1 Generative Effects: Orders and Galois Connections

In this book, we explore a wide variety of situations – in the world of science, engineering, and commerce – where we see something we might call *compositionality*. These are cases in which systems or relationships can be combined to form new systems or relationships. In each case we find category-theoretic constructs – developed for their use in pure math – which beautifully describe the compositionality of the situation.

This chapter, being the first of the book, must serve this goal in two capacities. First, it must provide motivating examples of compositionality, as well as the relevant categorical formulations. Second, it must provide the mathematical foundation for the rest of the book. Since we are starting with minimal assumptions about the reader's background, we must begin slowly and build up throughout the book. As a result, examples in the early chapters are necessarily simplified. However, we hope the reader will already begin to see the sort of structural approach to modeling that category theory brings to the fore.

1.1 More Than the Sum of Their Parts

We motivate this first chapter by noticing that while many real-world structures are compositional, the results of observing them are often not. The reason is that observation is inherently "lossy": in order to extract information from something, one must drop the details. For example, one stores a real number by rounding it to some precision. But if the details are actually relevant in a given system operation, then the observed result of that operation will not be as expected. This is clear in the case of roundoff error, but it also shows up in non-numerical domains: observing a complex system is rarely enough to predict its behavior because the observation is lossy.

A central theme in category theory is the study of structures and structure-preserving maps. A map $f\colon X \to Y$ is a kind of observation of object X via a specified relationship it has with another object, Y. For example, think of X as the subject of an experiment and Y as a meter connected to X, which allows us to extract certain features of X by looking at the reaction of Y.

Asking which aspects of X one wants to preserve under the observation f becomes the question "what category are you working in?" As an example, there are many functions f from $\mathbb{R}$ to $\mathbb{R}$ (where $\mathbb{R}$ is the set of real numbers), and we can think of them as observations: rather than view x "directly," we only observe $f(x)$. Out of all the functions $f: \mathbb{R} \to \mathbb{R}$, only some of them preserve the order of numbers, only some of them preserve the distance between numbers, only some of them preserve the sum of numbers, etc. Let's check in with an exercise; a solution can be found in the Appendix.

Exercise 1.1. Some terminology: a function $f: \mathbb{R} \to \mathbb{R}$ is said to be

(a) *order-preserving* if $x \leq y$ implies $f(x) \leq f(y)$, for all $x, y \in \mathbb{R}$;[1]
(b) *metric-preserving* if $|x - y| = |f(x) - f(y)|$;
(c) *addition-preserving* if $f(x + y) = f(x) + f(y)$.

For each of the three properties defined above – call it *foo* – find an f that is *foo*-preserving and an example of an f that is not *foo*-preserving. ◇

In category theory we want to keep control over which aspects of our systems are being preserved under various observations. As we said above, the less structure is preserved by our observation of a system, the more "surprises" occur when we observe its operations. One might call these surprises *generative effects*.

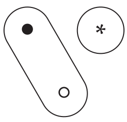
In using category theory to explore generative effects, we follow the basic ideas from work by Adam [Ada17]. He goes much more deeply into the issue than we can here; see Section 1.5. But as mentioned above, we must also use this chapter to give an order-theoretic warm-up for the full-fledged category theory to come.

1.1.1 A First Look at Generative Effects

To explore the notion of a generative effect we need a sort of system, a sort of observation, and a system-level operation that is not preserved by the observation. Let's start with a simple example.

A simple system

Consider three points; we'll call them $\bullet$, $\circ$, and $*$. In this example, a *system* will simply be a way of connecting these points together. We might think of our points as sites on a power grid, with a system describing connection by power lines, or as people susceptible to some disease, with a system describing interactions that can lead to contagion. As an abstract example of a system, there is a system where $\bullet$ and $\circ$ are connected, but neither is connected to $*$. We shall draw this like so:



[1] We are often taught to view functions $f: \mathbb{R} \to \mathbb{R}$ as plots in the (x, y)-coordinate system, where x is the domain (independent) variable and y is the codomain (dependent) variable. In this book, we do not adhere to that naming convention; e.g. in Example 1.1, both x and y are being "plugged in" as input to f. As an example consider the function $f(x) = x^2$. Then f being order-preserving would say that, for any $x, y \in \mathbb{R}$, if $x \leq y$ then $x^2 \leq y^2$; is that true?

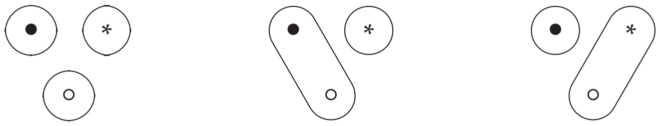
Connections are symmetric, so if a is connected to b, then b is connected to a. Connections are also transitive, meaning that if a is connected to b, and b is connected to c, then a is connected to c; that is, all a, b, and c are connected. Friendship is not transitive – my friend's friend is not necessarily my friend – but possible communication of a concept or a disease is.

Here we depict two more systems, one in which none of the points are connected, and one in which all three points are connected.

There are five systems in all, and we depict them below.

Now that we have defined the sort of system we want to discuss, suppose that Alice is observing this system. Her observation of interest, which we call Φ, extracts a single feature from a system, namely whether the point $\bullet$ is connected to the point $*$; this is what she wants to know. Her observation of the system will be an assignment of either `true` or `false`; she assigns `true` if $\bullet$ is connected to $*$, and `false` otherwise. So Φ assigns the value `true` to the following two systems:

and Φ assigns the value `false` to the three remaining systems:

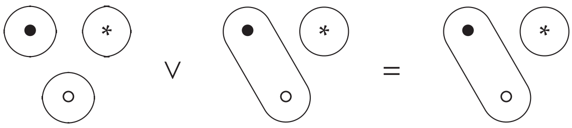


(1.1)

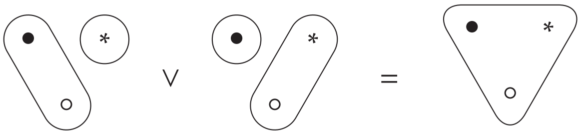
The last piece of setup is to give a sort of operation that Alice wants to perform on the systems themselves. It's a very common operation – one that will come up many times throughout the book – called *join*. If the reader has been following the story arc, the expectation here is that Alice's connectivity observation will not be compositional with respect to the operation of system joining; that is, there will be generative effects. Let's see what this means.

Joining our simple systems

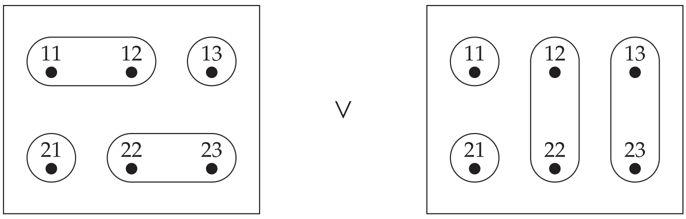
Joining two systems A and B is performed simply by combining their connections. That is, we shall say the *join* of systems A and B, denoted $A \vee B$, has a connection between points x and y if there are some points $z_1, \ldots, z_n$ such that each of the following is true in at least one of A or B: x is connected to z_1, z_i is connected to z_{i+1}, and z_n is connected to y. In a three-point system, the above definition is overkill, but we want to say something that works for systems with any number of elements. The high-level way to say it is "take the transitive closure of the union of the connections in A and B." In our three-element system, it means for example that



and (1.2)



Exercise 1.2. What is the result of joining the following two systems?



◇

We are now ready to see the generative effect. We don't want to build it up too much – this example has been made as simple as possible – but we shall see that Alice's observation fails to preserve the join operation. We've been denoting her observation – measuring whether • and ∗ are connected – by the symbol Φ; it returns a boolean result, either `true` or `false`.

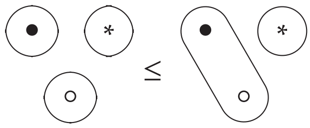
We see above in Eq. (1.1) that Φ() = Φ() = `false`: in both cases • is not connected to ∗. On the other hand, when we join these two systems as in Eq. (1.2), we see that Φ( ∨ ) = Φ() = `true`: in the joined system, • *is* connected to ∗. The question that Alice is interested in, that of Φ, is inherently lossy with respect to join, and there is no way to fix it without a more detailed observation, one that includes not only ∗ and • but also ∘.

While this was a simple example, it should be noted that whether the potential for such effects exist – i.e. determining whether an observation is operation-preserving – can be incredibly important information to know. For example, Alice could be in charge of putting together the views of two local authorities regarding possible contagion between an infected person • and a vulnerable person ∗. Alice has noticed that if they separately extract information from their raw data and combine the results, it gives a different answer than if they combine their raw data and extract information from it.

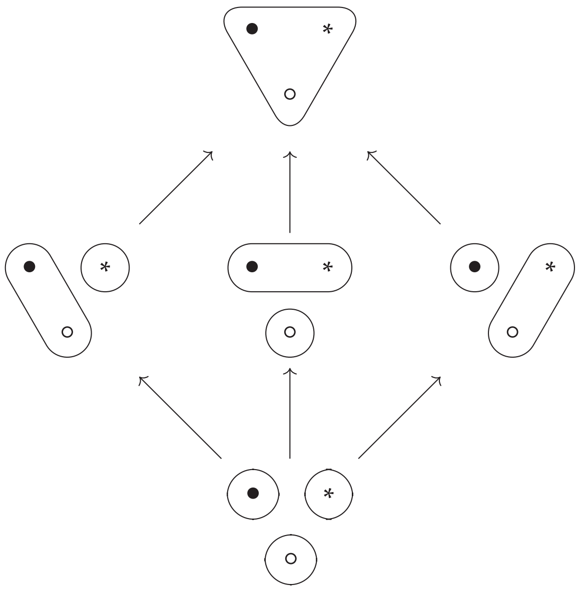
1.1.2 Ordering Systems

Category theory is all about organizing and layering structures. In this section we will explain how the operation of joining systems can be derived from a more basic structure: order. We shall see that while joining is not preserved by Alice's connectivity observation Φ, order is.

To begin, we note that the systems themselves are ordered in a hierarchy. Given systems A and B, we say that $A \leq B$ if, whenever x is connected to y in A, then x is connected to y in B. For example,



This notion of $\leq$ leads to the following diagram:



(1.3)

where an arrow from system A to system B means $A \leq B$. Such diagrams are known as *Hasse diagrams*.

As we were saying above, the notion of join is derived from this order. Indeed, for any two systems A and B in the Hasse diagram (1.3), the joined system $A \vee B$ is the smallest system that is bigger than both A and B. That is, $A \leq (A \vee B)$ and $B \leq (A \vee B)$, and for any C, if $A \leq C$ and $B \leq C$ then $(A \vee B) \leq C$. Let's walk through this with an exercise.

Exercise 1.3.

1. Write down all the partitions of a two-element set $\{\bullet, *\}$, order them as above, and draw the Hasse diagram.

2. Now do the same thing for a four-element set, say $\{1, 2, 3, 4\}$. There should be 15 partitions.

Choose any two systems in your 15-element Hasse diagram, call them A and B.

3. What is $A \vee B$, using the definition given in the paragraph above Eq. (1.2)?
4. Is it true that $A \leq (A \vee B)$ and $B \leq (A \vee B)$?
5. What are all the systems C for which both $A \leq C$ and $B \leq C$?
6. Is it true that in each case $(A \vee B) \leq C$? ◇

The set $\mathbb{B} = \{\texttt{true}, \texttt{false}\}$ of booleans also has an order, $\texttt{false} \leq \texttt{true}$:

$$\begin{array}{c}\texttt{true}\\ \uparrow \\ \texttt{false}\end{array}$$

Thus $\texttt{false} \leq \texttt{false}$, $\texttt{false} \leq \texttt{true}$, and $\texttt{true} \leq \texttt{true}$, but $\texttt{true} \not\leq \texttt{false}$. In other words, $A \leq B$ if A implies B.[2]

For any A, B in $\mathbb{B}$, we can again write $A \vee B$ to mean the least element that is greater than both A and B.

Exercise 1.4. Using the order $\texttt{false} \leq \texttt{true}$ on $\mathbb{B} = \{\texttt{true}, \texttt{false}\}$, what is:

1. $\texttt{true} \vee \texttt{false}$?
2. $\texttt{false} \vee \texttt{true}$?
3. $\texttt{true} \vee \texttt{true}$?
4. $\texttt{false} \vee \texttt{false}$? ◇

Let's return to our systems with $\bullet$, $\circ$, and $*$, and Alice's "$\bullet$ is connected to $*$" function, which we called Φ. It takes any such system and returns either $\texttt{true}$ or $\texttt{false}$. Note that the map Φ preserves the $\leq$ order: if $A \leq B$ and there is a connection between $\bullet$ and $*$ in A, then there is such a connection in B too. The possibility of a generative effect is captured in the inequality

$$\Phi(A) \vee \Phi(B) \leq \Phi(A \vee B). \tag{1.4}$$

We saw on page 4 that this can be a strict inequality: we showed two systems A and B with $\Phi(A) = \Phi(B) = \texttt{false}$, so $\Phi(A) \vee \Phi(B) = \texttt{false}$, but where $\Phi(A \vee B) = \texttt{true}$. In this case, a generative effect exists.

These ideas capture the most basic ideas in category theory. Most directly, we have seen that the map Φ preserves some structure but not others: it preserves order but not join. In fact, we have seen here hints of more complex notions from category theory, without making them explicit; these include the notions of category, functor, colimit, and adjunction. In this chapter we will explore these ideas in the elementary setting of ordered sets.

[2] In mathematical logic, $\texttt{false}$ implies $\texttt{true}$ but $\texttt{true}$ does not imply $\texttt{false}$. That is "P implies Q" means, "if P is true, then Q is true too, but if P is not true, I'm making no claims."

Cambridge University Press
978-1-108-48229-5 — An Invitation to Applied Category Theory
Brendan Fong , David I. Spivak
Excerpt
More Information

1.2 What is Order?

Above we informally spoke of two different ordered sets: the order on system connectivity and the order on booleans `false` $\leq$ `true`. Then we related these two ordered sets by means of Alice's observation Φ. Before continuing, we need to make such ideas more precise. We begin in Section 1.2.1 with a review of sets and relations. In Section 1.2.2 we will give the definition of a *preorder* – short for preordered set – and a good number of examples.

1.2.1 Review of Sets, Relations, and Functions

We will not give a definition of *set* here, but informally we will think of a set as a collection of things, known as elements. These things could be all the leaves on a certain tree, or the names of your favorite fruits, or simply some symbols a, b, c. For example, we write $A = \{h, 1\}$ to denote the set, called A, that contains exactly two elements, one called h and one called 1. The set $\{h, h, 1, h, 1\}$ is exactly the same as A because they both contain the same elements, h and 1, and repeating an element more than once in the notation doesn't change the set.[3] For an arbitrary set X, we write $x \in X$ if x is an element of X; so we have $h \in A$ and $1 \in A$, but $0 \notin A$.

Example 1.5. Here are some important sets from mathematics – and the notation we will use – that will appear again in this book.

- $\varnothing$ denotes the empty set; it has no elements.
- $\{1\}$ denotes a set with one element; it has one element, 1.
- $\mathbb{B}$ denotes the set of *booleans*; it has two elements, `true` and `false`.
- $\mathbb{N}$ denotes the set of *natural numbers*; it has elements $0, 1, 2, 3, \ldots, 90^{717}, \ldots$.
- $\underline{n}$, for any $n \in \mathbb{N}$, denotes the nth *ordinal*; it has n elements $1, 2, \ldots, n$. For example, $\underline{0} = \varnothing$, $\underline{1} = \{1\}$, and $\underline{5} = \{1, 2, 3, 4, 5\}$.
- $\mathbb{Z}$, the set of *integers*; it has elements $\ldots, -2, -1, 0, 1, 2, \ldots, 90^{717}, \ldots$.
- $\mathbb{R}$, the set of *real numbers*; it has elements like $\pi, 3.14, 5 * \sqrt{2}, e, e^2, -1457, 90^{717}$, etc.

Given sets X and Y, we say that X is a *subset* of Y, and write $X \subseteq Y$, if every element in X is also in Y. For example $\{h\} \subseteq A$. Note that the empty set $\varnothing := \{\}$ is a subset of every other set.[4] Given a set Y and a property P that is either true or false for each element of Y, we write $\{y \in Y \mid P(y)\}$ to mean the subset of those y's that satisfy P.

Exercise 1.6.

1. Is it true that $\mathbb{N} = \{n \in \mathbb{Z} \mid n \geq 0\}$?

[3] If you want a notion where "h, 1" is different from "$h, h, 1, h, 1$," you can use something called *bags*, where the number of times an element is listed matters, or *lists*, where order also matters. All of these are important concepts in applied category theory, but sets will come up the most for us.

[4] When we write $Z :=$ `foo`, it means "assign the meaning `foo` to variable Z," whereas $Z =$ `foo` means simply that Z is equal to `foo`, perhaps as discovered via some calculation. In particular, $Z :=$ `foo` implies $Z =$ `foo` but not vice versa; indeed it *would not* be proper to write $3 + 2 := 5$ or $\{\} := \varnothing$.

2. Is it true that $\mathbb{N} = \{n \in \mathbb{Z} \mid n \geq 1\}$?
3. Is it true that $\varnothing = \{n \in \mathbb{Z} \mid 1 < n < 2\}$? ◇

If both X_1 and X_2 are subsets of Y, their *union*, denoted $X_1 \cup X_2$, is also a subset of Y, namely the one containing the elements in X_1 and the elements in X_2 but no more. For example if $Y = \{1, 2, 3, 4\}$ and $X_1 = \{1, 2\}$ and $X_2 = \{2, 4\}$, then $X_1 \cup X_2 = \{1, 2, 4\}$. Note that $\varnothing \cup X = X$ for any $X \subseteq Y$.

Similarly, if both X_1 and X_2 are subsets of Y, then their *intersection*, denoted $X_1 \cap X_2$, is also a subset of Y, namely the one containing all the elements of Y that are both in X_1 and in X_2, and no others. So $\{1, 2, 3\} \cap \{2, 5\} = \{2\}$.

What if we need to union together or intersect a lot of subsets? For example, consider the sets $X_0 = \varnothing$, $X_1 = \{1\}$, $X_2 = \{1, 2\}$, etc. as subsets of $\mathbb{N}$, and we want to know what the union of all of them is. This union is written $\bigcup_{n \in \mathbb{N}} X_n$, and it is the subset of $\mathbb{N}$ that contains every element of every X_n, but no others. Namely, $\bigcup_{n \in \mathbb{N}} X_n = \{n \in \mathbb{N} \mid n \geq 1\}$. Similarly one can write $\bigcap_{n \in \mathbb{N}} X_n$ for the intersection of all of them, which will be empty in the above case.

Given two sets X and Y, the *product* $X \times Y$ of X and Y is the set of pairs (x, y), where $x \in X$ and $y \in Y$.

Finally, we may want to take a *disjoint* union of two sets, even if they have elements in common. Given two sets X and Y, their *disjoint union* $X \sqcup Y$ is the set of pairs of the form $(x, 1)$ or $(y, 2)$, where $x \in X$ and $y \in Y$.

Exercise 1.7. Let $A := \{h, 1\}$ and $B := \{1, 2, 3\}$.

1. There are eight subsets of B; write them out.
2. Take any two nonempty subsets of B and write out their union.
3. There are six elements in $A \times B$; write them out.
4. There are five elements of $A \sqcup B$; write them out.
5. If we consider A and B as subsets of the set $\{h, 1, 2, 3\}$, there are four elements of $A \cup B$; write them out. ◇

Relationships between different sets – for example between the set of trees in your neighborhood and the set of your favorite fruits – are captured using subsets and product sets.

Definition 1.8. Let X and Y be sets. A *relation between X and Y* is a subset $R \subseteq X \times Y$. A *binary relation on X* is a relation between X and X, i.e. a subset $R \subseteq X \times X$.

It is convenient to use something called *infix notation* for binary relations $R \subseteq A \times A$. This means one picks a symbol, say $\star$, and writes $a \star b$ to mean $(a, b) \in R$.

Example 1.9. There is a binary relation on $\mathbb{R}$ with infix notation $\leq$. Rather than writing $(5, 6) \in R$, we write $5 \leq 6$.

Other examples of infix notation for relations are $=, \approx, <, >$. In number theory, we are interested in whether one number divides without remainder into another number; this relation is denoted with infix notation $|$, so $5|10$.

Partitions and equivalence relations

We can now define partitions more formally.

Definition 1.10. If A is a set, a *partition* of A consists of a set P and, for each $p \in P$, a nonempty subset $A_p \subseteq A$, such that

$$A = \bigcup_{p \in P} A_p \qquad \text{and} \qquad \text{if } p \neq q \text{ then } A_p \cap A_q = \varnothing. \tag{1.5}$$

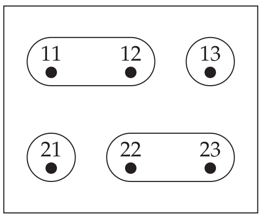
We may denote the partition by $\{A_p\}_{p \in P}$. We refer to P as the set of *part labels* and if $p \in P$ is a part label, we refer to A_p as the pth *part*. The condition (1.5) says that each element $a \in A$ is in exactly one part.

We consider two different partitions $\{A_p\}_{p \in P}$ and $\{A'_{p'}\}_{p' \in P'}$ of A to be the same if for each $p \in P$ there exists a $p' \in P'$ with $A_p = A'_{p'}$. In other words, if two ways to divide A into parts are exactly the same – the only change is in the labels – then we don't make a distinction between them.

Exercise 1.11. Suppose that A is a set and $\{A_p\}_{p \in P}$ and $\{A'_{p'}\}_{p' \in P'}$ are two partitions of A such that for each $p \in P$ there exists a $p' \in P'$ with $A_p = A'_{p'}$.

1. Show that for each $p \in P$ there is at most one $p' \in P'$ such that $A_p = A'_{p'}$.
2. Show that for each $p' \in P'$ there is a $p \in P$ such that $A_p = A'_{p'}$. ◇

Exercise 1.12. Consider the partition shown below:



For any two elements $a, b \in \{11, 12, 13, 21, 22, 23\}$, let's allow ourselves to write a twiddle (tilde) symbol $a \sim b$ between them if a and b are both in the same part. Write down every pair of elements (a, b) that are in the same part. There should be 10.[5] ◇

We shall see in Proposition 1.14 that there is a strong relationship between partitions and something called equivalence relations, which we define next.

[5] Hint: whenever someone speaks of "two elements a, b in a set A," the two elements may be the same!

Definition 1.13. Let A be a set. An *equivalence relation* on A is a binary relation, let's give it infix notation $\sim$, satisfying the following three properties:

(a) $a \sim a$, for all $a \in A$,
(b) $a \sim b$ iff[6] $b \sim a$, for all $a, b \in A$,
(c) if $a \sim b$ and $b \sim c$ then $a \sim c$, for all $a, b, c \in A$.

These properties are called *reflexivity*, *symmetry*, and *transitivity*, respectively.

Proposition 1.14. Let A be a set. There is a one-to-one correspondence between the ways to partition A and the equivalence relations on A.

Proof. We first show that every partition gives rise to an equivalence relation, and then that every equivalence relation gives rise to a partition. Our two constructions will be mutually inverse, proving the proposition.

Suppose we are given a partition $\{A_p\}_{p \in P}$; we define a relation $\sim$ and show it is an equivalence relation. Define $a \sim b$ to mean that a and b are in the same part: there is some $p \in P$ such that $a \in A_p$ and $b \in A_p$. It is obvious that a is in the same part as itself. Similarly, it is obvious that if a is in the same part as b then b is in the same part as a, and that if further b is in the same part as c then a is in the same part as c. Thus $\sim$ is an equivalence relation as defined in Definition 1.13.

Suppose we are given an equivalence relation $\sim$; we will form a partition on A by saying what the parts are. Say that a subset $X \subseteq A$ is ($\sim$)-closed if, for every $x \in X$ and $x' \sim x$, we have $x' \in X$. Say that a subset $X \subseteq A$ is ($\sim$)-connected if it is nonempty and $x \sim y$ for every $x, y \in X$. Then the parts corresponding to $\sim$ are exactly the ($\sim$)-closed, ($\sim$)-connected subsets. It is not hard to check that these indeed form a partition. □

Exercise 1.15. Let's complete the "it's not hard to check" part in the proof of Proposition 1.14. Suppose that $\sim$ is an equivalence relation on a set A, and let P be the set of ($\sim$)-closed and ($\sim$)-connected subsets $\{A_p\}_{p \in P}$.

1. Show that each part A_p is nonempty.
2. Show that if $p \neq q$, i.e. if A_p and A_q are not exactly the same set, then $A_p \cap A_q = \varnothing$.
3. Show that $A = \bigcup_{p \in P} A_p$. ◇

Definition 1.16. Given a set A and an equivalence relation $\sim$ on A, we say that the *quotient* $A/\sim$ of A under $\sim$ is the set of parts of the corresponding partition.

Functions

The most frequently used sort of relation between sets is that of functions.

[6] "Iff" is short for "if and only if."