Preliminaries: Sets & Functions

Jon Ander Mendia mendia@cornell.com
Problems in Semantics – Cornell University
September 3, 2019

We start off by reviewing some basic concepts of set theory, relations and functions. These notes are based on §1 and §2 of Partee et al. (1990) and Cable (2014). Another good source is §2 in Causey (1994). Both Partee et al. (1990) and Causey (1994) contain some review exercises with solutions at the end.

1 Basic Set Theory

1.1 Set specification

A set is a collection of things, concrete or abstract. To represent a set we can list its members by writing them in between curly brackets: $\{a,b,c\}$ denotes the collection of three things, the letters a, b and c. The order in which we write the objects that belong to a set does not affect its constituency, nor does writing the name of an object more than once. The following are all the same set: $\{a,b,c\} = \{c,b,a\} = \{a,a,a,a,b,b,c,c,c,c\}$.

The basic concept of set theory is the **membership** relation, $\in$. If A is a set and a is any object, then $a \in A$ means that a is an element (or member) of A. I the earlier example, $a \in \{a,b,c\}$, but $d \notin \{a,b,c\}$. Sets can also have other sets as members, as in $\{a,\{b,c\}\}$, the set with members a and $\{b,c\}$.

Set **identity** is likewise defined in term of membership: identical sets have identical members. Thus, we say that for any two sets A and B, $A = B$ *iff* for every $x \in A$, $x \in B$ too.

This is known as the Axiom of Extensionality. The Axiom of Extensionality provides a criterion of identity for sets, but it does *not* say which sets exist. (As usual, "*iff*" stands for the logical biconditional connective "if and only if".)

Although it is possible to list the members of some sets, this is often impractical (or impossible!). We define them by **abstraction**. Let D be a domain of discourse and p be *any* predicate that can be meaningfully applied to the elements of D. Then, $\{x \in D : p(x)\}$ denotes the set of elements x, such that $x \in D$ and $p(x)$. The predicate p can be anything, like the predicate "be a dog", $\{x : x \text{ is a dog}\}$, and it may be expressed however you see fit, e.g. the set $\{x : x > 0\}$ is identical to the set $\{x : x \text{ is greater than zero}\}$, although they are "written" differently. Notice that here we didn't specify what x is about, i.e. we did not specify that the property p only makes sense when it applies to numbers. Most of the time this is OK, we rely on our intuitions about what objects can be sensible predicated of p.

Some use of a vertical bar instead of the colon, $\{x \in D|p(x)\}$, and call this method the *predicate* notation. In linguistics sometimes we also see the variable written after the description of the set, $\{p(x) : x \in D\}$.

The number of members in a set A is called the *cardinality* of A, and written $|A|$ (sometimes $\#(A)$ as well). Infinite sets also have cardinalities, but these cardinalities are not expressed by natural numbers. For instance, the cardinality of the set of integers is $\aleph_0$, read "aleph-zero".

A set with cardinality zero is called the **empty set**, denoted by the symbol $\varnothing$. We can define it as the set of all things that are different from themselves,

Sometimes also called the null set, and denoted by empty curly brackets, $\{\}$.

$\varnothing = \{x : x \neq x\}$. Since all things are equal to themselves, the set contains no members, and thus has cardinality zero.

1.2 Relations between sets

Sets can be related to each other depending on whether they share some of their members. Let A and B be sets. A is a **subset** of B, written $A \subseteq B$ *iff* for every x, if $x \in A$, then $x \in B$. In words, A is a subset of B when every member of A is also a member of B. A and B may be identical: in fact, $A = B$ *iff* $A \subseteq B$ and $B \subseteq A$. If they are not identical, then A is a **proper subset** of B, written $A \subset B$. We can also use the denials of these two relations to express that A is not a subset of B, $A \nsubseteq B$, or that A is not a proper subset of B, $A \not\subset B$.

It follows from the Axiom of Extensionality.

Notice that given the definition of subset above, the empty set is a subset of every set: for any set A, $\varnothing \in A$. The reasoning could go as follows: how could $\varnothing$ *fail* to be a subset of some set A? There would have to be some member of $\varnothing$ that is not in A. But this is impossible, since $\varnothing$ does not have any members whatsoever, and thus (trivially) satisfies the definition of subsethood.

But **not** $\varnothing \in A$! This is a contingency, a set may or may not have $\varnothing$ as one of its members.

Notice that the empty set can be the sole member of a set, $\{\varnothing\}$, which is different from $\varnothing$ itself: $|\{\varnothing\}| = 1$ whereas $|\varnothing| = 0$.

Sometimes it is useful to refer to the set whose member are all the subsets of some set A. This set is called the **power set** of A, written $\wp(A)$. If $A = \{a, b, c\}$, then $\wp(A) = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$. The name "power set" comes from the fact that if the cardinality of A is n, then $\wp(A)$ has cardinality 2^n.

The funky "p" is a Weiestrass p also called pe. A common alternative is simply to write $POW(A)$, and sometimes 2^A.

Any subset of $\wp(A)$ is called a **family of sets** over A.

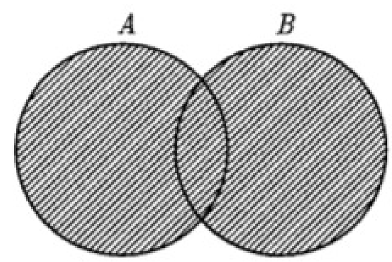

Figure 1: Venn diagram of set-theoretic union $A \cup B$.

1.3 Set operations

We now introduce two operations which take a pair of sets and produce another set. The **union** of any two sets A and B is:

(1) $A \cup B =_{def} \{x : x \in A \text{ or } x \in B\}$

The **intersection** of A and B is:

(2) $A \cap B =_{def} \{x : x \in A \text{ and } x \in B\}$

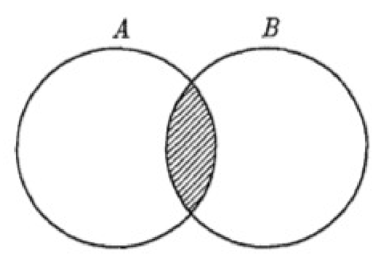

Figure 2: Venn diagram of set-theoretic intersection $A \cap B$.

Notice that the disjunction "or" in the definition of set union allows an object to be a member of both A and B. For this reason, the "or" is an inclusive disjunction.

The operations of set-theoretic union and intersection can easily be generalized to apply to more than two sets, in which case we write the corresponding union or intersection sign in front of the set of sets to be operated on: e.g. for any sets A, B and C, $\bigcup\{A, B, C\}$ = the set of all elements in A or in B or in C. Similarly, $\bigcap\{A, B, C\}$ = the set of all elements in A and in B and in C. Also, note that $A \cup \varnothing = A$ and $A \cap \varnothing = \varnothing$, $A \cup \varnothing = A$.

Another important binary operation on arbitrary sets A and B is the **dif-**

ference, written $A - B$, which "substracts" from A all the objects which are also in B. $A - B$ is also known as the **relative complement** of A and B.

The same operation is sometimes written with a backslash: $A \backslash B$.

(3) $A - B =_{def} \{x : x \in A \text{ and } x \notin B\}$

Thus, if $A = \{a, b, c\}$ and $B = \{b, c, d\}$, then $A - B = \{a\}$.

The operation that yields the difference of A and B must be distinguished from the **complement** of set A, written $\bar{A}$ (or A'), which is the set consisting of everything not in A.

(4) $\bar{A} =_{def} \{x : x \notin A\}$

Where do the objects that do not belong to A come from? Usually, statements about sets are made against a background of assumed objects which comprise the **universe of discourse** (or domain of discourse). This is declared by convention of using the symbol U for it. The complement of A is then set of all objects in U that are not in A: $\bar{A} = U - A$. For this reason, $\bar{A}$ is referred to as the **absolute complement** of A (as in Figure 4).

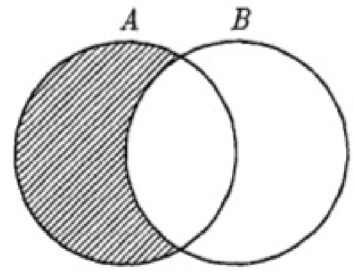

Figure 3: Venn diagram of set-theoretic difference $A - B$.

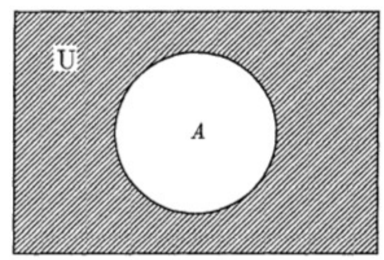

Figure 4: Venn diagram of set-theoretic complement $\bar{A}$.

1.4 Set-theoretic equalities

From the properties of sets and operations on sets described above there are a number of general equalities that follow. Below you will find some of the more common (taken from Partee et al. 1990).

Proofs that these equalities really hold can be found in Causey (1994) §2 and in Partee et al. (1990) §7.

1. *Idempotent Laws*
   (a) $X \cup X = X$ (b) $X \cap X = X$
2. *Commutative Laws*
   (a) $X \cup Y = Y \cup X$ (b) $X \cap Y = Y \cap X$
3. *Associative Laws*
   (a) $(X \cup Y) \cup Z = X \cup (Y \cup Z)$ (b) $(X \cap Y) \cap Z = X \cap (Y \cap Z$
4. *Distributive Laws*
   (a) $X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z)$
   (b) $X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$
5. *Identity Laws*
   (a) $X \cup \emptyset = X$ (c) $X \cap \emptyset = \emptyset$
   (b) $X \cup U = U$ (d) $X \cap U = X$
6. *Complement Laws*
   (a) $X \cup X' = U$ (c) $X \cap X' = \emptyset$
   (b) $(X')' = X$ (d) $X - Y = X \cap Y'$
7. *DeMorgan's Law*
   (a) $(X \cup Y)' = X' \cap Y'$ (b) $(X \cap Y)' = X' \cup Y'$
8. *Consistency Principle*
   (a) $X \subseteq Y$ iff $X \cup Y = Y$ (b) $X \subseteq Y$ iff $X \cap Y = X$

2 Relations and Functions

2.1 Ordered tuples

Although sets are collections of unordered objects, we can use ordinary sets to define an **ordered pair** of elements, written $\langle a, b\rangle$ for example, in which a is considered the **first member** and b is the **second member** of the pair. The definition is as in 5: the first element, a, is the element occurring in the singleton $\{a\}$ and the second element, b, is the member in the set $\{a, b\}$ that is not also in $\{a\}$.

In some texts ordered pairs are written between parenthesis: (a, b).

(5) $\langle a, b\rangle =_{def} \{\{a\}, \{a, b\}\}$

Because "order matters", we have that if $a \neq b$, then $\langle a, b\rangle \neq \langle b, a\rangle$, and, if $a = b$, then $\langle a, b\rangle = \langle b, a\rangle$.

We can generalize the notion of ordered pair to that of **ordered *n*-tuple**. An ordered triple, for instance, is defined as an ordered pair where the first element is itself an ordered pair:

(6) $\langle a, b, c\rangle =_{def} \langle\langle a, b\rangle, c\rangle$

Following the same strategy above we can further generalize to n-tuples of arbitrary arity:

(7) $\langle x_1, \ldots, x_n\rangle =_{def} \langle\langle x_1, \ldots, x_{n-1}\rangle, x_n\rangle$

An important kind of n-tuple is the **cartesian product**, a set of ordered pairs formed from two sets A and B by taking an element of A as the first member of the pair and an element of B as the second member. The cartesian product of A and B, written $A \times B$, is the set of all such ordered pairs:

(8) $A \times B =_{def} \{\langle x, y\rangle : x \in A \text{ and } y \in B\}$

Note that $A \times \varnothing = \varnothing$.

For instance, if $A = \{a, b\}$ and $B = \{1, 2\}$, then cartesian product of A and B is $A \times B = \{\langle a, 1\rangle, \langle a, 2\rangle, \langle b, 1\rangle, \langle b, 2\rangle\}$. As we did with ordered pairs, we can generalize the notion of cartesian product to that of ***n*-ary cartesian product**. Suppose that we have a series of n sets $A_1, \ldots, A_n$. Then:

(9) $A_1 \times \ldots \times A_n =_{def} \{\langle a_1, \ldots, a_n\rangle : a_i \in A_i\}$

That is, the n-ary carterian product of the series is such the set of all n-tuples $\langle a_1, \ldots, a_n\rangle$ where $a_1 \in A_1$, $a_2 \in A_2$, etc. all they way up to $a_n \in A_n$.

We write A^2 to represent the cartesian product of $A \times A$, and A^n for $A \times \ldots \times A$ (n times).

2.2 Relations

Ordered tuples can be useful for many things. One of them is representing **relations**. A relation is simply that, the intuitive connection that holds (or does not hold) of a set of objects. For instance, the relation "mother of" holds between all mothers and their children, but not between the children themselves. Transitive verbs, at least some of them, like "see" can also be regarded as denoting an sort of relation between a pair objects, where the first sees the second.

Relations that hold between pairs of objects are called **binary relations** and can naturally be represented as sets of ordered pairs. For instance, $\{\langle a, b\rangle : a \text{ is the mother of } b\}$ denotes the aforementioned relation between mother and children. Because these are ordered pairs, order matters, and thus we can collect all the first members of all pairs in a single set (i.e. the "mothers" in our earlier example), and all the second members in a different set (i.e. the children). These sets are so commonly used that they have their own names:

(10) Let R be a binary relation. Then:

a. **Domain**: $\{a : \langle a, b\rangle \in R\}$

b. **Range**: $\{b : \langle a, b\rangle \in R\}$

Another name for the range is codomain.

The **inverse** of a relation R, written R^{-1}, is so to say the "flip-image" of R, where all the elements in each ordered pair have swapped their positions in the pair.

(11) Let R be a relation. Then $R^{-1} =_{def} \{\langle b, a\rangle : \langle a, b\rangle \in R\}$

Note that $(R^{-1})^{-1} = R$.

Just like we represent binary relations as sets of ordered pairs, we can write sets of n-tuples in order to represent n-ary relations. For instance, we could represent the verb "give" as the set of triples $\{\langle a, b, c\rangle : a \text{ gave } b \text{ to } c\}$, and so on.

Because of the way we defined n-tuples in (7), it follows than an n-ary relation R is equivalent to a binary relation whose domain is an n-1-ary relation, that is, a set of n-1-tuples.

2.3 Functions

Functions are a special kind of relations:

(12) Let $R \subseteq A \times B$. R is a function from A to B if the following two conditions hold:

a. The domain of R is A.

b. If $\langle a, b\rangle \in R$ and $\langle a, c\rangle \in R$, then $b = c$.

The requirement that the domain of R be equal to A is a property of **total functions**. Later on in the semester we will see **partial functions**, where condition (12a) does not hold.)

We write $f : A \to B$ to express a function f from A to B, we say that $f(x)$ is the unique y such that $\langle x, y\rangle \in f$.

Functions can be of different types:

(13) Let $f : A \to B$. Then:

a. f is a **surjection** *iff* the range of $f = B$.

b. f is a **injection** *iff* each $a \in A$ is mapped to a different $b \in B$.

c. f is a **bijection** *iff* f is a surjection and an injection.

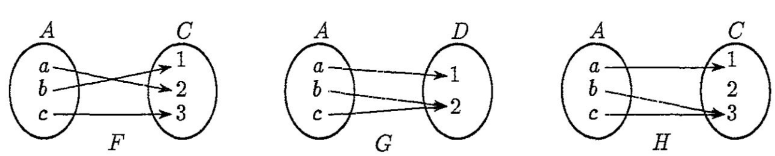

Figure 5: F is a bijection, G is a surjection (not an injection) and H is an injection (not a surjection).

Surjections are said to be "onto" functions, injections are "one-to-one" and bijections are "one-to-one correspondences".

Since a function f is just a special type of relation, we can also speak of the inverse of f, f^{-1}. What's key to understand is that the inverse of a function f isn't necessarily a function itself. The inverse of a surjection f that is not an injection is not itself a function, since some element in the domain of f^{-1} must be mapped to more than one element in its range. The inverse of an injection f that is not a surjection is not a function either, since there are elements in the domain of f^{-1} that are not mapped to any element in its range. However the inverse of a bijection is always a bijection, and thus a function.

But it could be a "partial" function.

An ***n*-ary function** is a function whose domain is a set of n-tuples rather than a set of (simplex) elements. We write n-ary functions as in following this schema: $f : (A_1 \times \ldots \times A_n) \to B$. Functions whose domain is a set of ordered pairs are called **binary functions** (e.g. $\{\langle\langle x,y\rangle,z\rangle : z = x+y\}$), those functions whose domain is a set of ordered triples are called **ternary functions** (e.g. $\{\langle\langle\langle x,y\rangle,z\rangle,w\rangle : w = x+y+z\}$), and so on.

The term "function" is sometimes restricted to unary functions; functions with greater arity than one are sometimes called "operations".

2.4 Composition

Given two unary functions $f : A \to B$ and $g : B \to C$, we may form a new function h from A to C, called the **composite**, or the **composition** of f and g, written $g \circ f$.

(14) $g \circ f =_{def} \{\langle x,z\rangle : \text{for some } y, \langle x,y\rangle \in f \text{ and } \langle y,z\rangle \in g\}$

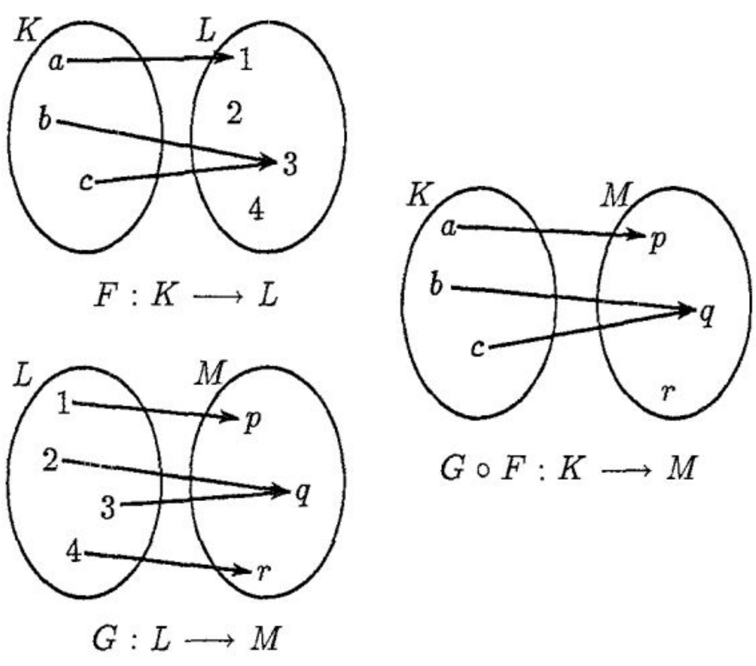

Figure 6: The composition $G \circ F$ of F and G.

The notation $g \circ f$ may seem to read backwards, but the value of a function f at an argument a is $f(a)$, and the value of g at the argument $f(a)$ is written

$g(f(a))$. By the definition of composition, $g(f(a))$ and $(g \circ f)(a)$ produce the same value. Thus, for all x, $g \circ f(x) = g(f(x)) = (g \circ f)(x)$.

Another way of writing the composition of f and g is $g\langle f\rangle$. This notation becomes more useful when we consider the composition of n-ary functions:

(15) Let g be an n-ary function and let $f_1, \ldots f_n$ be a series of n m-ary functions. Then:
$g\langle f_1, \ldots f_n\rangle =_{def}$ the m-ary function s.t. for any m-ary sequence $a_1, \ldots, a_m$,
$g\langle f_1, \ldots f_n\rangle(\langle a_1, \ldots, a_m\rangle) = g(f_1(\langle a_1, \ldots, a_m\rangle), \ldots f_n(\langle a_1, \ldots, a_m\rangle))$

For instance, take the following three functions $g = \{\langle\langle x, y\rangle, z\rangle : z = x + y\}$, $f = \{\langle x, y\rangle : y = x - 1\}$ and $h = \{\langle x, y\rangle : y = x + 2\}$. Then we have that $g\langle f, h\rangle(2) = g(f(2), h(2)) = g(1, 4) = 5$. That is, the binary function g takes as its two inputs the outputs of applying argument 2 to f and h respectively: $g\langle f, h\rangle = \{\langle x, y\rangle : y = (x - 1) + (x + 2)\}$.

2.5 Characteristic functions

One of the most useful functions is the **characteristic function**. A characteristic function singles out a subset A of some domain U and has value 1 at points of A and 0 at points of $U - A$.

Also known as the "indicator" function in some areas. 0 and 1 are also sometimes expressed as F and T, for "false" and "true"respectively.

(16) For every set $A \subseteq B$ there is a unique function $f_A : B \to \{0, 1\}$, s.t. for all $b \in B, f_A(b) = 1$ *iff* $b \in A$.

In other words, given some set B, we can determine a function f_A that maps all the elements that belong to some subset A of B to 1, and maps all the elements in $B - A$ to 0.

Because every set $A \subseteq B$ has a unique characteristic function, we can reason inversely as well and determine the **characteristic set** A_f of a some function $f : B \to \{0, 1\}$:

(17) For every function $f : B \to \{0, 1\}$, there is a unique set $A_f \subseteq B$ s.t. for all $b \in B, b \in A_f$ *iff* $f(b) = 1$.

That is, given some characteristic function f, we can determine the set of all objects that f maps to value 1. This is a very convenient correspondence that is oftentimes exploited in linguistics (mostly for convenience). But technically, sets and their corresponding characteristic functions are different objects. We can also generalize the notion of characteristic function to n-ary relations: for all $\langle a_1, \ldots, a_n\rangle \in A_1 \times \ldots \times A_n$, $f_R(\langle a_1, \ldots, a_n\rangle) = 1$ *iff* $\langle a_1, \ldots, a_n\rangle \in R$.

For instance the set $\{a, b, c\}$ is not the same as $\{\langle a, 1\rangle, \langle b, 1\rangle, \langle c, 1\rangle, \langle d, 0\rangle, \langle e, 0\rangle\}$.

2.6 Schönfinkelization / Currying of functions

We have talked about how functions may contain ordered n-tuples in their domain. Consider first a binary function $f : (A \times B) \to C$. **Currying** is the process by which from f we obtain a function g that describes f as a sequence of unary functions; i.e. where arguments are never n-tuples.

This is a **very** useful technique for linguistic purposes; after all, arguments of "linguistic relations" only come one at a time–assuming binary branching syntactic trees, that is.

(18) Let f be a binary function $f : (A \times B) \to C$. There's a unique function, $Cur(f) : A \to B \to C$, s.t.: for all $\langle a,b \rangle \in A \times B$, $f(\langle a,b \rangle) = c$ *iff* $Cur(f)(x)(y) = c$.

The idea is that instead of having a function that takes one element from A and one from B and maps them to some element in C, we now have a function that takes one element from A and returns *another function* from B to C. Schematically, assuming $f : \{a,b\} \times \{c,d\} \to \{0,1\}$:

(19)
$$\text{If } f = \begin{bmatrix} \langle a,c \rangle & \to & 1 \\ \langle a,d \rangle & \to & 0 \\ \langle b,c \rangle & \to & 0 \\ \langle b,d \rangle & \to & 1 \end{bmatrix}, \text{ then } Cur(f) = \begin{bmatrix} a & \to & \begin{bmatrix} c & \to & 1 \\ d & \to & 0 \end{bmatrix} \\ b & \to & \begin{bmatrix} c & \to & 0 \\ d & \to & 1 \end{bmatrix} \end{bmatrix}$$

Once again, we can generalize to cases with n-ary function. We can do this by exploiting the equivalence, introduced in (5), that any n-tuple $\langle a_1, \ldots, a_n \rangle$ can be represented as an ordered pair $\langle \langle a_1, \ldots a_{n-1} \rangle, a_n \rangle$. Then, if we had a function with ordered triples in its domain, $f(\langle \langle a,b \rangle, c \rangle) = d$, we could represent it as a binary function $Cur(f)(\langle a,b \rangle)(c) = d$. In that case, $Cur(f)$ would be a binary function, and by iterating the process, we could further "break down" $Cur(f)$ into the unary function $Cur(Cur(f))$, so that $f(\langle \langle a,b \rangle, c \rangle) = d$ *iff* $Cur(Cur(f))(a)(b)(c) = d$. This reasoning process generalizes to the following:

(20) Let f be an n-ary function $f : A_1 \times \ldots \times A_n \to C$. There is a unique function $CUR(f) : A_1 \to \ldots A_n \to C$ s.t.: for all $\langle a_1, \ldots a_n \rangle \in A_1, \ldots, A_n$, $f(\langle a_1, \ldots, a_n) = c$ *iff* $CUR(f)(a_1)\ldots(a_n) = c$.

Recall that every n-ary relation has a unique characteristic function. Then it follows from the characterization of currying that every n-ary relation has its own unique characteristic curried function, $CUR(f_R)$. Let $R \subseteq A_1 \times \ldots \times A_n$ be any n-ary relation. Then, the unique $CUR(f_R)$ function is such that $\langle a_1, \ldots, a_n \rangle \in R$ *iff* $CUR(f_R)(a_1)\ldots(a_n) = 1$.

3 Indexing

Throughout the notes above we have used the following informal notation: $A = \{a_1, \ldots, a_n\}$, meaning that "A is a set consisting of n different elements a_i, for all $0 < i \leq n$". This informal notation implies the existence of a bijection $f : \mathbb{N} \to A$, s.t. $f(i) = a_i$. This sets up the following general definition:

There is no special reason why we use numbers as indices, any symbol would do. But numbers are very convenient.

(21) Let N and A be sets such that there is a bijection $f : N \to A$. We can say that A is an indexed family, and that N is the index set.

Suppose A is an indexed family whose index set is N. We can represent A as the set for which the inverse of f, as mentioned above, returns a unique n: $\{a : \text{there is a } n \in N \text{ s.t. } f^{-1}(a) = n\}$. Or, more compactly: $\{a_n\}_{n \in N}$.

Throughout the notes above, we have also used numerical indices to represent n-tuples, as in $\langle a_1, \ldots, a_n \rangle$. We can adapt our notation above as means for compactly representing n-tuples. For instance, the following are two possible ways of expressing the tuple $\langle a_1, \ldots, a_n \rangle$: $\langle a_i \rangle_{i \in \{1,\ldots,n\}}$ and $\langle a_i \rangle_{i \leq n}$.

We can also perform operations on indexed sets. For instance, suppose S is an indexed family and I is an index set. Then, $\bigcup S = \{x : x \in S$ for some $i \in I\}$. Similarly, $\bigcap S = \{x : x \in S$ for all $i \in I\}$..

Thus, if $S = \{S_1, \ldots S_n\}$, $\bigcup S = S$, but $\bigcap S$ is not necessarily equal to S

Some popular alternative notations include the following:

- If S is an indexed family of sets with index set I, then $\bigcup S$ is written $\bigcup\limits_{i \in I} S_i$, and $\bigcap S$ is written $\bigcap\limits_{i \in I} S_i$.
- If S is an indexed family of sets with index set $\{1, \ldots, n\}$, then $\bigcup S$ is written $\bigcup\limits_{i=1}^{n} S_i$, and $\bigcap S$ is written $\bigcap\limits_{i=1}^{n} S_i$.

References

Cable, S. (2014). From Logic to Montague Grammar. Course Notes. University of Massachusetts Amherst. Available online: `https://ling.auf.net/lingbuzz/001989`.

Causey, R. L. (1994). *Logic, sets and recursion.* Jones and Barlett Publishers.

Partee, B., ter Meulen, A., and Wall, R. E. (1990). *Mathematical Methods in Linguistics.* Dordrecht: Kluwer.