordinary semantical equivalence taking $V = \Omega^X$. Observe that if we take $V = \varnothing$, then equivalence mosulo V is the total relation on L.

Example 2.80 (Congruence modulo n) Every natural number $n > 0$ defines an equivalence relation $\equiv_n$: $\mathbf{Z} \rightharpoonup \mathbf{Z}$, called *congruence modulo* n. The relation is defined by the formula

$$x \equiv_n y \Leftrightarrow n|(y - x) \Leftrightarrow \exists q \in \mathbf{N}(x + nq = y)\,. \tag{2.227}$$

The first condition is the actual definition, the second is more convenient for calculations. The relation is reflexive because

$$x + n0 = x \Rightarrow x \equiv_n x \tag{2.228}$$

It is symmetric because

$$x \equiv_n y \Rightarrow \exists q(x + nq = y) \Rightarrow \exists q(y + n(-q) = x) \Rightarrow \exists p(y + np = x) \Rightarrow y \equiv_n x \tag{2.229}$$

where the first implication follows from the definition of congruence; the second from arithmetical properties of $\mathbf{Z}$; the third by taking $p = -q$; the last again from the definition of congruence. Finally, congruence is transitive because

$$\begin{aligned}(x \equiv_n y) \wedge (y \equiv_n z) &\Rightarrow \exists p(x + np = y) \wedge \exists q(y + nq = z) && (2.230)\\ &\Rightarrow \exists pq(x + n(p + q) = z) && (2.231)\\ &\Rightarrow \exists r(x + nr = z) && (2.232)\\ &\Rightarrow x \equiv_n z\end{aligned}$$

where the first implication follows from the definition of congruence;; the second from arithmetical properties of $\mathbf{Z}$; the third by setting $r = p + q$; the last by definition of congruence modulo n.

Example 2.81 If $E : A \rightharpoonup A$ is an equivalence relation and B is any set, then E induces an equivalence relation E^B on the set $A^B = \{f : B \to A\}$ of all functions from B to A, defined by the formula

$$(f, g) \in E^B \Leftrightarrow \forall b \in B.(f(b), g(b)) \in E \tag{2.233}$$

- E^B is reflexive because given $f \in A^B$ we have

$$\forall a \in A(aEa) \Rightarrow \forall b \in B(f(b)\, Ef(b)) \Rightarrow fE^B f \tag{2.234}$$

  where the first implication follows from the fact that if the claim holds for all $a \in A$, it holds in particular for $f(b) \in A$; the second follows from the definition of E^B.
- E^B is symmetric because

$$fE^B g \Rightarrow \forall b(f(b)\, Eg(b)) \Rightarrow \forall b(g(b)\, Ef(b)) \Rightarrow gE^B f \tag{2.235}$$

  where the first implication follows from the definition of E^B; the second from symmetry of E; the last again from the definition of E^B.
- E^B is transitive because

$$(fE^B g) \wedge (gE^B h) \Rightarrow \forall b(f(b)\, Eg(b)) \wedge \forall b(g(b)\, Eh(b)) \tag{2.236}$$

$$\Rightarrow \forall b(f(b)\, Eg(b) \wedge g(b)\, Eh(b)) \Rightarrow \forall b(f(b)\, Eh(b)) \tag{2.237}$$

$$\Rightarrow fE^B h \tag{2.238}$$

  where the first implication follows from the definition of E^B; the second from properties of the (infinitary) conjunction; the third from transitivity of E; the last from the definition of E^B.

□ **The generated equivalence.** We now describe a procedure to produce a countless number of equivalence relations.

PROPOSITION 2.82. Given any relation $R : A \rightharpoonup A$ there is a smallest equivalence relation R^e on A caontaining R, the *equivalence closure* of R.

Proof. Observe first that the intersection of any set $S = \{E_i : i \in I\}$ of equivalence relations on A is an equivalence relations on A. For let $E = \bigcap_i E_i$.

- E is reflexive. Since every equivalence relation is reflexive, we have $A \subseteq E_i$ for every index i. Therefore $A \subseteq \bigcap_i E_i = E$ and E is reflevive.
- E is symmetric. This is proved observing that

$$E^{\mathrm{op}} = \left(\bigcap_i E_i\right)^{\mathrm{op}} = \bigcap_i E_i^{\mathrm{op}} = \bigcap_i E_i = E \tag{2.239}$$

  where the first equality follows from the definition of E; the second from compatibility of opposites with intersections (proposition 2.40); the third from the fact that every E_i is symmetric; the last from the definition of E.
- E is transitive. For this observe that

$$E^2 = \left(\bigcap_i E_i\right)\left(\bigcap_j E_j\right) \subseteq \bigcap_{ij} E_i E_j \subseteq \bigcap_i E_i E_i = \bigcap_i E_i = E \tag{2.240}$$

  where the first equality follows from the definition of E; the second from formula (2.111); the third from the fact that the intersection on the right is taken on a subset of the set of indices of the intersection on the left, $D = \{(i, j) : i = j\} \subseteq I \times J$, and thus contains the intersection on the left; the fourth from the fact that $E_i^2 = E$ because every E_i is an equivalence relation; the last from the definition of E.

Now let S be set of all equivalence relations on A that contain R. Observe that S is not empty, because the total relation $\top$ on A is an equivalence relation (example 2.77). Let R^e be the intersection of all elements of S. By the first part of the proof R^e is an equivalence relation. Also, all elements of S contain R by definition and therefore R is also contained in their intersection R^e. Finally, R^e is the smallest equivalence

on A containing R: for if E is an equivalence relation on A and $R \subseteq E$ then $E \in S$ and therefore $R^e \subseteq E$. □

R^e is also called the equivalence relation *generated* by R and, when no confusion could arise by considering other types of closure, it is also denoted by $\langle R \rangle$.
As it is, the proposition does not provide particular insight in the nature of the equivalence closure, nor on its size. It could be, after all, that all equivalence closures are equal to the total relation on A. A more analytic approach is to construct R^e in stages, adding successively the properties tha equivalence relations should have. This is done using the abstract concept of closure with respect to a property.

Definition 2.83. Let P be a property that a relation $R : A \rightharpoonup A$ may or may not have; we write $R \in P$ when R has — or satisfies — P. Given $R : A \rightharpoonup A$, the *P-closure* of R, if it exists, is the smallest relation $R^P : A \rightharpoonup A$ satisfying P and containing R. Explicitly,

1. $R \subseteq R^P \in P$,
2. $R \subseteq S \in P \Rightarrow R^P \subseteq S$.

Observe that the P-closure of a relation does not necessarily exist for all properties P.

Example 2.84 A relation $R : A \rightharpoonup A$ is *irreflexive* if $(a, a) \notin R$ for every $a \in A$; let P be the property of being irreflexive. Now suppose A is non-empty and let $R = \{(a, a)\}$ where a is an element of A. R does not satisfy P because $(a, a) \in R$; however, every relation $S \supseteq R$ must have $(a, a) \in S$ and therefore can not satisfy P. Thus, R has no P-closure.

However, when a P-closure exists, it is unique.

Proposition 2.85. For every property P, the P-closure of a relation $R : A \rightharpoonup A$ is unique if it exists.

Proof. Suppose S and T are two P-closures of R. Then $R \subseteq S \in P$ and since T is a P-closure of R, $T \subseteq S$. Dually, $R \subseteq T \in P$ and since S is a P-closure we also have $S \subseteq T$. From the two inclusions it follows that $S = T$. □

Our construction of the equivalence closure consists in forming the reflexive, symmetric and transitive closures in stages and in examining their stability properties.

Lemma 2.86. Every relation $R : A \rightharpoonup A$ has a reflexive closure

$$R^r = R \cup A. \tag{2.241}$$

Moreover R^r is compatible with symmetry and transitivity, in the sense that if R is symmetric so is R^r; and if R is transitive, so is R^r.

Proof. ① Since $A \subseteq R \cup A = R^r$, R^r is reflexive and clearly contains R. If now S is reflexive and contains R, we have both $A \subseteq S$ and $R \subseteq S$, so that $R^e = R \cup A \subseteq S$. This proves that $R \cup A$ is the reflexive closure of R. ② Suppose R is symmetric. We have

$$(R^e)^{\mathrm{op}} = (R \cup A)^{\mathrm{op}} = R^{\mathrm{op}} \cup A^{\mathrm{op}} = R \cup A = R^e \tag{2.242}$$

where the first equality follows from the definition of R^e; the second from compatibility of the opposite with union (proposition 2.40); the third from the assumption that R is symmetric and the fact that the identity relation is symmetric; the last by the definition of R^e. ③ Assume R is transitive. Then

$$(R^r)^2 = (R \cup A)(R \cup A) = RR \cup RA \cup AR \cup AA \subseteq R \cup R \cup R \cup A = R \cup A = R^r \tag{2.243}$$

where the first equality follows from the definition of R^r; the second from compatibility of the product with union (proposition 2.29); the third inclusion from the assumption that R is transitive, i.e. $R^2 \subseteq R$ and the fact that the identity is neutral for the product; the fourth from idempotency of union; the last form the definition of R^r. This proves that R^r is transitive. □

LEMMA 2.87. Every relation $R : A \rightharpoonup A$ has a symmetric closure

$$R^s = R \cup R^{\mathrm{op}}. \tag{2.244}$$

Moreover R^s is compatible with reflexivity, in the sense that if R is reflexive so is R^s.

Proof. ① By definition $R \subseteq R^s$; moreover, R^s is symmetric, because

$$(R^s)^{\mathrm{op}} = (R \cup R^{\mathrm{op}})^{\mathrm{op}} = R^{\mathrm{op}} \cup R^{\mathrm{op}\,\mathrm{op}} = R^{\mathrm{op}} \cup R = R^s \tag{2.245}$$

where the first equality follows from the definition of R^s; the second from compatibility of the opposite with unions (proposition 2.40); the third from the involutory property of the opposite (proposition 2.38); the last from commutativity of the union and the definition of R^s. If now S is symmetric and contains R we have

$$R^s = R \cup R^{\mathrm{op}} \subseteq S \cup S^{\mathrm{op}} = S \cup S = S \tag{2.246}$$

where the first equality follows from the definition of R^s; the second inclusion from the assumption that $R \subseteq S$ and from compatibility of the opposite with the order (proposition 2.39), so that $R^{\mathrm{op}} \subseteq S^{\mathrm{op}}$; the third equality from the fact that S is symmetric; the last from idempotency of union. Thus, R^s is the symmetric closure of R. ② Assume R is reflexive. Then

$$I \subseteq R \subseteq R \cup R^{\mathrm{op}} = R^s \tag{2.247}$$

where the first inclusion follows from the assumption that R is reflexive; the second from the definition of union; the last from the definition of R^s. Thus, R^s is symmetric if R is. □

The symmetric closure, however, is not compatible with transitivity: even if R is transitive, R^s need not be.

Example 2.88 On the set $A = \{0, 1\}$ consider the relation R represented by the matrix

$$M(R) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{2.248}$$

Observe that R is trivially transitive because there are no composable pairs of arrows. This can also be checked directly:

$$M(R^2) = M(R)\,M(R) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.249}$$

so that $R^2 = \bot \subseteq R$ and R is transitive. The symmetric closure of R is $R^s = R \cup R^{\mathrm{op}}$ and has matrix

$$M(R^s) = M(R \cup R^{\mathrm{op}}) = M(R) \vee M(R^{\mathrm{op}}) = M(R) \vee M(R)^{\mathrm{T}} \tag{2.250}$$

$$= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \vee \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{2.251}$$

However,

$$M(S^2) = M(S)\,M(S) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = M(A) \tag{2.252}$$

so that $S^2 = A \not\subseteq S$ so that S is not transitive. One can also argue directly as follows: $R = \{(0, 1)\}$ and therefore $S = \{(0, 1), (1, 0)\}$; since $0S1$ and $1S0$ we have $0S^0$, although $(0, 0) \notin S$, so that S is not transitive.

Finally, we address transitivity.

LEMMA 2.89. Every relation $R : A \rightharpoonup A$ has a transitive closure R^t defined by the formula

$$R^t = \bigcup_{n=1}^{\infty} R^n. \tag{2.253}$$

Moreover R^t is compatible with reflexivity and symmetry, in the sense that if R is reflexive so is R^t; and if R is symmetric, so is R^t.

Proof. ① Clearly $R = R^1 \subseteq R^t$. Also, R^t is transitive because

$$(R^t)^2 = \left(\bigcup_i R^i\right)\left(\bigcup_j R^j\right) = \bigcup_{ij} R^i R^j = \bigcup_{ij} R^{i+j} = \bigcup_i R^i = R^t \tag{2.254}$$

where the first equality follows from the definition of R^t; the second from formula (2.112); the third from the definition of powers of a relation; the fourth from idempotence of the union; the last from the definition of R^t. If now S is transitive, then

$$R \subseteq S \Rightarrow \forall n > 0 (R^n \subseteq S^n) \Rightarrow \forall n > 0 (R^n \subseteq S) \Rightarrow \bigcup_{n=1}^{\infty} R^n \subseteq S \Rightarrow R^t \subseteq S \tag{2.255}$$

where the first implication follows from compatibility of the product with inclusions (proposition 2.39); the second from the fact that $S^n \subseteq S$ by transitivity of S; the third from the definition of union; the last from the definition of R^t. This proves that R^t is the transitive closure of R. ② Assume R is reflexive, i.e. that $A \subseteq R$. Then R^t is reflexive because

$$A \subseteq R \subseteq \textstyle\bigcup_{n=1}^{\infty} R^n = R^t \tag{2.256}$$

where the first inclusion follows by the assumption; the second by the definition of union; the last from the definition of R^t. ③ Assume R is symmetric, i.e. $R^{\mathrm{op}} = R$. Then R^t is symmetric because

$$(R^t)^{\mathrm{op}} = \left(\bigcup_{n=1}^{\infty} R^n\right)^{\mathrm{op}} = \bigcup_{n=1}^{\infty} (R^n)^{\mathrm{op}} = \bigcup_{n=1}^{\infty} (R^{\mathrm{op}})^n = \bigcup_{n=1}^{\infty} R^n = R^t \tag{2.257}$$

where the first equality follows from the definition of R^t; the second from the compatibility of the opposite with unions (proposition 2.40); the third from compatibility of the opposite with products (proposition 2.41); the fourth from the assumption that R is symmetric; the last from the definition of R^t. □

We can now provide a more precise description of the equivalence closure of a relation.

PROPOSITION 2.90. Let $R : A \rightharpoonup A$ be a relation. Then the equivalence closure of R is the relation

$$R^e = R^{rst} \tag{2.258}$$

obtained by taking in succession the reflexive, symmetric and transitive closure. In fact R^e can be obtained from R applying the three closures in any order, as long as the transitive closure follow the symmetric one.

Proof. Since the reflexive closure R^r of R is a reflexive relation containing R, the symmetric closure R^{rs} of this relation is both reflexive and symmetric (proposition 2.87) and contains R^r and therefore R. The transtivie closure R^{rst} of R^{rs} is reflexive, symmetric and transitive (proposition 2.89) and contains R^{rs} and therefore also R. Thus, R^{rst} is an equivalence relation containing R. We claim that it is the equivalence closure. For suppose E is an equivalence relation containing R. Since E is reflexive, it must contain the reflexive closure R^r of R. However, E is also symmetric and since it contains R^r it must also contain its symmetric closure R^{rs}. Finally, since E is transitive and contains R^{rs}, it must contain its transitive closure R^{rst}. This proves that R^e is the equivalence closure of R.
The argument does not depend on the order in which the closures are applied, as long as the transitive closure follows the symmetric one; for if the transitive closure is taken before, there is no guarantee that the symmetric closure will preserve transitivity (example 2.89). □

We can use the proposition to provide an interpretation of R^e; this interpretation is best understood first in the finite case.

Example 2.91 Assume A is a finite set and $R : A \rightharpoonup A$. We describe the equivalence closure $E = R^e$ using the notion of finite paths induced by R on A.

- If $a \in A$, aEa because E is reflexive; we represent this with a single element a and call this a path of length 0 from a to a. Thus all endpoints of a path of length 0 are in E.
- If $a_1, a_2 \in A$, we say that there is a path of length 1 from a_1 to a_2 if either a_1Ra_2 or a_2Ra_1. We represent a path of length 1 in the form

$$a_1 \text{ —— } a_2 \tag{2.259}$$

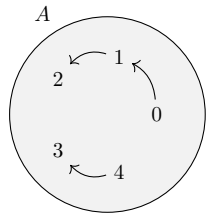
  Since $R \subseteq E$ and E is symmetric, if there is a path of length 1 from a_1 to a_2, then a_1Ea_2. We can think of such a path as an arrow of R with the head removed, because it can be trveled in both directions.
- If $a_1, a_n \in A$, we say that there is a path of length n from a_1 to a_n if there are n paths of length 1

$$a_1 \text{ —— } a_2 \text{ —— } a_3 \cdots\cdots \qquad \cdots\cdots a_{n-1} \text{ —— } a_n \tag{2.260}$$

  starting at a_1 and ending at a_n. If there is such a path, then a_1Ea_n, because every path of length 1 is in E and E is transitive.

The three constructions above correspond to the three steps in the construction of E using the reflexive, symmetric and transitive closures. Thus, given $a, b \in A$, we have aEb if and only if there is a finite path from a to b. We also say that two elements $a, b \in A$ belong to the same *component* if there is a finite path from a to b; note that this is equivalent to say that there is a path from b to a. Thus, aEb exactly when they belong to the same component.

Example 2.92 On $A = \{0, 1, 2, 3, 4\}$ consider the relation R whose graph is given below.



(2.261)

The matrices of R and R^e are

$$M(R) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix} \qquad M(R^e) = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \tag{2.262}$$

The matrix of R^e has been constructed by relating element that belong to the same component. For example, the third row of $M(R^e)$ records the elements in the direct image of 2; these alements are 0, to which 2 is connected by a path of lenght two; 1 with a connection via a path of length one and 2 itself with a path of length zero.

An alernative construction of the equivalence closure in the finite case is provided by the following

LEMMA 2.93. If $R : A \rightharpoonup A$ is a binary relation on A and if there exists $n \in \mathbf{N}$ such that the reflexive closure S of R stabilizes at some power n, in the sense that $S^n = S^{n+1}$, then S^n is the reflexive and transitive closure of R.

Proof. Observe that, since $S = A \cup R$, for every $m \in \mathbf{N}$ we have

$$S^m = S^m A \subseteq S^m(A \cup R) = S^m S = S^{m+1}. \tag{2.263}$$

Now assume $S^n = S^{n+1}$ for some n. We claim that $S^{n+m} = S^n$ for every $m \in \mathbf{N}$. This is proved by induction on m. The case $m = 0$ is obvious, because $S^{n+0} = S^n$. Assuming that the statement is true for some m, we have

$$S^{n+m+1} = S^{n+1}S^m = S^n S^m = S^{n+m} = S^n. \tag{2.264}$$

Now observe that if T is the reflexive and transitive closure of R,

$$T = R^{rt} = S^t = \bigcup_{m=1}^{\infty} S^m \subseteq \bigcup_{m=1}^{\infty} S^n = S^n \subseteq T \tag{2.265}$$

so that $T = S^n$. □

Since the reflexive and symmetric closure commute, we can apply the lemma to the reflexive and symmetric closure of R.

Example 2.94 For the relation R of example 2.92, the symmetric and reflexive closure is $S = R \cup R^{\mathrm{op}} \cup I$ and therefore its matrix is

$$M(S) = M(R) \vee M(R)^{\mathrm{T}} \vee M(I) \tag{2.266}$$

$$= \begin{pmatrix} 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&0&0\\ 0&0&0&0&0\\ 0&0&0&1&0 \end{pmatrix} \vee \begin{pmatrix} 0&0&0&0&0\\ 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&0&0&1\\ 0&0&0&0&0 \end{pmatrix} \vee \begin{pmatrix} 1&0&0&0&0\\ 0&1&0&0&0\\ 0&0&1&0&0\\ 0&0&0&1&0\\ 0&0&0&0&1 \end{pmatrix} = \begin{pmatrix} 1&1&0&0&0\\ 1&1&1&0&0\\ 0&1&1&0&0\\ 0&0&0&1&1\\ 0&0&0&1&1 \end{pmatrix}. \tag{2.267}$$

The successive powers of S are

$$S^2 = \begin{pmatrix} 1&1&1&0&0\\ 1&1&1&0&0\\ 1&1&1&0&0\\ 0&0&0&1&1\\ 0&0&0&1&1 \end{pmatrix}, \qquad S^3 = \begin{pmatrix} 1&1&1&0&0\\ 1&1&1&0&0\\ 1&1&1&0&0\\ 0&0&0&1&1\\ 0&0&0&1&1 \end{pmatrix}. \tag{2.268}$$

and since the powers stabilize at S^2, we have $R^e = S^2$.

□ Quotient sets.

DEFINITION 2.95. Assume $E : A \rightharpoonup A$ is an equivalence relation and $x \in A$. The *equivalence class* of x is the set $[x] \subseteq A$ of elements of A in relation with x in E.

$$[x] := \{y \in A : xEy\} = \{y \in A : yEx\} \tag{2.269}$$

The set of its equivalence classes is called the *quotient set* of A modulo E and is denoted by A/E.

$$A/E = \{[x] : x \in A\}. \tag{2.270}$$

Observe that the two variants of the definition in (2.269) are equivalent because $xEy \Leftrightarrow yEx$ by symmetry. Note also that the equivalence class of x is the direct image of x along E, which coincides with the inverse image of x along E. We collect the basic properties of equivalence classes.

PROPOSITION 2.96. Assume $E : A \rightharpoonup A$ is an equivalence relation and $x, y \in A$. The following conditions are equivalent.

1. $[x] = [y]$
2. $[x] \cap [y] \neq \varnothing$
3. xEy

Proof. We prove the equivalence by showing that $1 \Rightarrow 2 \Rightarrow 3$.

1⇨2. Since xEx by reflexivity, $x \in [x]$ by definition of equivalence class and therefore $[x] \neq \varnothing$. thus

$$[x] \cap [y] = [x] \cap [x] = [x] \neq \varnothing \tag{2.271}$$

2⇨3. Assume $z \in [x] \cap [y]$. By definition of equivalence class, we have xEz and zEy and by transitivity xEy.

3⇨1. Assume xEy. If $z \in [x]$ then zEx by definition of equivalence class; from zEx and xEy it follows zEy by transitivity and therefore $z \in [y]$. Thus $[x] \subseteq [y]$. From xEy by symmetry it follows that yEx and as above $[y] \subseteq [x]$. Thus, $[x] = [y]$. □

Example 2.97 (QUOTIENTS OF EXTREMAL EQUIVALENCES) We have seen in example 2.77 that on every set A the identity relation I is the smalles equivalence. For this relation and for every $x \in A$ we have

$$[x] = \{y \in A : y = x\} = \{x\} \tag{2.272}$$

so that eaxh equivalence class consists of a single element. Therefore, the function $f : A \to A/I$ defined by the formula $f(x) = \{x\}$ provides an isomorphism $A \simeq A/I$ and we can identify A/I with A.
At the opposite side, we also know that the total relation $\top$ is an equivalence relation. By definition of total relation,

$$[x] = \{y \in A : y \top x\} = A. \tag{2.273}$$

Thus, there is a single equivalence class consisting of the whole set A and $A/\top = \{A\}$ is the set with a single element, the set A itself.

Example 2.98 On the set $A = \{0, 1, 2, 3\}$ consider the binary relation E represented by the matrix

$$M = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \tag{2.274}$$

We proved in example 2.78 that E is an equivalence relation. To determine the equivalence classes recall that $a_i E a_j \Leftrightarrow m_{ij} = 1$, so that

$$[a_i] = \{a_j \in A : m_{ij} = 1\}. \tag{2.275}$$

In this case, we have

$$[0] = \{0, 3\} = [3], \qquad [1] = \{1, 2\} = [2] \tag{2.276}$$

so that $A/E = \{[0], [1]\}$ is the quotient set of A modulo E.

Example 2.99 From example 2.80 we know that for $n > 0$ in $\mathbf{N}$, congruence modulo n

$$x \underset{n}{\equiv} y \Leftrightarrow n|(y-x) \Leftrightarrow \exists q \in \mathbf{N}(x + nq = y)\,. \tag{2.277}$$

is an equivalence relation on $\mathbf{Z}$. The quotient set $\mathbf{Z}/\equiv_n$ is usually denoted by $\mathbf{Z}/(n)$ and is called the set of *residue classes* modulo n. If $x \in \mathbf{Z}$, its equivalence class is the set

$$[x] = \{y \in \mathbf{Z} : \exists q(y = x + nq)\} \tag{2.278}$$

of integers obtained from x by adding a multiple of n; in particular, $[0]$ is the set of integer multiples of n. To understand the structure of $\mathbf{Z}/(n)$, recall that the division theorem in $\mathbf{Z}$ states that for every $x \in \mathbf{Z}$ there is a unique pair (q, r) such that

$$x = nq + r, \qquad 0 \leq r < n; \tag{2.279}$$

q is the quotient and r the remainder of the division of x by n. The first formula in (2.279) and proposition 2.96 show that $[x] = [r]$, so that every equivalence class is the class of a remainder. Moreover, different remainders have different equivalence classes because

$$0 \leq r < s < n \Rightarrow 0 < s - r < n \Rightarrow n \nmid (s-r) \Rightarrow r \not\equiv_n s \Rightarrow [r] \neq [s]\,. \tag{2.280}$$

Thus, $\mathbf{Z}/(n)$ consists exactly of the equivalence classes of the remainders:

$$\mathbf{Z}/(n) = \{[0], [1], [2], \ldots, [n-1]\}\,. \tag{2.281}$$

As a special case, $\mathbf{Z}/(0) = \{[0], [1]\}$ where $[0]$ is cle class of even numbers and $[1]$ the class of odd numbers.

Definition 2.100. For every equivalence relation E on a set A, the function

$$\begin{aligned} A &\xrightarrow{\;p\;} A/E \\ x &\longmapsto [x] \end{aligned} \tag{2.282}$$

is called the *projection* on the quotient.

The projection on the quotient has interesting properties.

Proposition 2.101. Let E be an equivalence relation on A and let $p : A \to A/E$ be the projection on the quotient. Then

$$pp^{\mathrm{op}} = E, \qquad p^{\mathrm{op}}p = 1. \tag{2.283}$$

In particular: p is always surjective; and p is injective if and only if E is the euqality relation on A.

Proof. ① Observe first that, by definition of projection, $xpz \Leftrightarrow p(x) = z \Leftrightarrow [x] = z$. Therefore, $zp^{\mathrm{op}}y \Leftrightarrow ypz \Leftrightarrow z = [y]$ abd thus $xpp^{\mathrm{op}}y \Leftrightarrow \exists z(xpz \wedge zp^{\mathrm{op}}y) \Leftrightarrow \exists z([x] = z \wedge z = [y]) \Leftrightarrow [x] = [y]$. Putting everything together, we have

$$xpp^{\mathrm{op}}y \Leftrightarrow [x] = [y] \Leftrightarrow xEy. \tag{2.284}$$

② Likewise

$$upp^{\mathrm{op}}v \Leftrightarrow \exists x(up^{\mathrm{op}}x \wedge xpv) \Leftrightarrow \exists x(u = [x] \wedge [x] = v) \Leftrightarrow u = v \tag{2.285}$$

and therefore $p^{\mathrm{op}}p = 1$. ③ Since $p^{\mathrm{op}}p = 1$, p is surjective (proposition 2.61). And similarly, p is injective when $pp^{\mathrm{op}} = 1$, i.e. if $E = 1$. □

Example 2.102 Consider again the equivalence relation E of example 2.98. By definition of projection $p : A \to A/E$ we have

$$p(0) = [0], \quad p(1) = [1], \quad p(2) = [2] = [1], \quad p(3) = [3] = [0] \tag{2.286}$$

so that the matrix representing p with the chosen ordering of elements in A and A/E is

$$M(p) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{2.287}$$

For more complex cases on a finite set one can also argue as follows. We will prove below that $E = pp^{\mathrm{op}}$. To construct p from E, write C_i for the i-th column of M and R_j for the j-th row. Assuming E has order n, we have

$$M = \bigvee_{i,j=1}^{n} C_i R_j = \bigvee_{i,j \in I} C_i R_j = [C_I]\,[R_I] = [C_I]\,[C_I]^{\mathrm{T}}. \tag{2.288}$$

Here the first equality is the usual decomposition with matrices of rank one and is proved exactly as in linear algebra. The second equality follows from the fact that disjunction is idempotent and therefore we do not need to repeat summands of the disjunction which contain the same column and row; we can therefore choose a representative subset I of indices for which all the columns are different; since M is symmetric, I will also work for the rows and we can take the disjunction over products of columns and rows from I. For the third equality we write $[C_I]$ for the matrix whose columns are the columns of E with indices from I and likewise $[R_I]$ for the rows; the equality follows again from the decomposition with matrices of rank one. The last equality follows from the fact that M is symmetric and therefore the rows are obtained from the columns by transposition. Thus, $M(p) = [C_I]$. In our case we can take $I = 1, 2$ as this is a complete set of representatives for the columns; therefore

$$M = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} = M(p)\,M(p)^{\mathrm{T}} \tag{2.289}$$

and $M(p)$ is exactly as in (2.287).

The set of equivalence classes of an equivalence relation can also be characterized abstractly through the notion of partition.

Definition 2.103. Given a set A, a *partition* of A is assigned by a set $P = \{A_i \subseteq A : i \in I\}$ of subsets of A satisfying the following conditions.

$$\forall i (A_i \neq \varnothing) \qquad \text{Nonemptyness} \tag{2.290}$$

$$\forall ij (i \neq j \to A_i \cap A_j = \varnothing) \qquad \text{Disjointness} \tag{2.291}$$

$$A = \bigcup_{i \in I} A_i \qquad \text{Covering} \tag{2.292}$$

If P is a partition of A, the relation $E : A \rightharpoonup A$ induced by P is defined by the formula

$$xEy \Leftrightarrow \exists i (x, y \in A_i). \tag{2.293}$$

The characterization is the following.

Proposition 2.104. Partitions on A classify equivalences on A. More precisely:

1. If E is an equivalence relation on A, then its equivalence classes form a partition of A.
2. If P is a partition of A, then the induced relation is an equivalence relation.

The two constructions are mutually inverse.

Proof. ① Assume $E : A \rightharpoonup A$ is an equivalence relation and let $P = \{[x] : x \in A\}$ be the set of its equivalence classes. Every equivalence class is non-empty, because xEx and therefore $x \in [x]$. Different equivalence classes are disjoint by proposition 2.96. Since $x \in [x]$ we have $\{x\} \subseteq [x] \subseteq A$ and therefore

$$A = \bigcup_{x \in A} \{x\} \subseteq \bigcup_{x \in A} [x] \subseteq A \tag{2.294}$$

whence $A = \bigcup_x [x]$.This proves that P is a partition. ② Assume $P = \{A_i \subseteq A : i \in I\}$ is a partition and let $E : A \rightharpoonup A$ be the induced relation. If $x \in A$ then $x \in A_i$ for some $i \in I$ by the covering condition; hence $x, x \in A_i$, xEx and E is reflexive. If xEy then $x, y \in A_i$ for some i, hence yEx and E is symmetric. If xEy and yEz then there exist indices i and j suche that $x, y \in A_i$ and $y, z \in A_j$; however $y \in A_i \cap A_j$ and by disjointness, $i = j$; thus $x, z \in A_i$, xEz and E is transitive. This proves that E is an equivalence. ③ Assume now E is an equivalence relation, P is the partition on A induced by E and F is the equivalence relation induced by P. Then

$$xEy \Leftrightarrow [x] = [y] \Leftrightarrow xFy. \tag{2.295}$$

so that $F = E$. Conversely, assume P is a partition on A, E is the equivalence induced by P and Q is the partition induced by E. If $A_i \in P$ there exists $x \in A_i$ by the nonemptyness condition. Now $y \in [x] \Leftrightarrow xEy \Leftrightarrow y \in A_i$; thus $A_i = [x]$, the sets of the two partitions are the same and $P = Q$. □

Thus, quotient sets of A by equivalence relations are essentially the same as partitions of A.

Example 2.105 Let $A = \{0, 1, 2, 3, 4\}$. Then

$$P = \{\{0, 1\}, \{2, 3\}, \{4\}\} \tag{2.296}$$

is a partition of A. The corresponding equivalence relation $E : A \rightharpoonup A$ has incidence matrix

$$M(E) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.297}$$

□ **Kernel pairs.** The most important equivalence relation on any set A is equality. In fact, every equivalence relation is obtained from equality by a base change.

PROPOSITION 2.106. Equivalence relations are stable under inverse images. More precisely, if $f : B \to A$ is any function, the inverse image along f of every equivalence relation on A is an equivalence relation on B

Proof. Assume, in the diagram below, that E is an equivalence relation on A and that $F := fEf^{\mathrm{op}}$ is its inverse image along f.

$$\begin{array}{ccc} B & \xrightarrow{f} & A \\ {\scriptstyle F}\,\vdots & & \Big\downarrow{\scriptstyle E} \\ B & \xrightarrow[f]{} & A \end{array} \tag{2.298}$$

We prove that F is an equivalence relation on B. ① F is reflexive because

$$B \subseteq ff^{\mathrm{op}} = fAf^{\mathrm{op}} \subseteq fEf^{\mathrm{op}} = F \tag{2.299}$$

where the first inclusion follows from the fact f is a function (corollary 2.50); the second equality from the fact that A is neutral for the product; the last inclusion from the fact that $A \subseteq E$ because E is reflexive

and from compatibility of the product with inclusion. Since $B \subseteq F$, F is reflexive. ② F is symmetric because

$$F^{\mathrm{op}} = (fEf^{\mathrm{op}})^{\mathrm{op}} = f^{\mathrm{op}\,\mathrm{op}}E^{\mathrm{op}}f^{\mathrm{op}} = fEf^{\mathrm{op}} = F \tag{2.300}$$

where the first equality follows from the definition of F; the second from the compatibility of the opposite with products; the third from the fact that the opposite is involutory and E is symmetric; the last from the definition of F. ③ F is transitive, because

$$F^2 = fEf^{\mathrm{op}}fEf^{\mathrm{op}} \subseteq fEAEf^{\mathrm{op}} = fE^2f^{\mathrm{op}} = fEf^{\mathrm{op}} = F \tag{2.301}$$

where the first equality follows from the definition of F; the second inclusion from the fact that $f^{\mathrm{op}}f \subseteq A$ because f is a function (corollary 2.50) and from compatibility of the product with inclusion; the third equality from the fact that A is neutral for the product; the fourt from the fact that E is an equivalence relation; the last from the definition of F. □

Definition 2.107. If $f : A \to B$ is any function, the *kernel pair* of f is the inverse image along f of the equality relation on B. It is denoted by $\ker(f)$. Thus,

$$\ker(f) = f^*B = ff^{\mathrm{op}}. \tag{2.302}$$

Proposition 2.108. Equivalente relations on a set A are exactly the kernel pairs of functions with domain A.

Proof. ① Every kernel pair of a function $f : A \rightharpoonup B$ is an equivalence relation on A. This follows from the fact that the equality relation on B is an equivalence relation, $\ker(f)$ is the inverse image of this relation along f by definition, and equivalence relations are stable under inverse images (proposition 2.106). ② If E is an equivalence relation on E and $p : A \to A/E$ is the projection on the quotient, then $E = pp^{\mathrm{op}}$ (proposition 2.101) which means that $E = \ker(p)$. □

Observe that given an equivalence relation E on A there are many funcions f with domain A having E as their kernel pair; we will see later, however, that if we add the restriction that f be surjective, then the function is essentially unique. Also observe that if we start with an equivalence E, represnt it as the kernel pair of some function f and then compute ther kernel pair of f we get back the relation E. Howevere, the function we use to recosntruct the kernel pair of any function f is the projection and not the function itself, as there is no way to reconstruct a function only knowing its kernel pair.

Example 2.109 We have already seen that semantical equivalence is an equivalence relation (example 2.79). We provide a different perspective and examine further details here. If L is apropositional language with variables X and if $V = \Omega^X$ is the set of valuations of L, we can define a function $[\![\_]\!] : L \to \Omega^V$ which associates to any formula φ its truth function $[\![\varphi]\!]$. By definition, two formulas are equivalent when they have the same truth function (definition 1.33), $\varphi \equiv \psi \Leftrightarrow [\![\varphi]\!] = [\![\psi]\!]$; this shows that semantical equivalence is the kernel pair of $[\![\_]\!]$ and therefore an equivalence relation on L (proposition 2.108).

Now call a function $t : V \to \Omega$ a *truth function*. Every formula $\varphi \in L$ induces a truth function, but not all truth functions come from propositional formulas. This depends on the fact that every formula φ has finitely many variables and therefore $[\![\varphi]\!](a) = [\![\varphi]\!](b)$ if a and b are valuations attaining the same values on all variables of φ. If we order valuations by increasing binary enumeration, this means that once we have reached the last variable of φ, say at step 2^n, the pattern of truth values attained by $[\![\varphi]\!]$ will repeat with periodicity 2^n. Thus, any truth function which does not have a periodicity of 2^n for some n can not be induced by a formula. One such function can be easily defined setting $t(a_{2^n}) = 1$ and zero otherwise, where a_{2^n} is the 2^n-th valuation. In other words, the function $[\![\_]\!]$ is not surjective.

If, however, X is finite of cardinality n, then V has only 2^n elements. If t is a truth function, we can regard it as a connective of arity n and by functional completeness, there is a formula that realizes t, i.e. such that $t = [\![\varphi]\!]$. Thus, when X is finite $[\![\_]\!]$ is surjective and therefore

$L/\equiv \simeq \Omega^V$: the set of semantically equivalent classes of formulas can be identified with the set of truth functions — or truth tables. In particular, since $|V| = |\Omega^X| = 2^n$, L has 2^{2^n} equivalence classes of formulas.

Example 2.110 The example above can be generalized as we did in example 2.79. Consider an arbitrary subset $V \subseteq \Omega^X$ of the set of all valuations on L. Again we have a function $[\![\_]\!]_V : L \to \Omega^V$ which associates to $\varphi \in L$ the restriction of its truth function to V. Then

$$(\varphi, \psi) \in \ker([\![\_]\!]_V) \Leftrightarrow \forall v([\![\varphi]\!]_V(v) = [\![\psi]\!]_V(v)) \tag{2.303}$$

$$\Leftrightarrow \forall v \in V(v(\varphi) = v(\psi)) \tag{2.304}$$

$$\Leftrightarrow \varphi \underset{V}{\equiv} \psi \tag{2.305}$$

This proves that equivalence modulo V is the kernel pair of $[\![\_]\!]_V$ and therefore an equivalence relation. When X is finite, every truth function $t \in \Omega^V$ is the restriction of a truth function on all valuation and therefore comes from a formula. Thus, $L/\equiv_V \simeq \Omega^V$ and the quotient set can be identified with the set of partial truth tables.
This example includes the case of equivalence modulo a theory T, where we can take $V = \{vin\Omega^X : v \vDash T\}$. For a concrete case, let L be the propsitional language generated by variables $\{x, y\}$ and by the standard set of connectives C. Further, let T be the theory in L axiomatised by $A = \{x \vee y\}$. Observe that since $|X| = 2$ there are $2^2 = 4$ valuations on L, but only three of them satisfy T, those in the rows of the table below.

x	y	$[\bot]$	$[\neg x]$	$[\neg y]$	$[\neg(x \wedge y)]$	$[x \wedge y]$	$[y]$	$[x]$	$[\top]$
0	1	0	1	0	1	0	1	0	1
1	0	0	0	1	1	0	0	1	1
1	1	0	0	0	0	1	1	1	1

(2.306)

With 3 valuations, we have $2^3 = 8$ truth functions, or truth tables, corresponding to the columns of the table. The headings of the column provide formulas in the equivalence class whose truth table is the given column. Observe that every equivalence class contains infinitely many formulas. However, the quotient set is finite:

$$L/\equiv_T = \{[\bot], [\neg x], [\neg y], [\neg(x \wedge y)], [x \wedge y], [y], [x], [\top]\}. \tag{2.307}$$

There is another interesting aspect of kernel pairs: from proposition 2.61 we know that a function $f : A \to B$ is injective if and only if its kernel pair is the equality relation on A. One may wonder if a similar characterization exists or surjectivity. The problem here is that the kernel pair of f is an equivalence relation; however, f^{op} is not a function, in general, and its kernel pair is not, by consequence, an equivalence relation. One can remedy this using the generated equivalence, but this is outside the scope of our discussion.

2.5 Orders

□ Partial orders.

Definition 2.111. A binary relation $R : A \rightharpoonup A$ on a set is a *partial order* if it is reflexive, transitive and antisymmetric, i.e. if it satisfies the following formulas.

$$\forall x(xRx) \qquad \text{Reflexivity} \tag{2.308}$$

$$\forall xyz(xRy \wedge yRz \to xRz) \qquad \text{Transitivity} \tag{2.309}$$

$$\forall xy(xRy \wedge yRx \to x = y) \qquad \text{Antisymmetry} \tag{2.310}$$

A *partially ordered set* is a pair (A, R) consisting of a set A together with a partial order relation R on A.

A partial order is also simply called an *order*. The standard symbol used for a generic partial order relation is $\leq$. The notion of partial order can be formulated in a more algebraic fashion as follows.

PROPOSITION 2.112. A binary relation $R : A \rightharpoonup A$ is a partial order if and only if it satisfies the following formulas.

$$R^2 = R \qquad \text{Idempotency} \tag{2.311}$$

$$R \cap R^{\mathrm{op}} = I \qquad \text{Strict antisymmetry} \tag{2.312}$$

Proof. ① Necessity. Assume R is a partial order. Reflexivity means that $I \subseteq R$ and transitivity that $R^2 \subseteq R$. Therefore,

$$R = RI \subseteq RR \subseteq R \tag{2.313}$$

where the first equality follows from the fact that the identity relation is neutral for the product; the second inequality follows from the fact that $I \subseteq R$ and compatibility of the product with inclusion; the last from transitivity of R. From the double inclusion we deduce that $R^2 = R$. Next, observe that antisymmetry can be restated as $R \cap R^{\mathrm{op}} \subseteq I$. Thus

$$I = I \cap I^{\mathrm{op}} \subseteq R \cap R^{\mathrm{op}} \subseteq I \tag{2.314}$$

where the first equality follows from the fact that $I^{\mathrm{op}} = I$ by compatibility of the opposite with the multiplicative structure and idempotency of intersection; the second inclusion follows from the fact that $I \subseteq R$ by reflexivity, that $I^{\mathrm{op}} \subseteq R^{\mathrm{op}}$ by compatibility of the opposite with inclusion; the last inclusion follows from antisymmetry. The double inclusion implies $R \cap R^{\mathrm{op}} = I$. ② Sufficiency. Since $R^2 = R \subseteq R$ the relation is transitive. Since $I = R \cap R^{\mathrm{op}} \subseteq R$ the relation is reflexive. Finally, since $R \cap R^{\mathrm{op}} = I \subseteq I$, it is also antisymmetric. □

It is worth observing that, despite the name, antisymmetry is not the opposite concept of symmetry: a relation can be both symmetric and antisymmetric or it can be neither.

Example 2.113 On the set $A = \{0, 1, 2\}$ conside the relations R and S represented by the following matrices.

$$M(R) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \qquad M(S) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{2.315}$$

Then R is both symmetric and antisymmetric, whereas S is neither.
More generally, observe that if $R : A \rightharpoonup A$ is both symmetric and antisymmetric and xRy, then yRx by symmetry and therefore $x = y$ by antisymmetry. Thus, the only pairs belonging to R are of the form (x, x) and R must be a subset of the equality relation.
On the other hand, for R to be neither symmetric nor antisymmetric it must have at least one pair $(x, y) \in R$ such that $(y, x) \notin R$ — which implies that $x \neq y$ — and at least one pair $(x, y) \in R$ such that $(y, x) \in R$ with $x \neq y$.

The following are basic examples of partial orders.

Example 2.114 (Strict orders) A relation $R : A \rightharpoonup A$ is *irreflexive* if $\forall x(\neg xRx)$. We say that R is a *strict order* if it is irreflexive and transitive. The standard symbol for a generic strict order relation is $<$. Orders and strict orders are essentially the same concept.
For suppose we have a partial order relation $\leq$ on A; we can then define a strict order setting $x < y \leftrightarrow (x \leq y) \wedge (x \neq y)$. Irreflexivity is immediate, because by definition $x < x \leftrightarrow x \leq y \wedge x \neq x$; the last condition in the conjunction is not satisfied, so $x \not< x$. For transitivity, assume $x < y$ and $y < z$. Then $x \leq y$ and $y \leq z$, so that $x \leq z$ by transitivity of the partial order; if now $x = z$ we would have $y \leq z = x$ and $x = y$ by antisimmetry of the partial order, contradicting the assumption $x < y$; thus $x \neq z$ and therefore $x < z$.
Conversely, assume $<$ is a strict order. Define a partial order setting $x \leq y \leftrightarrow (x < y) \vee (x = y)$. Reflexivity is immediate, because $x \leq x \leftrightarrow (x < x) \vee (x = x)$ and the last part of the disjunction is satisfied, so that $x \leq x$. To prove transitivity, assume $x \leq y$ and $y \leq z$. If $x = y$ or $y = z$ then we immediately obtain $x \leq z$ by substitution in the second or in the first inequality respectively. Otherwise, $x < y$ and $y < z$, hence $x < z$ by transitivity of the strict order and therefore $x \leq z$. To prove antisymmetry, assume $x \leq y$ and $y \leq x$. If $x \neq y$ we would have $x < y$ and $y < x$, hence $x < x$ by transitivity, which contradicts irreflexivity; therefore $x = y$ and $\leq$ is a partial order.

Example 2.115 On the set $A = \{0, 1, 2, 3\}$ consider the relation R represented by the matrix

$$M = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.316}$$

To prove that R is a partial order it suffices to verify that $R \cap R^{\mathrm{op}} = A$ and $R^2 = R$; recalling that $M(R \cap R^{\mathrm{op}}) = M(R) \wedge M(R^{\mathrm{op}}) = M \wedge M^{\mathrm{T}}$ and that $M(R^2) = M(R)\,M(R) = M^2$, the verification amounts to

$$M \wedge M^{\mathrm{T}} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \wedge \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = I, \tag{2.317}$$

$$M^2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = M. \tag{2.318}$$

Example 2.116 (Hasse diagrams) A partial order relation $R : A \rightharpoonup A$ on a finite set is represented by a *Hasse diagram* which is constructed from the graph of R in three steps.

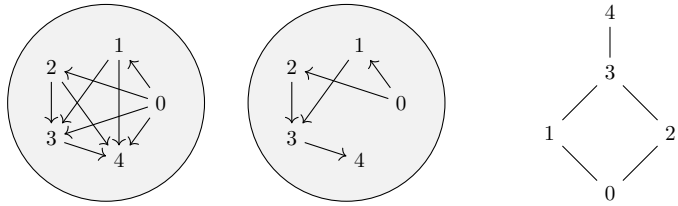
1. Since we know that R is reflexive, we will consider all the loops on every element as understood and will omit them. This amounts to say that instead of representing R we are representing the irreflexive relation S whose reflexive closure is R.
2. Since R is transitive so is S and we know that whenever we have a sequence of arrows $x_1 \to x_2 \to \cdots \to x_n$ in the graph of S, there must be an arrow $x_1 \to x_n$ between the endpoints. Therefore, we will omit an arrow $x_1 \to x_n$ if there is a path of greater length joining x_1 to x_n in S. This means that instead of representing S, we are representing the smallest relation $T \subseteq S$ whose transitive closure is S.
3. Finally, since xRy in a partial order is usually understood as '*x is smaller than or equal to y*', we will rearrange the arrows so that if $x \to y$ in the graph of T, then y sits above x, the idea being that elements in a higher position are bigger than elements lying below them. Since the higher position already indicates the direction of the arrow, we will omit arrow heads and simply draw segments.

Conversely, any Hasse diagram in which distinct elements of a set A are placed on top of each other with connecting segments and no segment replaceable by a composable path defines a partial order R: declare that xTy if there is an ascending segment from x to y and then take the reflexive and transitive closure R of T. Observe that Hasse diagrams can not be used to represent relations other than orders.

Example 2.117 For a concrete example on Hasse diagrams, consider the set $A = \{0, 1, 2, 3, 4\}$ and the partial order relation $R : A \rightharpoonup A$ with incidence matrix

$$M(R) = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \tag{2.319}$$

The construction of the Hasse diagram of R is illustrated below. In the first graph we remove all the loops; that is, we consider the relation S such that $SR = R$. In the second graph we remove all arrows if they can be replaced by a path of greater length; this means that we are replacing S with the smallest T such that $R = T^{rt}$. The final image is the Hasse diagram, in which we have placed y above x when xTy.



(2.320)

Example 2.118 For every set A, the inclusion relation $\subseteq$ on the powerset PA is a partial order relation,

$$x \subseteq y := \forall w(w \in x \to w \in y). \tag{2.321}$$

The relation is reflexive because

$$\forall w(w \in x \to w \in x) \Rightarrow x \subseteq x, \tag{2.322}$$

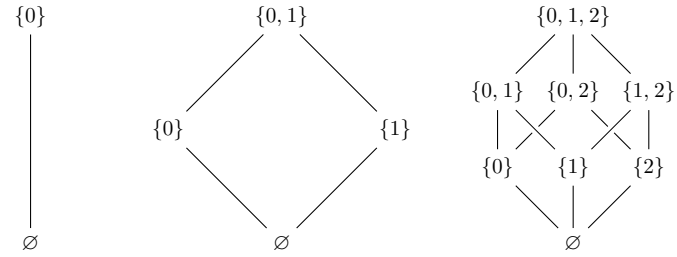
it is transitive because

$$\begin{aligned} (x \subseteq y) \wedge (y \subseteq z) &\Rightarrow \forall w(w \in x \to w \in y) \wedge \forall w(w \in y \to w \in z) && (2.323)\\ &\Rightarrow \forall w((w \in x \to w \in y) \wedge (w \in y \to w \in z)) && (2.324)\\ &\Rightarrow \forall w(w \in x \to w \in z)) && (2.325)\\ &\Rightarrow x \subseteq z && (2.326) \end{aligned}$$

and it is antisymmetric because

$$\begin{aligned} x \subseteq y \wedge y \subseteq x &\Rightarrow \forall w(w \in x \to w \in y) \wedge \forall w(w \in y \to w \in x) && (2.327)\\ &\Rightarrow \forall w((w \in x \to w \in y) \wedge (w \in y \to w \in x)) && (2.328)\\ &\Rightarrow \forall w(w \in x \leftrightarrow w \in y)) && (2.329)\\ &\Rightarrow x = y && (2.330) \end{aligned}$$

The following are the Hasse diagrams of the inclusion when $A = \{0\}$, $\{0, 1\}$ and $\{0, 1, 2\}$.



(2.331)

Example 2.119 (Divisibility) The divisibility relation on the set $\mathbf{N}$ of natural numbers is defined by the formula

$$x|y := \exists u(xu = y) \tag{2.332}$$

and is a partial order. Reflexivity follows from the observation that $x1 = x \to x|x$. Transitivity is proved observing that

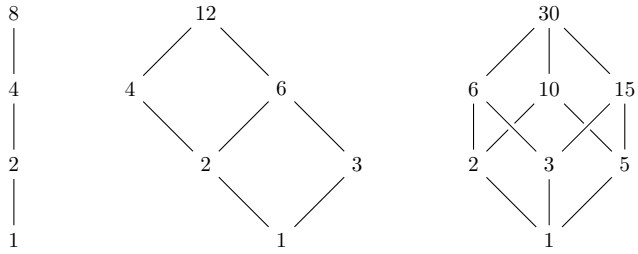
$$\begin{aligned}(x|y) \wedge (y|z) &\Rightarrow \exists uv((xu = y) \wedge (yv = z)) && (2.333)\\ &\Rightarrow \exists uv(x(uv) = z) && (2.334)\\ &\Rightarrow \exists w(xw = z) && (2.335)\\ &\Rightarrow x|z && (2.336)\end{aligned}$$

For antisymmetry we have

$$\begin{aligned}(x|y) \wedge (y|x) &\Rightarrow \exists uv((xu = y) \wedge (yv = x)) && (2.337)\\ &\Rightarrow \exists uv(xuv = x) && (2.338)\\ &\Rightarrow x = 0 \vee uv = 1 && (2.339)\\ &\Rightarrow x = 0 \vee (u = v = 1) && (2.340)\\ &\Rightarrow x = y. && (2.341)\end{aligned}$$

It is worth observing that divisibility is not a partial order relation on the set $\mathbf{Z}$ of integers, because antisymmetry fails; in fact, for every $x \neq 0$ we have $x|-x$ and $-x|x$, yet $x \neq -x$.

Example 2.120 If $S \subseteq \mathbf{N}$ is any subset, the divisibility relation is also a partial order on S because the proof given in exercise 2.119 is local in nature. This is of particular interest when S is the set $D(n)$ of natural divisors of $n \in \mathbf{N}$. Below are the Hasse diagrams of $D(8)$, $D(12)$ and $D(30)$.



(2.342)

Observe that the shape of the Hasse diagram does not depend on n but rather on its prime factorization and more precisely on the number of primes and their multiplicity. Thus, the first diagram applies to all natural numbers of type $n = p^3$; the second to $n = p_1^2 p_2$ where p_1 and p_2 are distinct primes; the third to $n = p_1 p_2 p_3$.

Example 2.121 Assume (A, R) is a partially ordered set and B is an arbitrary set. Then the set $A^B = \{f : B \to A\}$ of all functions from B to A is partially ordered by the relation R^B defined by

$$f R^B g := \forall b(f(b)\, R g(b))\,. \tag{2.343}$$

The relation is reflexive because $f(b) = f(b)$ for every $b \in B$ and since R is reflexive $f R^B g$. Transitivity is proved observing that

$$\begin{aligned}(f R^B g) \wedge (g R^B h) &\Rightarrow \forall b(f(b)\, R g(b)) \wedge \forall b(g(b)\, R h(b)) && (2.344)\\ &\Rightarrow \forall b(f(b)\, R g(b) \wedge g(b)\, R h(b)) && (2.345)\\ &\Rightarrow \forall b(f(b)\, R h(b)) && (2.346)\\ &\Rightarrow f R^B h. && (2.347)\end{aligned}$$

(2.348)

Antisymmetry is proved observing that

$$(fR^B g) \wedge (gR^B f) \Rightarrow \forall b(f(b)\,Rg(b)) \wedge \forall b(g(b)\,Rf(b)) \tag{2.349}$$
$$\Rightarrow \forall b(f(b)\,Rg(b) \wedge g(b)\,Rf(b)) \tag{2.350}$$
$$\Rightarrow \forall b(f(b) = g(b)) \tag{2.351}$$
$$\Rightarrow f = g \tag{2.352}$$

More examples can be obtained from the following observation.

PROPOSITION 2.122. If $R : A \rightharpoonup A$ is a partial order relation on A, so is R^{op}.

Proof. Observe that

$$R^{\mathrm{op}} R^{\mathrm{op}} = (RR)^{\mathrm{op}} = R^{\mathrm{op}} \tag{2.353}$$

so that R^{op} is idempotent. Also,

$$R^{\mathrm{op}} \cap R^{\mathrm{op}\,\mathrm{op}} = R^{\mathrm{op}} \cap R = R \cap R^{\mathrm{op}} = I \tag{2.354}$$

so that R^{op} satisfies strict antisymmetry. Thus, R^{op} is a partial order. □

□ Total orders.

DEFINITION 2.123. A partial order relation $R : A \rightharpoonup A$ is a *total order* if it satisfies the formula

$$\forall xy(xRy \vee yRx) \qquad \text{Comparability} \tag{2.355}$$

A *totally ordered set* is a pair (A, R) consisting of a set A and a total order relation on A.

The comparability condition means that given any two elements one of them will be greater than or equal to the other. Since yRx is equivalent to $xR^{\mathrm{op}}y$, the comparability condition can be reformulated as $R \cup R^{\mathrm{op}} = \top$, the total relation.

Example 2.124 On the set $A = \{0, 1, 2, 3\}$ consider the relation R defined by the incidence matrix

$$M = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix}. \tag{2.356}$$

R is a total order relation because

$$M \wedge M^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} \wedge \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = I \tag{2.357}$$

$$M^2 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} = M \tag{2.358}$$

$$M \vee M^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} \vee \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = M(\top) \tag{2.359}$$

The Hasse diagram of R, shown below, is a *chain* with all elements in a single strand; this is a consequence of the comparability axioms: given any two distinct elements, one of them must be greater than the other and this prevents the existence of elements at the same height, as these would not satisfy this condition.

$$\begin{matrix} 2 \\ | \\ 0 \\ | \\ 3 \\ | \\ 1 \end{matrix} \tag{2.360}$$

Example 2.125 There is a total order relation on the set $\mathbf{N}$ of natural numbers defined by the formula

$$x \leq y := \exists u(x + u = y)\,. \tag{2.361}$$

The relation is reflexive because $x + 0 = x \to x \leq x$. For transitivity we have

$$\begin{aligned}(x \leq y) \wedge (y \leq z) &\Rightarrow \exists uv((x + u = y) \wedge (y + v = z)) && (2.362)\\ &\Rightarrow \exists uv(x + (u + v) = z) && (2.363)\\ &\Rightarrow \exists w(x + w = z) && (2.364)\\ &\Rightarrow x \leq z && (2.365)\end{aligned}$$

The relation is antisymmetric because

$$\begin{aligned}(x \leq y) \wedge (y \leq x) &\Rightarrow \exists uv((x + u = y) \wedge (y + v = x)) && (2.366)\\ &\Rightarrow \exists uv(x + (u + v) = x) && (2.367)\\ &\Rightarrow u + v = 0 && (2.368)\\ &\Rightarrow u = v = 0 && (2.369)\\ &\Rightarrow x = y. \end{aligned}$$

Antisymmetry is proved as follows:

$$\begin{aligned}(x \leq y) \wedge (y \leq x) &\Rightarrow \exists uv[(x + u = y) \wedge (y + v = x)] && (2.370)\\ &\Rightarrow x + (u + v) = x && (2.371)\\ &\Rightarrow u + v = 0 && (2.372)\\ &\Rightarrow u = v = 0 && (2.373)\\ &\Rightarrow x = y && (2.374)\end{aligned}$$

□ **Generated and induced orders.** We have seen that every binary relation R on a set A is contained in a smallest equivalence relation. This is not true for partial orders: although partial orders on a A are closed under intersections, we can not replicate the argument used for equivalence relations becuase there may be no partial order containing R. This happens, for example, if R is not antisymmetric, i.e. if there exist distinct elements $x, y \in A$ such that $(x, y), (y, x) \in R$. For in this case every relation caontaining R would necessarily have these two elements and would therefore not be antisymmetric either. The situation requires more analysis.

Proposition 2.126. Let $R : A \rightharpoonup A$ be a binary relation and let $T := R^{rt}$ be its reflexive and transitive closure. If T is antisymmetric, then it is the smallest partial order relation containing R; otherwise, R is not contained in any order relation.

Proof. Observe that every partial order U containing R is reflexive and transitive by definition, and must therefore contain the reflexive and transitive closure T of R. If now T is antisymmetric then it is a partial order and therefore the smallest partial order containing R. If, however, T is not antisymmetric there exist two distinct elemenets $x, y \in A$ such that $(x, y), (y, x) \in T$; any relation containing T must then have this two pairs as elements and can not be antisymmetric; thus, there is no partial order containing R in this case. □

To compute the reflexive and transitive closure we can use lemma 2.93 when A is finite, for we have an ascending chain $S \subseteq S^2 \subseteq S^3 \subseteq \cdots \subseteq A^2$ and since A^2 is finite the sequence must necessarily stabilize after a finite number of steps.

Example 2.127 On the set $A = \{0, 1, 2, 3\}$ consider the binary relation R defined by the incidence matrix

$$M(R) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{2.375}$$

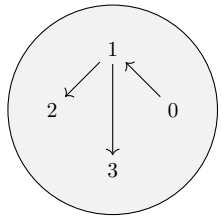
To compute the reflexive and transitive closure T of R we compute successive powers of the matrix $N = M(A \cup R)$ of the reflexive closure of R. We have

$$N = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad N^2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad N^3 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.376}$$

and since $N^2 = N^3$, we have $M(T) = N^2$. Now observe that $N^2 \wedge (N^2)^{\mathrm{T}} = A$; thus T is antisymmetric and is the order closure of R. One can also compute T from the graph of R as follows: since

$$T = R^{tr} = R^t \cup A = \left(\bigcup_{m=1}^{\infty} R^m \right) \cup A = \bigcup_{m=0}^{\infty} R^m \tag{2.377}$$

then xTy when there is a finite path $x \to \cdots \to y$ built from arrows of R. From the graph of R we then immediately obtain the matrix of T.

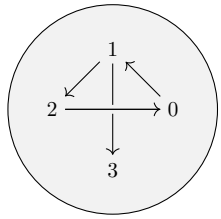


$$M(T) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.378}$$

Example 2.128 On the set $A = \{0, 1, 2, 3\}$ Consider the binary relation R represented by the matrix

$$M(R) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{2.379}$$

We compute its reflexive and transitive closure T from the graph.



$$M(T) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.380}$$

Observe that T is not antisymmetric — it suffices to observe that $0T1$ and $1T0$, for example — hence R is not contained in any order relation.

We are left with the problem of constructing an order when T is not antisymmetric. The idea here is to collapse antisymmetric pairs into a single element using an equivalence relation.

Definition 2.129. Assume R and S are binary relations as shown in the diagram below. A *morphism* of relations from R to S is a function f as in the diagram such that $fS = Rf$. If f is an isomorphism, we say that it is an isomorphism of relations.

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle R}\downarrow & & \downarrow{\scriptstyle S} \\ A & \xrightarrow[f]{} & B \end{array} \tag{2.381}$$

If f is an isomorphism of relations from R to S then R and S behave in the same way with f acting as a translation layer. For example, if $f(x) = x'$ and $f(y) = y'$, then from $Rf = fS$ and the fact that f is an isomorphism, we have $R = Rff^{\mathrm{op}} = fSf^{\mathrm{op}}$ and therefore

$$xRy \Leftrightarrow xfSf^{\mathrm{op}}y \Leftrightarrow x'Sy'. \tag{2.382}$$

Theorem 2.130. Let $R : A \rightharpoonup A$ be a binary relation and let $T = R^{rt}$ be its reflexive and transitive closure. Then the relation $E := T \cap T^{\mathrm{op}}$ is an equivalence relation on A and if $p : A \to A/E$ is the projection on the quotient, the direct image

$$p_*T = p^{\mathrm{op}}Tp \tag{2.383}$$

of T along p is a partial order on A/E, called the partial order *induced* by R. If T is antisymmetric and is therefore the generated order, then p_*T is isomorphic to T.

Proof. We refer to the diagram below.

$$\begin{array}{ccc} A & \xrightarrow{T} & A \\ {\scriptstyle p}\downarrow & & \downarrow{\scriptstyle p} \\ A/E & \underset{U}{\dashrightarrow} & A/E \end{array} \tag{2.384}$$

① E is an equivalence relation. For reflexivity observe that

$$A = A \cap A = A \cap A^{\mathrm{op}} \subseteq T \cap T^{\mathrm{op}} = E \tag{2.385}$$

where the first equality follows from idempotency of intersection; the second from compatibility of the opposite with the multiplicative structure; the third from the fact that $A \subseteq T$ because T is reflexive, therefore $A^{\mathrm{op}} \subseteq T^{\mathrm{op}}$ by compatibility of the opposite with inclusions and finally from the compatibility of products with inclusions; the last equality from the definition of E. Thus $A \subseteq E$ and E is reflexive. To prove symmetry observe that

$$E^{\mathrm{op}} = (T \cap T^{\mathrm{op}})^{\mathrm{op}} = T^{\mathrm{op}} \cap T^{\mathrm{op}\,\mathrm{op}} = T^{\mathrm{op}} \cap T = T \cap T^{\mathrm{op}} = E \tag{2.386}$$

where the first equality follows from the definition of E; the second from compatibility of opposites with intersections; the third from the involution property of opposites; the fourth from commutativity of intersections; the last from the definition of E. To prove transitivity observe that

$$E^2 = (T \cap T^{\mathrm{op}})\,(T \cap T^{\mathrm{op}}) \subseteq TT \cap TT^{\mathrm{op}} \cap T^{\mathrm{op}}T \cap T^{\mathrm{op}}T^{\mathrm{op}} \subseteq TT \cap T^{\mathrm{op}}T^{\mathrm{op}} \subseteq T \cap T^{\mathrm{op}} = E \tag{2.387}$$

where the first equality follows from the definition of E; the second inclusion from compatibility of orders with intersections(formula (2.105)); the third inclusion from properties of the intersection; the fourth from

the fact that $TT \subseteq T$ by transitivity and therefore $T^{\mathrm{op}}T^{\mathrm{op}} \subseteq T^{\mathrm{op}}$ by compatibility of opposites with products and inclusions; the last from the definition of E. ② We let $U = p^{\mathrm{op}}Tp$ and prove that U is a partial order. It is transitive because

$$U^2 = (p^{\mathrm{op}}Tp)\,(p^{\mathrm{op}}Tp) = p^{\mathrm{op}}T1Tp = p^{\mathrm{op}}TTp \subseteq p^{\mathrm{op}}Tp = U \tag{2.388}$$

To prove that it is both reflexive and antisymmetric we show that $U \cap U^{\mathrm{op}} = 1$. Suppose V is a binary relation on A/E. Then

$$\begin{aligned} V \subseteq U \cap U^{\mathrm{op}} &\Leftrightarrow (V \subseteq U) \wedge (V \subseteq U^{\mathrm{op}}) && (2.389)\\ &\Leftrightarrow (V \subseteq p^{\mathrm{op}}Tp) \wedge (V \subseteq p^{\mathrm{op}}T^{\mathrm{op}}p) && (2.390)\\ &\Leftrightarrow (pVp^{\mathrm{op}} \subseteq ETE) \wedge (pVp^{\mathrm{op}} \subseteq E^{\mathrm{op}}T^{\mathrm{op}}E^{\mathrm{op}}) && (2.391)\\ &\Leftrightarrow (pVp^{\mathrm{op}} \subseteq T) \wedge (pVp^{\mathrm{op}} \subseteq T^{\mathrm{op}}) && (2.392)\\ &\Leftrightarrow pVp^{\mathrm{op}} \subseteq T \cap T^{\mathrm{op}} && (2.393)\\ &\Leftrightarrow pVp^{\mathrm{op}} \subseteq pp^{\mathrm{op}} && (2.394)\\ &\Leftrightarrow V \subseteq 1 \end{aligned}$$

where the first equivalence follows from the properties of intersections; the second from the definition of U; the third from multiplication by p on the left and p^{op} on the right and the fact that $pp^{\mathrm{op}} = E$ in one direction, and from multiplication by p^{op} on the left and by p on the right in the other direction; the fourth from the fact that $ETE = T$ and therefore $E^{\mathrm{op}}T^{\mathrm{op}}E^{\mathrm{op}} = T^{\mathrm{op}}$, because $E = T \cap T^{\mathrm{op}} \subseteq T$ and therefore $ETE \subseteq T^3 \subseteq T$ by transitivity of T and conversely, since $1 \subseteq E$ we have $T = 1T1 \subseteq ETE$; the fifth equivalence follows from properties of the intersection; the sixth from the fact that $T \cap T^{\mathrm{op}} = E = pp^{\mathrm{op}}$; the last from multiplication by p^{op} on the left and p on the right and the fact that $p^{\mathrm{op}}p = 1$ in one direction and from multiplication by p on the left and p^{op} on the right in the other direction. Thus, $U \cap U^{\mathrm{op}}$ and 1 have the same subsets and are therefore equal. ③ Assume now that T is antisymmetric. Since it is also reflexive, $\ker(p) = E = T \cap T^{\mathrm{op}} = 1$, so that p is injective; However, p is always surjective and therefore it is an isomorphism from T to U. □

The fact that $U = p^{\mathrm{op}}Tp$ means that p is a morphism of relations from T to U because

$$U = p^{\mathrm{op}}Tp \Leftrightarrow pU = ETp \Leftrightarrow pU = Tp \tag{2.395}$$

where the first equivalence follows from multiplication on the left by p in the right direction and by p^{op} in the left direction, plus the observation that $pp^{\mathrm{op}} = E$ and $p^{\mathrm{op}}p = 1$; the second equivalence follows from the fact that $ET = T$ because

$$T = 1T \subseteq ET = (T \cap T^{\mathrm{op}})T \subseteq TT \subseteq T. \tag{2.396}$$

We can then compute explicitly wity U because formula (2.383) amounts to

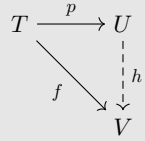
$$U[y] \Leftrightarrow xTu. \tag{2.397}$$

In fact,

$$uUv \Leftrightarrow upTp^{\mathrm{op}}v \Leftrightarrow \exists xy(up^{\mathrm{op}}x \wedge xTy \wedge yp^{\mathrm{op}}v) \Leftrightarrow \exists xy(u = [x] \wedge xTy \wedge [y] = v) \tag{2.398}$$

The induced order is the best order one can construct from R in the sense made explicit by the following proposition, whose proof is omitted.

PROPOSITION 2.131. Assume $R : A \rightharpoonup A$ is any relation, T its reflexive and transitive closure, $E = T \cap T^{\mathrm{op}}$ the induced equivalence and U the induced order on A/E. Then for every order relation V on A/E and and every morphism of relations f as in the diagram below, there exists a unique morphism of relations h such that $ph = f$.



(2.399)

Example 2.132 On the set $A = \{0, 1, 2, 3, 4\}$ consider the relation R represented by the matrix $M(R)$ below.

$$M(R) = \begin{pmatrix} 1&0&0&0&0 \\ 0&0&1&1&0 \\ 1&0&0&0&1 \\ 0&0&0&0&0 \\ 0&0&0&0&0 \end{pmatrix} \qquad M(T) = \begin{pmatrix} 1&1&1&1&1 \\ 1&1&1&1&1 \\ 1&1&1&1&1 \\ 0&0&0&1&0 \\ 0&0&0&0&1 \end{pmatrix} \tag{2.400}$$

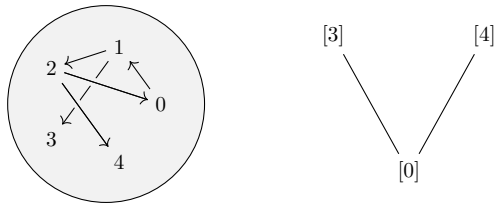
The reflexive and transitive closure of R is the relation T represented by $M(T)$ and is not antisymmetric because, for example, $0S1$ and $1S0$. The induced equivalence $E = T \cap T^{\mathrm{op}}$ on A and the projection $p : A \to A/E$ are represented by the matrices

$$M(E) = M(T) \wedge M(T)^{\mathrm{T}} = \begin{pmatrix} 1&1&1&0&0 \\ 1&1&1&0&0 \\ 1&1&1&0&0 \\ 0&0&0&1&0 \\ 0&0&0&0&1 \end{pmatrix}, \qquad M(p) = \begin{pmatrix} 1&0&0 \\ 1&0&0 \\ 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{pmatrix}. \tag{2.401}$$

The equivalence classes are therefore $[0] = \{0, 1, 2\} = [1] = [2]$, $[3] = \{3\}$ and $[4] = \{4\}$ and the quotient set is $A/E = \{[0], [3], [4]\}$. The induced order on A/E is therefore

$$M(U) = M(p^{\mathrm{op}} T p) = \begin{pmatrix} 1&1&1&0&0 \\ 0&0&0&1&0 \\ 0&0&0&0&1 \end{pmatrix} \begin{pmatrix} 1&1&1&1&1 \\ 1&1&1&1&1 \\ 1&1&1&1&1 \\ 0&0&0&1&0 \\ 0&0&0&0&1 \end{pmatrix} \begin{pmatrix} 1&0&0 \\ 1&0&0 \\ 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{pmatrix} = \begin{pmatrix} 1&1&1 \\ 0&1&0 \\ 0&0&1 \end{pmatrix}. \tag{2.402}$$

The graph of R and the Hasse diagram of U are shown below. Observe that the elements in T that violate antisymmetry have been collapsed into the single element $[0]$.



(2.403)

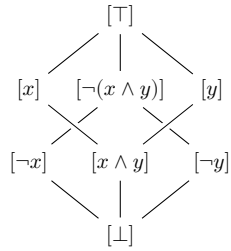
Example 2.133 Let L be a propositional language generated by a set X of variables and a set C of connectives. The satisfaction relation $\vDash\colon L \rightharpoonup L$ between formulas, as in $\varphi \vDash \psi$, is reflexive and transitive. To see this, recall that $\varphi \vDash \psi$ when $v(\varphi) \leq v(\psi)$ for every valuation $v \in \Omega^X$. Since $v(\varphi) \leq v(\varphi)$ for every valuation v, we have $\varphi \vDash \varphi$ and the relation is reflexive. For transitivity, assume $\varphi \vDash \chi$ and $\chi \vDash \psi$; then for every $v \in \Omega^X$ we have $v(\varphi) \leq v(\chi)$ by the first assumption and $v(\chi) \leq v(\psi)$ by the second; since $(\Omega, \leq)$ is a totally ordered set, we infer that $v(\varphi) \leq v(\psi)$ and therefore $\varphi \vDash \psi$, which proves transitivity.
Satisfaction is not antisymmetric, though: if we take $\varphi = x$ and $\psi = \neg\neg x$, then $\varphi \leq \psi$ and $\psi \leq \varphi$, because the two formulas are equivalent, although $\varphi \neq \psi$. We have assumed the existence of at least one variable in L, but we could replace x with $\bot$ or $\top$ and if none of these were in L, then L would be empty. In any case, satisfaction is not a partial order in L.
We can, however, construct the order relation induced by satisfaction. Since $\vDash$ is already reflexive and transitive, it is its own reflexive and transitive closure. The associated equivalence relation on L is precisely semantical equivalence, because if $\varphi \leq \psi$ and $\psi \leq \varphi$, we have $v(\varphi) \leq v(\psi)$ and $v(\psi) \leq v(\varphi)$ for every $v \in \Omega^X$; therefore $v(\varphi) = v(\psi)$ for every v and therefore $\varphi \equiv \psi$. If L is generated by a finite number of variables, the quotient set $L/\!\equiv$ can be identified with the set of truth tables of formulas (example 2.109) and the induced order on the quotient is defined by the formula

$$[\![\varphi]\!] \leq [\![\psi]\!] \Leftrightarrow (\varphi \vDash \psi)\,. \tag{2.404}$$

Example 2.134 The previous example extends to the case of semantical consequence $\varphi \vDash_V \psi$ relative to any subset $V \subseteq \Omega^X$ of valuations on L. If we take L to be the propositional language generated by variables $X = \{x, y\}$ and by the standard set of connectives C and T be the theory in L axiomatised by $A = \{x \vee y\}$, with $V = \big\{v \in \Omega^X : v \vDash T\big\}$, we have again that $\varphi \vDash_V \psi$ is reflexive and transitive and the associated equivalence is $\varphi \equiv_V \psi$. The quotient set was computed in example 2.110 as

$$L/\!\equiv_T = \ \{[\bot], [\neg x], [\neg y], [\neg(x \wedge y)], [x \wedge y], [y], [x], [\top]\}\,. \tag{2.405}$$

and the induced order on the quotient set is given by formula (2.404). We then obtain the following Hasse diagram for the order induced by $\vDash_T$.



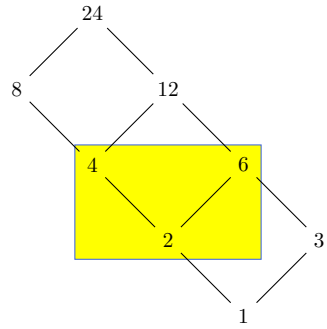
(2.406)

□ **The structure of partially ordered sets.** In this section we examine at a fairly superficial level the constituent elements of partially ordered sets.

Definition 2.135. Let $(A, \leq)$ be a partially ordered set and let $B \subseteq A$ be a subset. An element $x \in A$ is an *upper bound* for B if $y \leq x$ for every element $y \in B$. Dually, x is a *lower bound* for B if $x \leq y$ for every element $y \in B$.

Otherwise stated, the upper bounds for B in A are all elements of A which are greater than or equal to all elements of B, whereas the lower bounds are the elements of A which are less than or equal to all elements of B. Observe that lower bounds for B with respect to $\leq$ are exactly the upper bounds for B with respect to the opposite order $\leq^{\mathrm{op}}$.

Example 2.136 Consider the partially ordered set $(D(24), |)$ of natural divisors of 24 in $\mathbf{N}$, with the partial order relation given by divisibility. Consider also the subset $B = \{2, 4, 6\} \subseteq D(24)$.



(2.407)

Here x is an upper bound of B if it is divisible by all elements of B and it is a lower bound of B if it divides all elements of B. Therefore, the sets of upper and lower bounds for B in A are respectively

$$U = \{8, 12, 24\}, \qquad L = \{1, 2\}. \tag{2.408}$$

DEFINITION 2.137. Let $(A, \leq)$ be a partially ordered set and $B \subseteq A$ a subset. A *maximum* for B is an element of B that is greater than or equal to all elements of B. Dually, a *minimum* for B is an element of B that is less than or equal to all elements of B.

Thus, a maximum can also be defined as an upper bound for B which belongs to B; likewise, a minimum is a lower bound for B that belongs to B. As for upper bounds, a minimum for B with respect to $\leq$ is the same as a maximum for B with respect to the opposite order.

Example 2.138 In the situation described in example 2.136, the subset B does not have maximum, because none of the elements of the set U of upper bounds belongs to B. On the other hand, since $L \cap B = \{2\}$, 2 is the only minimum of B.

PROPOSITION 2.139. Assume $)A, \leq)$ is a partially ordered set and $B \subseteq A$. The maximum of B, it it exists, is unique; we denote it by $\max(B)$. Likewise, the minimum of B, it if it exists, is unique; we denote it by $\min(B)$.

Proof. ① Suppose $x, y \in A$ are both maxima for B. Since $x \in B$ and y is an upper bound for B, we have $x \leq y$. On the other hand, $y \in B$ and x is an upper bound, so that $y \leq x$. By antisimmetry, $x = y$. ② Observe that a minimum for B with respect to $\leq$ is the same thing as a maximum for B with respect to the opposite order $\leq^{\mathrm{op}}$. Since the maximum is unique, if it exists, so is the minimum. □

DEFINITION 2.140. Let $(A, \leq)$ be a partially ordered set and $B \subseteq A$ a subset. An element $x \in A$ is a *least upper bound* or *supremum* for B if x is a minimum in the set of upper bounds for B in A; it is denoted by $\sup(B)$. Dually, x is a *greatest lower bound* or *infimum* for B if x is a maximum in the set of lower bounds for B in A; it is denoted by $\inf(B)$.

Since maximum and minimum are unique when they exist, so are the supremum and infimum of B. As the other cases, the greatest lower bound is the least upper bound for the opposite order and conversely.

PROPOSITION 2.141. Let $(A, \leq)$ be a partially ordered set and $B \subseteq A$ a subset. If B has maximum x in A, then $x = \sup(B)$. Dually, if B has minimum x, then $x = \inf(B)$.

Proof. If $x = \max(B)$ then, by definition, $x \in U$ where U is the set of upper bound of B. Moreover, if $y \in U$ then $y \geq x$ because $x \in B$. Thus x is the minimum of U and therefore $x = \sup(B)$. The proof for the greatest lower bound is identical using the opposite order. □

Example 2.142 Consider the totally ordered set $(\mathbf{Z}, \leq)$ of the integers with

$$x \leq y := \exists n \in \mathbf{N}(x + n = y) \tag{2.409}$$

and let $B = \{x \in \mathbf{Z} : -1 \leq x < 3\}$. The sets of upper and lower bounds of B in $\mathbf{Z}$ are respectively

$$U = \{x \in \mathbf{Z} : x \geq 3\}, \qquad V = \{x \in \mathbf{Z} : x \leq -1\}. \tag{2.410}$$

Observe that $\min(U) = 3$ and therefore $3 = \sup(U)$. Likewise $\max(V) = -1$ and therefore $-1 = \inf(B)$. Observe that B does not have maximum but still has a least upper bound; on the other hand, it has minimum -1 which coincides with the greatest lower bound. Observe that least upper bound and greatest lower bound do not always exist, though. If we take $B = \{x \in \mathbf{Z} : x \leq 1\}$ then the set of its lower bounds is empty and therefore does not have maximum; thus, $\inf(B)$ does not exist in this case.

Example 2.143 Let A be an arbitrary set. Condier the partially ordered set $(PA, \subseteq)$ of subsets of A ordered by inclusion. If $Q \subseteq PA$ is a set consisting of subsets of A, then the least upper bound of Q is the union of all subsets of A belonging to Q and the greatest lower bound of Q is the intersection of all subsets of A belonging to Q.

$$\sup(Q) = \bigcup_{S \in Q} S, \qquad \inf(Q) = \bigcap_{S \in Q} S. \tag{2.411}$$

Example 2.144 Consider the partially ordered set $(\mathbf{N}, |)$ of natural numbers equipped with the divisibility relation. If $S \subseteq \mathbf{N}$ is any subset, an upper bound for S is an element $x \in \mathbf{N}$ which is a multiple of all the elements of S. Thus, the least upper bound of S, if it exists, is the least common multiple of the elements of S. Dually, the greatest lower bound of S is the greatest common divisor of the elements of S.

□ **Ordinals.** The von Neumann definition of natural numbers starts from $0 = \varnothing$ and defines the successor of n as the set $s(n) = n \cup \{n\}$. This produces the sets

$$0 = \{\}, \quad 1 = \{0\}, \quad 2 = \{0, 1\}, \quad 3 = \{0, 1, 2\}, \quad \cdots \tag{2.412}$$

Ordinal numbers generalize this construction to the infinite case.

DEFINITION 2.145. A totally ordered set $(A, \leq)$ is *well ordered* if every non-empty subset $B \subseteq A$ has a minimum. If A is well ordered and $x \in A$, the set

$$(x) = \{y \in A : y < x\} \tag{2.413}$$

is the *initial segment* generated by x.

It can be proved that any two well ordered sets are either isomorphic — with a unique isomorphism — or one and only one of them is isomorphic to an initial segment of the other. Thus, on the class of well ordered sets we have the equivalence relation of being isomorphic and on the quotient class of *order types* we have a total ordering provided by the initial segment embedding. An ordinal number should be an order type. If we wish to avoid classes, we have to choose a representative in each order type.

DEFINITION 2.146. A set A is *transitive* if $x \in y \in A \to x \in A$. An *ordinal number* is a transitive set which is well ordered by the membership relation

$$x < y \leftrightarrow x \in y. \tag{2.414}$$

All natural numbers are ordinals. For ordinal numbers we have $\alpha \in \beta \Leftrightarrow \alpha \subset \beta$; when this happens we write $\alpha < \beta$ and this defines a well ordering on the class of ordinals. Moreover, for every ordinal α we have

$$\alpha = \{\beta : \beta < \alpha\}. \tag{2.415}$$

It can be proved that every well ordered set is isomorphic to exactly one ordinal number and that there are two types of ordinal numbers:

$$\alpha + 1 = \alpha \cup \{\alpha\} = \inf\{\beta : \beta > \alpha\}, \qquad \alpha = \sup\{\beta : \beta < \alpha\} = \bigcup \alpha. \tag{2.416}$$

Ordinals of the first type are called *successor* ordinals, those of the second type are *limit* ordinals.

Definition 2.147. Let $(A, \leq)$ be a partially ordered set. An element $x \in A$ is *maximal* if there is no element of A which is strictly greater than x, i.e. if

$$\forall y(x \leq y \to x = y). \tag{2.417}$$

Dually, $x \in A$ is minimal if there is no element of A which is strictly less than x, i.e.

$$\forall y(y \leq x \to y = x). \tag{2.418}$$

Definition 2.148. Let $(A, \leq)$ be a partially ordered set. A *chain* in A is an increasing sequence

$$x_0 \leq x_1 \leq x_2 \leq \cdots \tag{2.419}$$

of elements of A. A is *inductive* if every chain of elements in A has an upper bound.

Lemma 2.149. (Kuratowsky, Zorn) Every inductive partially ordered set $(A, \leq)$ has a maximal element.

Proof. Let $c_0 \in P$ be any element — note that at least one exists, because the empty chain has an upper bound by hypothesis. Next, for every ordinal α, let $c_\alpha \in A$ be any element such that for every $\beta < \alpha$ $c_\beta < c_\alpha$, if such an element exists; this uses the axiom of choice. If now α is a limit ordinal, then

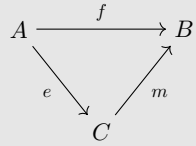
$$C_\alpha = \{c_\beta : \beta < \alpha\} \tag{2.420}$$

is a chain in A and since A is inductive there exists c_α. If α is any sufficiently large ordinal, however, there can be no $c_{\alpha+1} \in A$, which implies that $c_\alpha \in A$ is maximal. □

2.6 Factorization and isomorphism theorems

□ **Factorizations.** The purpose of factorization is to decompose an arbitrary function as a product of two functions of a more special nature. Examining nature and properties of the factor functions is usually simpler and still provides inportant information on the original function.

Definition 2.150. Let $f : A \to B$ be a function, as in the diagram below. A *factorization* of f is a pair (e, m) consisting of a pair of composable functions as in the diagram such that $f = em$.

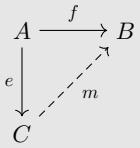


(2.421)

If e is surjective and m is injective we say that the pair (e, m) is an epi-mono factorization.

If, in diagram 2.421, f end e are given, we say tht f factors through e if we can find m such that $f = em$ and dually, if f and m are given we say that f factors through m if we can find an e such that $f = em$. Also the direct imega of the domain A along f is called the *image* of f and denoted $\mathrm{im}(f)$.
Although there are many types of factorization, here we are only interested in epi-mono factorizations. Our first goal is to prove the existence of at least one such factorization for every function. We use two lemmas.

LEMMA 2.151. (EPI FACTORIZATION) Assume, in the diagram below, that f is an arbitrary function and e is an epimorphism.



(2.422)

Then f factors through e if and only if $\ker(e) \subseteq \ker(f)$. In this case $f = em$ for a unique m; moreover

$$\ker(m) = e_*(\ker(f)), \qquad \mathrm{im}(m) = \mathrm{im}(f)\,. \tag{2.423}$$

Proof. ① Assume a factorization $f = em$ exists. Then

$$\ker(e) = ee^{\mathrm{op}} = e1e^{\mathrm{op}} \subseteq emm^{\mathrm{op}}e^{\mathrm{op}} = (em)\,(em)^{\mathrm{op}} = ff^{\mathrm{op}} = \ker(f) \tag{2.424}$$

where the first equality follows from the definition of kernel pair; the second from the fact that the identity is neutral for the product; the third inclusion from the fact that $1 \subseteq mm^{\mathrm{op}}$ because m is a function (corollary 2.50); the fourth equality from compatibility of the opposite with products; the fifth from the fact that (e, m) is a factorization of f; the last from the definition of kernel pair. ② Assume $\ker(e) \subseteq \ker(f)$. If a factorization $f = em$ exists then m is uniquely determined because

$$m = 1m = e^{\mathrm{op}}em = e^{\mathrm{op}}f \tag{2.425}$$

where the first equality follows from the fact that the identity is neutral for the product; the second from the fact that e is an epimorphism (proposition 2.61); the last from the fact that (e, m) is a factorization of f. It remains to prove that m is a function. First,

$$m^{\mathrm{op}}m = (e^{\mathrm{op}}f)^{\mathrm{op}}\,(e^{\mathrm{op}}f) = f^{\mathrm{op}}e^{\mathrm{op}\,\mathrm{op}}e^{\mathrm{op}}f = f^{\mathrm{op}}ee^{\mathrm{op}}f \subseteq f^{\mathrm{op}}ff^{\mathrm{op}}f \subseteq 1 \cdot 1 = 1 \tag{2.426}$$

where the first euality follows from the definition of m; the second from compatibility of the opposite with products; the third from the fact that the opposite is involutory; the fourth from the assumption that $\ker(e) \subseteq \ker(f)$; the fifth from the fact that f is a function; the last from the fact that the identity is neutral for the product. Moreover,

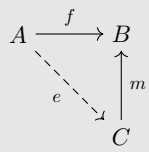
$$mm^{\mathrm{op}} = (e^{\mathrm{op}}f)\,(e^{\mathrm{op}}f)^{\mathrm{op}} = e^{\mathrm{op}}ff^{\mathrm{op}}e^{\mathrm{op}\,\mathrm{op}} = e^{\mathrm{op}}ff^{\mathrm{op}}e \supseteq e^{\mathrm{op}}1e = e^{\mathrm{op}}e = 1 \tag{2.427}$$

where the first equality follows from the definition of m; the second from compatibility of the opposite with products; the third from the fact that the opposite is involutory; the fourth from the fact that $1 \subseteq ff^{\mathrm{op}}$ because f is a function; the fifth from the fact that the identity is neutral for the product; the last from the fact that e is an epimorphism. Observe that formula (2.427) also proves that the kernel pair of m is the direct image along e of the kernel pair of f. As to $\mathrm{im}(m)$, observe that since e is surjective we have $e_*(A) = C$ and therefore

$$\mathrm{im}(m) = m_*(C) = m_*(e_*(A)) = (em)_*(A) = f_*(A) = \mathrm{im}(f)\,. \tag{2.428}$$

With a similar technique one proves the following, dual assertion.

LEMMA 2.152. (MONO FACTORIZATION) Assume, in the diagram below, that f is an arbitrary function and m a monomorphism.



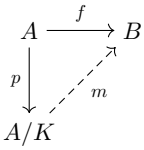
(2.429)

Then f factors through m if and only if $\mathrm{im}(f) \subseteq \mathrm{im}(m)$. In this case e is unique and

$$\ker(e) = \ker(f), \qquad \mathrm{im}(e) = m^*(\mathrm{im}(f)). \tag{2.430}$$

Proof. □

COROLLARY 2.153. Every function $f : A \to B$ admits an epi-mono factorization.



(2.431)

Proof. We refer to diagram 2.431. Let $K = ff^{\mathrm{op}}$ be the kernel pair of f and let p be the projection on the quotient set. p is surjective and its kernel pair is K (proposition 2.101), therefore there exists a unique m such that $f = em$ (lemma 2.151). Since $\ker(p) = K$, by the lemma we have

$$\ker(m) = p_*(K) = p^{\mathrm{op}}pp^{\mathrm{op}}p = 1 \cdot 1 = 1 \tag{2.432}$$

where the first equility follows from lemma 2.151; the second from the fact that $\ker(f) = \ker(p)$; the third from the fact that p is surjective; the last from the fact that the identity is neutral for the product. Thus, m is injective (proposition 2.61). □

The factorization provided by the corollary is known as the *canonical* epi-mono factorization of f.

Example 2.154 Let $A = \{0, 1, 2, 3\}$ and $B = \{0, 1, 2\}$. We compute an epi-mono factorization of the function $f : A \to B$ with incidence matrix

$$M(f) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{2.433}$$

If $E = \ker(f)$, the epimorphism is the projection on the quotient A/E and therefore

$$E = ff^{\mathrm{op}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} = pp^{\mathrm{op}}. \tag{2.434}$$

For the monomorphism we have

$$m = p^{\mathrm{op}}f = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{2.435}$$

Thus, the canonical epi-mono factorization of f is

$$f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = em. \tag{2.436}$$

We now address uniqueness of the factorization.

DEFINITION 2.155. Let $f : A \to B$ be any function. Given any two factorizations $em = f = e'm'$ of f as in the diagram below, a *morphism of factorizations* from first factorization $f = em$ to the second factorization $f = e'm'$ is a function h such that the diagram commutes, i.e. such that $eh = e'$ and $hm' = m$. If h is an isomorphism we say that it is an isomorphism of factorizations.

$$\begin{array}{ccc} A & \xrightarrow{e'} & D \\ {\scriptstyle e}\downarrow & \overset{h}{\nearrow} & \downarrow{\scriptstyle m'} \\ C & \xrightarrow[m]{} & B \end{array} \tag{2.437}$$

In particular an isomorphism of factorizations essentially menas that the two factorizations are the same, as the behaviour of e can be translated into the bahaviour of e' and conversely; and likewise for m and m'. We can now prove uniqueness of epi-mono factorizations.

PROPOSITION 2.156. Assume, in the diagram below, that $h = fg$.

$$\begin{array}{ccc} A & \xrightarrow{h} & C \\ & {\scriptstyle f}\searrow \quad \nearrow{\scriptstyle g} & \\ & B & \end{array} \tag{2.438}$$

1. If f and g are surjective, so is h. Conversely, if h is surjective, so is g.
2. If f and g are injective so is h. Conversely, if h is injective , so is f.

Proof. ① If f and g are surjective they have left inverses u and v respectively (proposition 2.64). Therefore,

$$(vu)\,h = (vu)\,(fg) = v(uf)\,g = v1g = vg = 1 \tag{2.439}$$

so that vu is a left inverse of h and h is surjective. Conversely if h is surjective it has a left inverse w. Then

$$(wf)\,g = w(fg) = wh = 1 \tag{2.440}$$

so that wf is a left inverse of g and g is surjective. ② Assume first that $A \neq \varnothing$. If f and g are injective and A is non-empty. Then B is nonempty and both f and g have right inverses u and v respectively. We have

$$h(vu) = (fg)\,(vu) = f(gv)\,u = f1u = fu = 1 \tag{2.441}$$

so that vu is a right inverse of h and h is injective. Conversely, if h is injective and w is a right inverse of h we have

$$f(gw) = (fg)\,w = hw = 1 \tag{2.442}$$

so that f has a right inverse and is injective. If $A = \varnothing$ then f and h are necessarily inkective and the claim is satisfied. □

PROPOSITION 2.157. Given any function $f : A \to B$, its epi-mono factorization is essentially unique, in the sense that given any two epi-mono factorizations $em = e'm'$ of f, they are isomorphic via a unique isomorphism h.

$$\begin{array}{ccc} A & \xrightarrow{e'} & D \\ {\scriptstyle e}\downarrow & \overset{h}{\nearrow} & \downarrow{\scriptstyle m'} \\ C & \xrightarrow[m]{} & B \end{array} \tag{2.443}$$

Proof. We refer to diagram 2.443. Since $f = em$ is an epi-mono factorization, we have

$$\ker(f) = ff^{\mathrm{op}} = (em)\,(em)^{\mathrm{op}} = emm^{\mathrm{op}}e^{\mathrm{op}} = e1e^{\mathrm{op}} = ee^{\mathrm{op}} = \ker(e) \tag{2.444}$$

where the first equality follows from the definition of kernel pair; the second from the factorization; the third from compatibility of the product with opposites; the fourth from the fact that e is an epimorphism; the fifth from the fact that the identity is neutral for the product; the last from the definition of kernel pair. Likewise, since $f = e'm'$ is another epi-mono factorization, $\ker(f) = \ker(e')$. Thus, $\ker(e) = \ker(e')$ and there exists a unique h such that $eh = e'$ (lemma 2.151). Moreover, sice e' is surjective, so is h (proposition 2.156). Next, $em = e'm' = ehm'$ and since e is an epimorphism and therefore left cancellable, $m = hm'$, so that h is a morphism of factorizations. Finally, since $hm' = m$ which is a monomorphism, h is a monomorphism (proposition 2.156). Thus, h is an isomorphism of factorizations. □

□ Isomorphism theorems.

THEOREM 2.158. (FIRST ISOMORPHISM THEOREM) If $f : A \to B$ is a function, then f induces an isomorphism

$$h : A/\ker(f) \to \mathrm{im}(f) \tag{2.445}$$

defined by the formula $h([x]) = f(x)$.

Proof. Consider the diagram below, in which p is the projection on the quotient set, defined by the formula $p(x) = [x]$, and i is the inclusion of the image of f in B defined by the formula $i(x) = x$.

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle p}\downarrow & \overset{g}{\nearrow} & \uparrow{\scriptstyle i} \\ A/\ker(f) & \xrightarrow[h]{} & \mathrm{im}(f) \end{array} \tag{2.446}$$

Observe that p is surjective and $\ker(p) = \ker(f)$; by the epi factorization lemma, there exists a unique function g such that $f = pg$. Moreover, $\ker(g) = e_*(\ker(f)) = 1$, so that g is a monomorphism and $\mathrm{im}(g) = \mathrm{im}(f)$. Since i is injective and $\mathrm{im}(i) = \mathrm{im}(f) = \mathrm{im}(g)$, the mono factorization lemma yields a unique function h such that $g = hi$. By the lemma, $\ker(h) = \ker(g) = 1$ so that h is a monomorphism and $\mathrm{im}(h) = i^*(\mathrm{im}(f)) = \mathrm{im}(f)$, so that h is an epimorphism. Thus, h is an isomorphism. Since $h = gi^{\mathrm{op}} = p^{\mathrm{op}}fi^{\mathrm{op}}$, we have $h([x]) = f(x)$. □

The relevance of the theorem is that instead of working on the quotient set A/E, one can work on the subset $\mathrm{im}(f)$, which has a asimpler structure, and then transfer the information back to the quotient set using h.

Assume now that A is a set and E is an equivalence relation on A. If $B \subseteq A$ is a subset, we let

$$BE = \bigcup_{x \in A} \{[x]_E : [x] \cap B \neq \varnothing\} \tag{2.447}$$

the set obtained by adjoining to B all the elements of the equivalence classes which intersect B. For any subset $S \subseteq A$, we also let $E|S := E \cap S^2$ be the *restriction* of E to S. With this notation, we have the following.

THEOREM 2.159. (SECOND ISOMORPHISM THEOREM) If A is a set, E is an equivalence relation on A and $B \subseteq A$ is a subset. Then $E|BE$ is an equivalence relation on BE and the function

$$h : B/(E|B) \to BE/(E|BE) \tag{2.448}$$

defined by the formula $h([x]) = [x]$ is an isomorphism.

THEOREM 2.160. (THIRD ISOMORPHISM THEOREM) Let A be a set and let $E \subseteq F$ be equivalence relations on A. Let $F/E := p_* F$ be the direct image of F along the projection $p : A \to A/E$. Then F/E is an equivalence relation on A/E and the function

$$h : (A/E)/(F/E) \to A/F \tag{2.449}$$

defined by the formula $h([[x]_E]_{F/E}) = [x]_F$ is an isomorphism.

2.7 Relations via functions

We have defined functions from relations. Conversely, one can define relations from functions. This point of view, which has some important advantages, will be examined not only for binary relations, but for relations of arbitrary arity.

□ **General products.** Binary relations can be generalized to relations of arity n for any $n \in \mathbf{N}$ and in fact even for infinite ordinal numbers. This is done by first generalizing the notion of cartesian product. Given sets a, b and c we could use the notion of pair to define a *triple* as

$$(a, b, c) = ((a, b), c) \tag{2.450}$$

using the notion of pair twice. One could then define the cartesian product of three sets as

$$A \times B \times C = \{(a, b, c) : (a \in A) \wedge (b \in B) \wedge (c \in C)\}. \tag{2.451}$$

Notice however, that one could make a different choice and define a triple as $(a, b, c) = (a, (b, c))$ and then would have a different product $A \times B \times C$. One can still prove that the two products we have defined are isomorphic, but it is quite clear that the notion and the notation would become more and more complex as arity increases. Insted, there is an alternative which applies uniformly to all cases. To motivate the definition, we will provide an equivalent description of the product $A \times B$. Let $2 = \{0, 1\}$ be a set with two elements; a function $s : 2 \to X$, where X is an arbitrary set, is called a *sequence* of length 2. It is customary to write s_0 and s_1 instead of $s(0)$ and $s(1)$. With this notation, let

$$C = \{s : 2 \to A \cup B : s_0 \in A \wedge s_1 \in B\} \tag{2.452}$$

be the set of sequences of length 2 in the union of A and B such that s_i belongs to the i-th factor of the product. C is not exactly the set $A \times B$, but there is an isomorphism $C \to A \times B$ which assigns to the sequence $s \in C$ the pair $(s_0, s_1) \in A \times B$. This suggests the following definition, in which $n = \{0, 1, \dots, n-1\}$ is the set of the first n natural numbers.

Definition 2.161. Given $n \in \mathbf{N}$ and n sets $A_0, A_1, \dots, A_{n-1}$, their *product* is the set

$$\prod_{i=0}^{n-1} A_i = \left\{ s : n \to \bigcup_{i=0}^{n-1} A_i : s_i \in A_i \right\} \tag{2.453}$$

of all sequences of length n into the union of the A_i's such that the i-th coordinate of the sequence s_i belongs to A_i.

Although we have given the definition in the case of a finite number n of factors, there is no difference whatsoever in the infinite case when we replace n by an arbitrary, possible infinite ordinal number — see definition 2.146 and the subsection it belongs to. Note that when $n = 0$, n is the empty set and there is exactly one function from 0 to any set; thus the product of zero sets is a set with 1 element.
In particular if $A_i = A$ is the same set for every index i, we have the n-power A^n, whose elements are the n-sequences of elements of A,

$$s = (s_0, s_1, \dots, s_{n-1}) . \tag{2.454}$$

When $n = 0$ we have $A^0 = \{\varnothing\} = \{0\} = 1$.

□ **Relations as functions.** Although we have defined functions starting from binary relations, one could go the other way around and define binary relations starting from functions. The key concept in this approach is that of characteristic function.

Definition 2.162. If A is a set and $B \subseteq A$ a subset, the *characteristic function* of B is the function $B_\Omega : A \to \Omega$ defined by the formula

$$B_\Omega(x) = \begin{cases} 1 & \text{if } x \in B \\ 0 & \text{if } x \notin B. \end{cases} \tag{2.455}$$

Immediately from the definition we have

Proposition 2.163. Given any set A, there is a natural bijection

$$\mathrm{Sub}(A) \to \Omega^A \tag{2.456}$$

which assigns to every subset $B \subseteq A$ its characteristi function. The inverse bijection assigns to every function $A \to \Omega$ the inverse image of $1 \in \Omega$.

Not only can we replace subsets with function. The replacement is compatible with inclusion. To see what this means, recall that Ω is totally ordered by $0 < 1$. Therefore, we have a partially ordered set $(\Omega^A, \leq)$ for the pointwise order relation (example 2.121).

Proposition 2.164. Under the isomorphism (2.456), inclusion on $\mathrm{Sub}(A)$ translates into inequality in Ω^A, i.e.,

$$U \subseteq V \Leftrightarrow U_\Omega \leq V_\Omega \tag{2.457}$$

Proof. It suffices to observe that

$$U \subseteq V \Leftrightarrow \forall a(a \in U \to a \in V) \tag{2.458}$$
$$\Leftrightarrow \forall a(U_\Omega(a) = 1 \to V_\Omega(a) = 1) \tag{2.459}$$
$$\Leftrightarrow \forall a(U_\Omega(a) = V_\Omega(a)) \tag{2.460}$$
$$\Leftrightarrow U_\Omega = V_\Omega \tag{2.461}$$

where the first equivalence follows from the definition of inclusion; the second from the fact that the characteristic functions U_Ω and V_Ω are two-valued; the last from the extensionality axiom. □

The next step is to prove that the Boolean algebra structure of $\mathrm{Sub}(A)$ can be replaced by one on Ω^A. To see how this can be done, observe that truth functions make Ω into a Boolean algebra. This structure transfers to Ω^A, which becomes a Boolean algebra with pointwise operations

$$(u \wedge v)(a) = u(a) \wedge v(a), \qquad (u \vee v)(a) = u(a) \vee v(a) \qquad (\neg u)(a) = \neg(u(a)), \tag{2.462}$$
$$\top(a) = \top, \qquad \bot(a) = \bot. \tag{2.463}$$

This is how we replace the boolean structure.

Proposition 2.165. Under the isomorphism (2.456), the Boolean structure on $\mathrm{Sub}(A)$ translates into the Boolean structure of Ω^A, i.e.,

$$U = \varnothing \Leftrightarrow U_\Omega = \bot \qquad U = A \Leftrightarrow U_\Omega = \top \tag{2.464}$$
$$U' = W \Leftrightarrow \neg U_\Omega = V_\Omega \tag{2.465}$$
$$U \cap V = W \Leftrightarrow U_\Omega \wedge V_\Omega = W_\Omega \qquad U \cup V = W \Leftrightarrow U_\Omega \vee V_\Omega = W_\Omega \tag{2.466}$$

Proof. We prove the claim for intersection. We have

$$U \cap V = W \Leftrightarrow \forall a(a \in U \cap V \leftrightarrow a \in W) \tag{2.467}$$
$$\Leftrightarrow \forall a((a \in U) \wedge (a \in V) \leftrightarrow a \in W) \tag{2.468}$$
$$\Leftrightarrow \forall a((U_\Omega(a) = 1) \wedge (V_\Omega(a) = 1) \leftrightarrow W_\Omega(a) = 1) \tag{2.469}$$
$$\Leftrightarrow \forall a((U_\Omega \wedge V_\Omega)(a) = 1) \leftrightarrow W_\Omega(a) = 1) \tag{2.470}$$
$$\Leftrightarrow \forall a((U_\Omega \wedge V_\Omega)(a) = W_\Omega(a)) \tag{2.471}$$
$$\Leftrightarrow U_\Omega \wedge V_\Omega = W \tag{2.472}$$

where the first equivalence follows from extensionality; the second from the definition of intersection; the third from the definition of characteristic function; the fourth from the definition of the $\wedge$-operation in Ω^A; the fifth from the fact that characteristic functions are only two-valued; the last from extensionality. □

The result of all this is that instead of working with subsets of A we can work with functions in Ω^A. Since binary relations $R : A \rightharpoonup B$ are precisely the subsets of $A \times B$, it follows that all the local structure on $\mathbf{Rel}(A, B)$ can be replaced by the structure on $\Omega^{A \times B}$. The situation for the multiplicative structure requires additional work.

□ **Relations of arity n.** What we have said so far about binary relations can be generalized to relations of arbitrary arity.

Definition 2.166. If $n \in \mathbf{N}$ is a natural number and $(A_1, \dots, A_n)$ is a sequence of n sets, not necessarily distinct, a *relation* of arity n on the A_i's is a subset

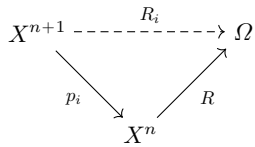
$$R \subseteq \prod_{i=1}^{n} A_i. \tag{2.473}$$

This definition is mostly used when all the elementas of the sequence are equal for a fixed set A, so that $R \subseteq A^n$. In this case we say that R is a relation of arity n on A. In particular, there are only two relations of arity 0 on A, namely 0 and 1 which do not depend on the choice of A and which we can identify with the elements of Ω. A relation of arity 1 on A is simply a subset of A and a relation of arity 2 is exactly what we have called a binary relation on A. Note that talking about domain and codomain for $n \neq 2$ is not meaningful, as no factor of the product of n copies of A has a particular meaning.
By proposition 2.163 we can identify a relation $R \subseteq A^n$ with a function $A^n \to \Omega$. This point of view has the additional benefit that we can extend the definition to the case $n = 0$: for $A^0 = 1$, the set with a single element, and a function $1 \to \Omega$ is simply a truth value; thus relaions of arity 0 are truth vaues.
There is a similar construction for functions, except that one needs to keep track of the codomain, in this case. Thus, a function of n arguments is simply a function $f : \prod_{i=1}^{n} A_i \to B$. Nothe that if we let $A = \prod_{i=1}^{n} A_i$, this is simply a function $f : A \to B$. To emphasize the dependence on n arguments, we write $f(a_1, \dots, a_n) = b$ instead of $f(a) = b$ if $a = (a_1, \dots, a_n)$.
A function $f : A^n \to A$ is called an n-ary operation on A.

□ **Change of base along projections.** Given $n \in \mathbf{N}$, let $p_i : X^{n+1} \to X^n$ be the projection that misses the i-th coordinate:

$$p_i(x_1, \dots, x_{n+1}) = (x_1, \dots, \hat{x}_i, \dots, x_{n+1}) \tag{2.474}$$

where the symbol $\hat{x}_i$ indicates that x_i has beeen omitted. If R is a relation of arity n on X, as shown in the diagram below, we write $R_i := p_i R$ for the relation of arity $n+1$ on X obtained by omitting the i-th coordinate.



(2.475)

Thus, composing with a projection increases the arity of a relation. There are two important constructions in the opposite direction:

Definition 2.167. Given a relation R of arity $n+1$ on X, define relations of arity n on X as follows

$$\begin{aligned} \forall_i R(x_1, \dots, \hat{x}_i, \dots, x_{n+1}) &:= \inf\{R(x_1, \dots, x_{n+1}) : x_i \in X\} \\ \exists_i R(x_1, \dots, \hat{x}_i, \dots, x_{n+1}) &:= \sup\{R(x_1, \dots, x_{n+1}) : x_i \in X\} \end{aligned} \tag{2.476}$$

These two constructions are characterized by the following proprty.

Theorem 2.168. (Change of base adjoints) Assume S is an n-ary relations on X and R, T are relations of arity $n+1$ on X. Then

$$\exists_i R \leq S \Leftrightarrow R \leq S_i \qquad\qquad S_i \leq T \Leftrightarrow S \leq \forall_i T \tag{2.477}$$

Proof. In the proof wewrite x_i to indicate an element in the i-th coordinate and $x \in X^n$ for an n-tuple of elements in the remaining coordinates. ① For the first inequality it suffices to observe that

$$\begin{aligned} \exists_i R \leq S &\Leftrightarrow \forall x[\sup_{x_i} R(x, x_i) \leq S(x)] && (2.478)\\ &\Leftrightarrow \forall x, x_i[R(x, x_i)) \leq S(x)] && (2.479)\\ &\Leftrightarrow \forall x, x_i[R(x, x_i) \leq S_i(x, x_i)] && (2.480)\\ &\Leftrightarrow R \leq S_i \end{aligned}$$

where the first equivalence follows from the definition of $\exists_i R$; the second from the definition of least upper bound; the third from the fact that $S(x) = S_i(x, x_i)$ by the definition of S_i; the last from the definition of order on relations. ② In this case we have

$$\begin{aligned} S \leq \forall_i T &\Leftrightarrow \forall x[S(x) \leq \inf_{x_i} T(x, x_i)] && (2.481)\\ &\Leftrightarrow \forall x, x_i[S(x) \leq T(x, x_i)] && (2.482)\\ &\Leftrightarrow \forall x, x_i[S_i(x, x_i) \leq T(x, x_i)] && (2.483) \end{aligned}$$

where the first equivalence follows from the definition of $\forall_i T$; the second from the definition of greatest lower bound; the third from the fact that $S(x) = S_i(x, x_i)$ by the definition of S_i; the last from the definition of order on relations. □

An application of quantification of relations is the description of the product of relations as a function. Suppose we have relation $R : A \rightharpoonup B$ and $S : B \rightharpoonup C$ with corresponding functions $R_\Omega : A \times B \to \Omega$ and $S_\Omega : B \times C \to \Omega$. To describe the function associated to RS, consider first the relation $R \times_B S$ whose associated function is defined by the following diagram.

$$\begin{array}{ccc} A \times B \times C & \overset{(R \triangle S)_\Omega}{\dashrightarrow} & \Omega \\ {\scriptstyle 1 \times \Delta \times 1}\downarrow & & \uparrow{\scriptstyle \wedge} \\ A \times B \times B \times C & \xrightarrow[R_\Omega \times S_\Omega]{} & \Omega \times \Omega \end{array} \tag{2.484}$$

Then

$$(RS)_\Omega = \exists b((R \triangle S)_\Omega) \tag{2.485}$$

3 First order logic

Not available yet

4 Algebraic theories

Not available yet

Bibliography

1 R.M. Smullyan, *First-order logic*, 1995.