Mathematical Background

LING 553

September 3, 2008

1. SETS

Definition: A **SET** is simply a collection of objects.

Definition: An object in a set is called an **ELEMENT**. "x is an element of S" (or "x is a member of S") is written $x \in S$.

Definition: The number of elements in a set is called the **CARDINALITY** of the set. "The cardinality of S" is written $|S|$.

1.1. *Some special sets*

- The set of everything being considered is called **THE UNIVERSE** and is written U. (This will vary from context to context.)
- Any set with one member is called a **SINGLETON SET**.
- The set with no members at all is called the **EMPTY SET** (or **NULL SET**) and written $\varnothing$.
- Occasional named sets, e.g. D_e, D_t, $\mathbb{N}$, $\mathbb{Z}$, etc., which will be described when named.

1.2. *Things to do with sets*

Definition: A is **IDENTICAL TO** B, written A = B, iff they have exactly the same members.

Definition: A is a **SUBSET** of B ($A \subseteq B$) iff every element of A is an element of B.
A is a **PROPER SUBSET** of B ($A \subset B$) iff $A \subseteq B$ and $A \neq B$.[1]

Definition: The **INTERSECTION** of A and B ($A \cap B$) is the set that contains all and only those elements that are in *both* A *and* B.

Definition: The **UNION** of A and B ($A \cup B$) is the set that contains all and only those elements that are in A *or* B (or both).[2]

Definition: The **DIFFERENCE** of A and B ($A - B$) is the set of all individuals that are in A and not in B.

Definition: The **COMPLEMENT** of A (A') is the set of all individuals that are not in A, with respect to some "universe of discourse" U: that is, $U - A$.

Definition: The **POWER SET** of A, $\wp(A)$, is the set of all subsets of A.

---

[1] Also: A is a (**PROPER**) **SUPERSET** of B iff B is a (proper) subset of A, written $A \supseteq B$, $A \supset B$. But this is rare.

[2] Additional notation: $A \cap B \cap C$ is often written as $\bigcap\{A, B, C\}$, and $A \cup B \cup C$ as $\bigcup\{A, B, C\}$.

1.3. How to specify a set

1.3.1. List Notation

Simply write out a list of the elements of the set.

- *A* = {Brown, Columbia, Cornell, Dartmouth, Harvard, Penn, Princeton, Yale}
- *B* = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, ...}

Advantages
- It's immediately clear what's in the set.
- It works for pretty much any set.

Disadvantages
- The larger the cardinality of the set, the harder it is to write out the full list.
- Once the set is infinite, literally listing the elements becomes impossible. It can still be done using ellipses, but it may be ambiguous or unclear:

  F = {1, 3, 5, 7, ...}
  P = {6, 28, 496, ...}

1.3.2. Predicate Notation

Describe the members of the set, rather than naming each one. Frequently used notation: "the set of all *x* for which it's true that [condition]" is written {*x* | [condition]}. Examples:

- *A* = the set of Ivy League schools, *or*
  {the Ivy League schools}, *or*
  {*x* such that *x* is an Ivy League school}, *or*
  {*x* | *x* is an Ivy League school}

- *B* = {*x* | *x* is a positive integer}

Advantages
- Concise and unambiguous, both of which are quite handy for infinite sets

Disadvantages
- Doesn't work for every set
- Russell's Paradox

<u>Russell's Paradox</u>

Let *S* be the set of sets which do not contain themselves, i.e. $S = \{x \mid x \notin x\}$.
Question: is *S* a member of *S*?

(i) If *S* is a not a member of *S*, then it matches the criterion for inclusion in *S*, and therefore it must be a member of *S*.
(ii) If *S* is a member of *S*, then that's because it matches the criterion for inclusion in *S*, i.e. because it's not a member of itself, i.e., *S* is not a member of *S*.

Therefore, *S* is neither a member of *S*, nor not a member of *S*.

Moral of the story: just because you can write it down doesn't mean you *should*. Or, a little more formally: just because you can write it in predicate notation doesn't make it a well-defined set.

- One solution: everything in a set should be the same "type", e.g.
  - Integers; people; sets of integers; sets of sets of integers...

  This will happen to be a feature of what we do, though we won't require it.

2. TUPLES

Definition: A **SEQUENCE** is a list of objects in a particular order.

Finite sequences are also called **TUPLES**; a sequence with *n* elements is an *n*-tuple. (A 2-tuple is usually called an "ordered pair".)

Definition: The **CARTESIAN PRODUCT** of two sets A and B ($A \times B$) is the set of ordered pairs:
$\{<x, y> \mid x \in A \text{ and } y \in B\}$

Technical note: tuples can be defined in terms of sets by saying $<x, y> =_{def} \{x, \{x, y\}\}$. It's handy to know—it means that, by introducing tuples, we haven't really added anything new to our system beyond just sets. We've only added some notation. On the other hand, the notation is handy and we'll never really need to use the $\{x, \{x, y\}\}$ underpinning, so you can pretty much ignore it.

3. RELATIONS

Definition: A **RELATION** is a set of pairs.
The **DOMAIN** of a relation R is $\{x \mid \text{there is some } y \text{ such that } \langle x, y\rangle \in R\}$.
The **RANGE** of a relation R is $\{y \mid \text{there is some } x \text{ such that } \langle x, y\rangle \in R\}$.

Specifically: If $R \subseteq A \times B$, then R is a *relation from A to B*.
If $R \subseteq A \times A$, then R is a *relation in A*.
If $\langle x, y\rangle \in R$, also written $R(x, y)$, Rxy, or xRy, then R *holds* between x and y.

Definition: The **COMPLEMENT** of a relation $R \subseteq A \times B$, written R', is the set of pairs in $A \times B$ that are not in R, i.e. $\{\langle x, y\rangle \mid \langle x, y\rangle \notin R\}$

Definition: The **INVERSE** of a relation $R \subseteq A \times B$, written R^{-1}, is the set of pairs in R with their elements reversed, i.e. $\{\langle y, x\rangle \mid \langle x, y\rangle \in R\}$

4. FUNCTIONS

Definition: F is a **FUNCTION** from A to B, written $F : A \rightarrow B$, if

F is a relation from A to B such that:
(a) each element in the domain of F maps to only one element in the range, and
(b) domain(F) = A

Note: if domain(F) $\subset$ A, then F is called a *partial function*. In general, "function" by itself is used for complete functions only, though we may occasionally see partial functions.

- If $\langle x, y\rangle \in F$, then $F(x) = y$, read "F maps x to y".
  In $F(x) = y$: x is the **ARGUMENT**, y is the **VALUE**.

- If each element in the range of F is mapped to by only one element in the domain—i.e., the converse of (a) in the definition—then F is **ONE-TO-ONE**. (If not, F is MANY-TO-ONE.)

- If range(F) = B—i.e., the converse of (b)—then F is **ONTO** (or a function "onto B"). (If not, F is INTO or a function "into B".)

- If F is one-to-one and onto, F is called a **ONE-TO-ONE CORRESPONDENCE**. (Note that in this case F^{-1} is also a function.)

4.1. Functions and Sets

Suppose that $U = \{a, b, c, d, e, f, g, h, i, j\}$, $D_t = \{\text{TRUE}, \text{FALSE}\}$, and $S = \{a, b, c, d, e\}$.

- $a \in S$: True. $b \in S$: True. $f \in S$: False. (etc. etc.)

Rather than giving this kind of list, we can write a function $\varphi{:}U \times D_t$ to represent it:

$$\varphi = \begin{bmatrix} a \rightarrow \text{TRUE} \\ b \rightarrow \text{TRUE} \\ c \rightarrow \text{TRUE} \\ d \rightarrow \text{TRUE} \\ e \rightarrow \text{TRUE} \\ f \rightarrow \text{FALSE} \\ g \rightarrow \text{FALSE} \\ h \rightarrow \text{FALSE} \\ i \rightarrow \text{FALSE} \\ j \rightarrow \text{FALSE} \end{bmatrix}$$

This is the **characteristic function** of S: the function such that

$$\text{for all } x \in U,\ \varphi(x) = \begin{cases} \text{TRUE} & \text{if } x \in S \\ \text{FALSE} & \text{otherwise} \end{cases}.$$

This turns a set into a function. We can turn a function φ into the set "characterized by" φ:

> S = {the individuals x such that φ maps x to TRUE} *or*
> $S = \{x \in U \mid \varphi(x) = \text{TRUE}\}$

i.e., <u>φ(a) = TRUE</u> is equivalent to <u>$a \in S$</u>.

- Quick moral: functions are different from sets. But we can sometimes talk about them interchangeably. (And thus we might write $a \in \varphi$ as a shorthand.)

4.2. What about algebra?

The kinds of functions that are more familiar....$\psi{:}\mathbb{N} \times \mathbb{N}$.[3]

$$\psi = \begin{bmatrix} 1 & \rightarrow & 2 \\ 2 & \rightarrow & 3 \\ 3 & \rightarrow & 4 \\ 4 & \rightarrow & 5 \\ & \dots & \end{bmatrix}$$

i.e., $\psi(x) = x + 1$, i.e. the function from natural numbers to natural numbers such that it maps each natural number to its successor.

There's not really a set characterized by this function. Generally: a set containing elements of S can be expressed as a function $f{:}S \times D_t$. And a function can be expressed as a set if the range of the function is D_t.

But: since functions (for us) are ordered pairs, we can instead use...

$$\psi = \{<1,2>, <2, 3>, <3, 4>, <4, 5>, \dots\}$$

(i.e., $\psi(<1,2>) = \text{TRUE}$, $\psi(<1,3>) = \text{FALSE}$, $\psi(<2,2>) = \text{FALSE}$, ...)

---

[3] $\mathbb{N}$ = the set of positive integers = {1, 2, 3, 4, ...}

4.3. *How to specify a function*

4.3.1. List Notation

Write out the mappings of the function, as above.

Advantages
- It's immediately clear what's in the ~~set~~ function.
- It works for pretty much any ~~set~~ function.

Disadvantages
- The larger the cardinality of the ~~set~~ function, the harder it is to write out the full ~~list~~ mapping.
- Once the ~~set~~ function is infinite, literally listing the ~~elements~~ mappings becomes impossible.

Note: that should look familiar, from the "list notation for sets" earlier.

4.3.2. Descriptively

A function must specify what it's a function *from*, what it's a function *to*, and which function it is. For example:

- φ is that function from the universe *U* to the set {TRUE, FALSE} which, given an element in the universe, returns TRUE if the element is a member of set S and FALSE otherwise. (See previous page for list notation.)

- ψ is that function from the (set of) positive integers to the (set of) positive integers which, given a positive integer, returns its successor.

- χ is that function from the (set of) human beings to the (set of) cities which, given any human being, returns the city that that human being was born in.

List notation for ψ and *h* requires ellipses (and may not be perspicuous):

$$\psi = \begin{bmatrix} 1 \rightarrow 2 \\ 2 \rightarrow 3 \\ 3 \rightarrow 4 \\ \vdots \end{bmatrix} \qquad \chi = \begin{bmatrix} \text{Lance Nathan} \rightarrow \text{Atlanta, GA} \\ \text{Bill Clinton} \rightarrow \text{Hope, AR} \\ \text{Marco Polo} \rightarrow \text{Venice, Italy} \\ \vdots \end{bmatrix}$$

For example, $\psi(1) = 2$; χ(Lance Nathan) = Atlanta, GA.

In general, we'll prefer descriptions.

4.4. Functions within functions

The value a function returns can also be another function, or a set. For instance:

- *m* is that function from human beings to functions from movie trilogies to the power set of {1, 2, 3} which, given a human being, returns that function which, given a movie trilogy, returns the set of numbers such that the human being has seen the *n*th movie in the trilogy if and only if *n* is in the set.

In list notation:
$$m = \begin{bmatrix} \text{Lance Nathan} \rightarrow \begin{bmatrix} \text{Pirates of the Caribbean} \rightarrow \{1,3\} \\ \text{The Lord of the Rings} \rightarrow \{1,2,3\} \\ \text{The Godfather} \rightarrow \varnothing \\ \vdots \end{bmatrix} \\ \text{Roger Ebert} \rightarrow \begin{bmatrix} \text{Pirates of the Caribbean} \rightarrow \{1,2,3\} \\ \text{The Lord of the Rings} \rightarrow \{1,2,3\} \\ \text{The Godfather} \rightarrow \{1,2,3\} \\ \vdots \end{bmatrix} \\ \vdots \end{bmatrix}$$

Thus,
$$m(\text{Lance Nathan}) = \begin{bmatrix} \text{Pirates of the Caribbean} \rightarrow \{1,3\} \\ \text{The Lord of the Rings} \rightarrow \{1,2,3\} \\ \text{The Godfather} \rightarrow \varnothing \\ \vdots \end{bmatrix}$$

And because
$$\begin{bmatrix} \text{Pirates of the Caribbean} \rightarrow \{1,3\} \\ \text{The Lord of the Rings} \rightarrow \{1,2,3\} \\ \text{The Godfather} \rightarrow \varnothing \\ \vdots \end{bmatrix}(\text{PotC}) = \{1,3\}$$
, we can write

m(Lance Nathan)(Pirates of the Caribbean) = {1,3}.[4]

Of course, neither list notation nor the description is especially easy to read....

[4] Note that functions always operate in this order: *m*(A)(B) means "give A to *m* as an argument, and give B to the result". It never means "give B to *m* as an argument, and give A to the result", or "give B to A as an argument, and give the result to *m*". If we wanted the latter, we'd write *m*(A(B)).

5. LAMBDA NOTATION

In semantics, we use a particular notation involving lambdas, where *lambda* (λ) indicates "this thing is a function."

The general format of a function in λ-notation is [λA : B . C]. In this:

A is called the **argument variable**
(a letter standing for an arbitrary argument),
B is called the **domain condition**, introduced by a colon
(which puts a condition on the possible values of the argument variable),
C is called the **value description**, introduced by a period
(which gives the value assigned by the function)

For example: function ψ:

> ψ is that function from the (set of) positive integers to the (set of) positive integers which, given a positive integer, returns its successor.

can be written:

	that function...	from positive integers	to	positive integers, which,	
	given a	positive integer,	returns	its successor	
[	λx	$: x \in \mathbb{N}$	**.**	$x + 1$	**]**

Similarly:

> χ is that function from the (set of) human beings to the (set of) cities which, given any human being, returns the city that that human being was born in. ≡
>
> χ = [λ*x* : *x* is a human being . the city *x* was born in]

This becomes even more helpful with functions embedded in functions:

> *m* is that function from human beings to functions from movie trilogies to the power set of {1, 2, 3} which, given a human being, returns that function which, given a movie trilogy, returns the set of numbers such that the human being has seen the *n*th movie in the trilogy if and only if *n* is in the set. ≡
>
> *m* = [λ*x* : *x* is a human being .
> [λ*T* : *T* is a movie trilogy . {*n* ∈ {1,2,3} | *x* has seen the *n*th movie in trilogy *T*}]]

5.1. What about truth values?

U = {a, b, c, d, e, f, g, h, i, j}　　　　S = {a, b, c, d, e}

$$\varphi = \begin{bmatrix} \texttt{a} \rightarrow \textsc{true} & \texttt{b} \rightarrow \textsc{true} \\ \texttt{c} \rightarrow \textsc{true} & \texttt{d} \rightarrow \textsc{true} \\ \texttt{e} \rightarrow \textsc{true} & \texttt{f} \rightarrow \textsc{false} \\ \texttt{g} \rightarrow \textsc{false} & \texttt{h} \rightarrow \textsc{false} \\ \texttt{i} \rightarrow \textsc{false} & \texttt{j} \rightarrow \textsc{false} \end{bmatrix}$$

φ is that function from the universe U to the set {TRUE, FALSE} which, given an element in the universe, returns TRUE if the element is a member of set S and FALSE otherwise.

Just as:

- $\psi = [\lambda x : x \in \mathbb{N} \, . \, x + 1]$
  takes a number, returns a number

- $\chi = [\lambda x$: x is a human being . the city x was born in]
  takes a human being, returns a city

We want φ to take an individual and return a truth value:

- $\varphi = [\lambda x : x \in U$. TRUE if $x \in S$, FALSE otherwise]

This can be abbreviated to simply $f = [\lambda x : x \in U \, . \, x \in S]$, where any time there's a clause as the value description, it means "TRUE if this clause is true, FALSE otherwise".

5.2. Lambda-conversion

To apply a function written in λ-notation to its argument: remove the λ and its argument variable and domain condition; remove the argument; and substitute the argument wherever the argument variable occurs within the value description,

For example: $[\lambda x \in \mathbb{N}.\, x^2 + 2x + 1](5)$

$= [\cancel{\lambda x \in \mathbb{N}} \, . \, x^2 + 2x + 1]\cancel{(5)} = 5^2 + 2 \cdot 5 + 1 = 36$

Another example: $[\lambda x \in \mathbb{N} \, . \, [\lambda y \in \mathbb{N} \, . \, x^2 - y^2]](5)(4)$

$= [\cancel{\lambda x \in \mathbb{N}} \, . \, [\lambda y \in \mathbb{N} \, . \, x^2 - y^2]]\cancel{(5)}(4) = [\lambda y \in \mathbb{N} \, . \, 5^2 - y^2](4)$

$= [\cancel{\lambda y \in \mathbb{N}} \, . \, 5^2 - y^2]\cancel{(4)} = 5^2 - 4^2 = 25 - 16 = 9$