Lessons for Amie

Doc Benton's Brief *Introduction to the* Joys and *the Art of* Theorem Proving

Christopher P. Benton, PHD

INTRODUCTION

I once had a very bright student named Amie who did so well in her community college courses on calculus and differential equations that she decided to major in mathematics when she transferred to a senior university. However, I immediately knew that she would face a problem. Namely, the fact that upper level college courses in mathematics are often very different from the courses you encounter at the freshman and sophomore levels. Once you reach the junior and senior levels, mathematics becomes much more abstract, a discipline in its own right rather than a tool that must always have an application somewhere else, and there is an increased emphasis on being able to prove theorems. Thus, I decided the best thing I could do for her was to put together a short introduction to theorem proving, and I knew from experience that even a little familiarity with material on theorem proving can make her transition to higher level mathematics a lot easier. And so I put together for her this brief series of lessons to serve as an introduction to more advanced mathematics, and I tried to craft the lessons so that she could take them and explore, to a great degree, on her own. There are certainly many important topics that I have completely left out of this material such as topology and linear algebra, but then again, an introduction can only do so much, and after that each person must define their own journey. Thus, for anyone who comes across these lessons, spend a little time with each one, ponder and explore on your own, and enjoy what insights you acquire!

Lesson 1

Symbolic Logic

This lesson is an introduction to symbolic logic and what we actually mean in mathematics by statements such as "*a* implies *b*" and "*a* if and only if *b*." Below are some common symbols that are used in logic followed by the corresponding math symbols that I will use instead.

LOGIC	**MATH**
$\sim$ or $\neg$	not
$\vee$	or
$\wedge$	and
$a \rightarrow b$	$a \Rightarrow b$
$a \leftrightarrow b$	$a \Leftrightarrow b$

The statement "$a \Rightarrow b$" can be read as "*a implies b*" or "*if a then b*" or "*a is a sufficient condition for b*" or "*b is a necessary condition for a.*"

The statement "$a \Leftrightarrow b$" can be read or written as "*a iff b*" or "*a if and only if b*" or "*a is a necessary and sufficient condition for b.*"

Using the logical connectives above, we can rewrite "$a \Rightarrow b$" as "not (a & not-b)." Similarly, since "$a \Leftrightarrow b$" means "$a \Rightarrow b$ & $b \Rightarrow a$," we can rewrite "$a \Leftrightarrow b$" as "[not (a & not-b)] & [not (b & not-a)]."

In mathematics, for a compound statement *"A & B"* to be true, both of the statements *A* and *B* must be true. On the other hand, for the compound statement *"A or B"* to be true, only one of the statements must be true. Construct some simple examples to convince yourself that this is the correct way to proceed. Also, in mathematics, unless stated otherwise, we always use an *inclusive or*. That means that for *"A or B"* to be true, we either have *A* true or *B* true or both *A* and *B* true. In an *exclusive or*, either *A* or *B* can be true, but not both at the same time.

Now, here are some things for you to either do or look up and respond to.

1. What is a *truth table?*

2. What is the *law of the excluded middle?*

3. What is a *tautology?*

4. What is a *contradiction?*

5. Complete the following *truth tables*. Your final values should be in the column shaded yellow.

Not	A
	T
	F

A	&	B

A	or	B

Not	[A	&	(Not	B)]

{ Not	[A	&	(Not	B)] }	&	{ Not	[B	&	(Not	A)] }

6. What is *modus ponens?* Give an example.

7. What is *modus tollens"* Give an example.

8. Explain why a false statement can imply anything? In other words, why is a statement such as *"If the moon is made of green cheese, then I am the smartest man on the planet"* considered to be true?

9. Use truth tables to show that " not (*a* & not-*b*) " and "not-*a* or b" are *logically equivalen*t. In this context, logical equivalence means that if *a* and *b* are assigned the same truth values, then the truth values of our final compound statements are the same.

10. Consider the statement *"This sentence is false."* What are the implications of this statement being true? What are the implications of it being false? Read http://en.wikipedia.org/wiki/Paradox.

11. What is the *inverse* of $a \Rightarrow b$?

    What is the *converse* of $a \Rightarrow b$?

    What is the *contrapositive* of $a \Rightarrow b$?

    Use truth tables to show that $a \Rightarrow b$ is *logically equivalent* to $\sim b \Rightarrow \sim a$ (not-*b* implies not-*a*).

Lesson 1 – Answers

Symbolic Logic

This lesson is an introduction to symbolic logic and what we actually mean in mathematics by statements such as "*a* implies *b*" and "*a* if and only if *b*." Below are some common symbols that are used in logic followed by the corresponding math symbols that I will use instead.

LOGIC	**MATH**
~ or ¬	not
$\vee$	or
$\wedge$	and
$a \to b$	$a \Rightarrow b$
$a \leftrightarrow b$	$a \Leftrightarrow b$

The statement "$a \Rightarrow b$" can be read as "*a implies b*" or "*if a then b*" or "*a is a sufficient condition for b*" or "*b is a necessary condition for a.*"

The statement "$a \Leftrightarrow b$" can be read or written as "*a iff b*" or "*a if and only if b*" or "*a is a necessary and sufficient condition for b.*"

Using the logical connectives above, we can rewrite "$a \Rightarrow b$" as "not (a & not-b)." Similarly, since "$a \Leftrightarrow b$" means "$a \Rightarrow b$ & $b \Rightarrow a$," we can rewrite "$a \Leftrightarrow b$" as "[not (a & not-b)] & [not (b & not-a)]."

In mathematics, for a compound statement *"A & B"* to be true, both of the statements *A* and *B* must be true. On the other hand, for the compound statement *"A or B"* to be true, only one of the statements must be true. Construct some simple examples to convince yourself that this is the correct way to proceed. Also, in mathematics, unless stated otherwise, we always use an *inclusive or*. That means that for *"A or B"* to be true, we either have *A* true or *B* true or both *A* and *B* true. In an *exclusive or*, either *A* or *B* can be true, but not both at the same time.

Now, here are some things for you to either do or look up and respond to.

1. What is a *truth table?*

   A truth table is a 2-dimensional array that indicates the final truth value (true or false) of a statement based upon the truth values assigned to its component propositions.

   See http://mathworld.wolfram.com/TruthTable.html

2. What is the *law of the excluded middle?*

The Law of the Excluded Middle assumes that for any proposition P either it or its negation is true and that there are no other possibilities. In other words, A or not-A. A proposition is assumed to be either true or false.

See http://mathworld.wolfram.com/LawoftheExcludedMiddle.html

3. What is a *tautology?*

A tautology is a statement that is always true regardless of whether its component propositions are true or false. Hence, in a truth table, the final truth value will always be "true." For example, consider the statement, "I will study math or I won't study math."

See http://mathworld.wolfram.com/Tautology.html

4. What is a *contradiction?*

A contradiction is a statement of the form (P & not-P). If we have an argument that implies both P and not-P, then the argument has led to a contradiction. Alternatively, in a truth table, the final truth value will always be "false."

See http://mathworld.wolfram.com/Contradiction.html

5. Complete the following *truth tables*. Your final values should be in the column shaded yellow.

Not	**A**
F	T
T	F

A	&	B
T	T	T
T	F	F
F	F	T
F	F	F

A	or	B
T	T	T
T	T	F
F	T	T
F	F	F

Not	[A	&	(Not	B)]
T	T	F	F	T
F	T	T	T	F
T	F	F	F	T
T	F	F	T	F

{ Not	[A	&	(Not	B)] }	&	{ Not	[B	&	(Not	A)] }
T	T	F	F	T	T	T	T	F	F	T
F	T	T	T	F	F	T	F	F	F	T
T	F	F	F	T	F	F	T	T	T	F
T	F	F	T	F	T	T	F	F	T	F

6. What is *modus ponens?* Give an example.

Modus ponens is a form of argument that essential says that if *F implies G* (i.e. $F \Rightarrow G$) and if *F* is true, then *G* is true. The steps are,

(1) $F \Rightarrow G$

(2) F

(3) Therefore, *G*.

Example:

If someone is a man, then they are mortal.

Socrates is a man.

Therefore, Socrates is mortal.

7. What is *modus tollens"* Give an example.

Modus tollens is a form of argument that essential says that if *F implies G* (i.e.

$F \Rightarrow G$) and if G is false, then F is false. The steps are,

(1) $F \Rightarrow G$

(2) not G

(3) Therefore, not F.

Example:

If I am at home, then I will study math.

I am not studying math.

Therefore, I am not at home.

8. The statement $A \Rightarrow B$ means "not (A & not B)." Use a truth table to explain why a false statement can imply anything? In other words, why is a statement such as *"If the moon is made of green cheese, then I am the smartest man on the planet"* considered to be true?

Below is the truth table for $A \Rightarrow B$ (i.e. not (A & not B)). From this table we can see that if A is false, then the final truth value of the statement is always "true." The implication is false only if A is true while B is false. Hence, a false statement can imply anything.

Not	[A	&	(Not	B)]
T	T	F	F	T
F	T	T	T	F
T	F	F	F	T
T	F	F	T	F

9. Use truth tables to show that "not (*a* & not-*b*)" and "not-*a* or b" are *logically equivalen*t. In this context, logical equivalence means that if *a* and *b* are assigned the same truth values, then the truth values of our final compound statements are the same.

   We have previously completed the truth table for "not (*a* & not-*b*)."

Not	[A	&	(Not	B)]
T	T	F	F	T
F	T	T	T	F
T	F	F	F	T
T	F	F	T	F

   We just need to complete the corresponding truth table for "not-*a* or b."

(Not	A)	or	B
F	T	T	T
F	T	F	F
T	F	T	T
T	F	T	F

From the results we can see that the same true/false assignments made to *A* and *B* result in the same final truth value for both "not (*a* & not-*b*)" and "not-*a* or b." Therefore, there are logically equivalent.

10. Consider the statement "*This sentence is false.*" What are the implications of this statement being true? What are the implications of it being false? Read http://en.wikipedia.org/wiki/Paradox.

In general, we often think of a paradox as something that is simply unusual or unexpected, but in particular, we tend to think of a mathematical paradox as a statement that implies its negation such as $A \Leftrightarrow$ not-A. This is the type of paradox we have above. If the sentence is true, then by definition, it is false, and if the sentence is false, then it must be true. Paradox! In many respects, I think of paradoxes as important because they reveal to us what may be either flaws or unexpected consequences of our logic. Sometimes they also seem to reveal the limitations of logical discourse. Furthermore, their presence suggests that sometimes

something will exist other than just “true” or “false.”

11. What is the *inverse* of $a \Rightarrow b$?

The *inverse of* $a \Rightarrow b$ is not-$a \Rightarrow$ not-b.

What is the *converse* of $a \Rightarrow b$?

The *converse of* $a \Rightarrow b$ is $b \Rightarrow a$.

What is the *contrapositive* of $a \Rightarrow b$?

The *contrapositive of* $a \Rightarrow b$ is not-$b \Rightarrow$ not-a.

Use truth tables to show that $a \Rightarrow b$ is *logically equivalent* to not-$b \Rightarrow$ not-a.

As stated previously, $a \Rightarrow b$ means “not (a & not-b),” and the truth table for this statement is below.

Not	**[A**	**&**	**(Not**	**B)]**
T	T	F	F	T
F	T	T	T	F
T	F	F	F	T
T	F	F	T	F

Similarly, not-$b \Rightarrow$ not-a means "not (not-b & not-not-a)" which we can simplify to "not (not-b & a)." The truth table for this statement is below

Not	**[(not**	**B)**	**&**	**A]**
T	F	T	F	T
F	T	F	T	T
T	F	T	F	F
T	T	F	F	F

From the above tables we can see that when *A* and *B* are given the same truth values in each table, then the truth values of the resulting statements are identical. Therefore, $a \Rightarrow b$ is *logically equivalent* to not-$b \Rightarrow$ not-a.

Lesson 2

Set Theory

In general, a mathematical proof is a convincing argument conforming to standard rules of logic that begins with a premise and ends with a conclusion. In the good ol' days (back in the seventies when I was young) there was a tendency to use as much mathematical shorthand notation as possible. In particular, two symbols from logic known, respectively, as the *universal quantifier* ($\forall$, *for every* …) and the *existential quantifier* ($\exists$, *there exists* …) were frequently employed as well as the symbol $\therefore$ for *therefore*. These days, however, there is a greater tendency to write proofs in plain English and to use the shorthand symbols a little more sparingly.

When you write a proof, think in terms of trying to write a convincing argument that a colleague could easily understand. This also means that when professional research mathematicians are writing proofs to be read by other researchers, they can be very brief in their arguments. On the other hand, when one is writing a proof for someone with less training in formal mathematics, a little more verbosity is often needed in order to make the argument convincing.

Each branch of mathematics tends to have its own style and technique of doing proofs. In particular, in set theory if one is trying to show that for two sets A and B that $A = B$, then one generally utilizes the following: "THEOREM: If A and B are sets such that

$A \subseteq B$ and $B \subseteq A$, then $A = B$." Hence, proofs involving the equality of two sets *A* and *B* generally take the following form:

PROOF: Let $x \in A$. … Thus, $x \in B$ and, hence, $A \subseteq B$. Now let $x \in B$. … Thus, $x \in A$ and, hence, $B \subseteq A$. Therefore, $A = B$. □

(NOTE: Mathematicians used to end their proofs with the letters *QED* which stands for *quod erat demonstandum*, which means *that which was to be demonstrated*. However, a twentieth century mathematician named Paul Halmos felt it was a little presumptuous to always assume that one's proof was correct, and he introduced the practice of using a square (usually shaded) to indicate the end of a proof.)

Before we continue, here are a few basic definitions regarding set notation.

$\in$ - is an element of

U - The *universal set*. Whatever our universe of discourse is, i.e. real numbers, complex numbers, etc.

$\varnothing$ - The *null* or *empty set*. The set containing no elements.

$A \cup B = \{x \in U \mid x \in A \text{ or } x \in B\}$ (This is read as "*A* union *B*.")

$A \cap B = \{x \in U \mid x \in A \text{ and } x \in B\}$ (This is read as "*A* intersect *B*.")

$A' = \{x \in U \mid x \notin A\}$ (This is read as "*A*-complement.")

$A \subseteq B$ if and only if $\forall x \in A, x \in B$ (This is read as "*A* is a subset of *B*.")

$A \subset B$ if and only if $\forall x \in A, x \in B$ and $\exists y \in B$ such that $y \notin A$ (This is read as "*A* is a proper subset of *B*.")

The cardinality of a set *A* is the number of elements in *A*, and this is denoted by $|A|$. For example, if $A = \{a,b,c\}$, then $|A| = 3$. Also, $|\varnothing| = 0$.

1. Prove De Morgan's Laws.
    a. PROVE: $(A \cup B)' = A' \cap B'$.
    b. PROVE: $(A \cap B)' = A' \cup B'$

2. PROVE: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

3. Explain why the *null set* is a subset of every set.

4. List the subsets of the following sets: $\varnothing, \{\varnothing\}, \{a\}.\{a,b\}, \{a,b,c\}$. Do you see a pattern with respect to the number of subsets?

5. Who was Georg Cantor? How did he live? How did he die?

6. Explain *Russell's Paradox*. What does it tell us about set theory? How do mathematicians "weasel out" of this paradox?

Lesson 2 – Answers

Set Theory

In general, a mathematical proof is a convincing argument conforming to standard rules of logic that begins with a premise and ends with a conclusion. In the good ol' days (back in the seventies when I was young) there was a tendency to use as much mathematical shorthand notation as possible. In particular, two symbols from logic known, respectively, as the *universal quantifier* ($\forall$, *for every* …) and the *existential quantifier* ($\exists$, *there exists* …) were frequently employed as well as the symbol $\therefore$ for *therefore*. These days, however, there is a greater tendency to write proofs in plain English and to use the shorthand symbols a little more sparingly.

When you write a proof, think in terms of trying to write a convincing argument that a colleague could easily understand. This also means that when professional research mathematicians are writing proofs to be read by other researchers, they can be very brief in their arguments. On the other hand, when one is writing a proof for someone with less training in formal mathematics, a little more verbosity is often needed in order to make the argument convincing.

Each branch of mathematics tends to have its own style and technique of doing proofs. In particular, in set theory if one is trying to show that for two sets A and B that $A = B$, then one generally utilizes the following: "THEOREM: If A and B are sets such that

$A \subseteq B$ and $B \subseteq A$, then $A = B$." Hence, proofs involving the equality of two sets *A* and *B* generally take the following form:

> PROOF: Let $x \in A$. … Thus, $x \in B$ and, hence, $A \subseteq B$. Now let $x \in B$. Thus, $x \in A$ and, hence, $B \subseteq A$. Therefore, $A = B$. □
>
> (NOTE: Mathematicians used to end their proofs with the letters *QED* which stands for *quod erat demonstandum*, which means *that which was to be demonstrated*. However, a twentieth century mathematician named Paul Halmos felt it was a little presumptuous to always assume that one's proof was correct, and he introduced the practice of using a square (usually shaded) to indicate the end of a proof.)

Before we continue, here are a few basic definitions regarding set notation.

$\in$ - is an element of

U - The *universal set*. Whatever our universe of discourse is, i.e. real numbers, complex numbers, etc.

$\varnothing$ - The *null* or *empty set*. The set containing no elements.

$A \cup B = \{x \in U \mid x \in A \text{ or } x \in B\}$ (This is read as "*A* union *B*.")

$A \cap B = \{x \in U \mid x \in A \text{ and } x \in B\}$ (This is read as "*A* intersect *B*.")

$A' = \{x \in U \mid x \notin A\}$ (This is read as "*A*-complement.")

$A \subseteq B$ if and only if $\forall x \in A, x \in B$ (This is read as "*A* is a subset of *B*.")

$A \subset B$ if and only if $\forall x \in A, x \in B$ and $\exists y \in B$ such that $y \notin A$ (This is read as "*A* is a proper subset of *B*.")

The cardinality of a set *A* is the number of elements in *A*, and this is denoted by $|A|$. For example, if $A = \{a, b, c\}$, then $|A| = 3$. Also, $|\varnothing| = 0$.

1. Prove De Morgan's Laws.

   a. PROVE: $(A \cup B)' = A' \cap B'$.

   PROOF: Let $x \in (A \cup B)'$. Then $x \notin A$ and $x \notin B$ (since, otherwise, we would have $x \in (A \cup B)$). Thus, if $x \notin A$ and $x \notin B$, then $x \in A'$ and $x \in B'$ which implies that $x \in A' \cap B'$. Therefore, $(A \cup B)' \subseteq A' \cap B'$.

   Now suppose that $x \in A' \cap B'$. Then $x \in A'$ and $x \in B'$ which implies that $x \notin A$ and $x \notin B$, and, hence, $x \notin A \cup B$. Therefore, $x \in (A \cup B)'$, and, thus, $A' \cap B' \subseteq (A \cup B)'$. Furthermore, since $(A \cup B)' \subseteq A' \cap B'$ and $A' \cap B' \subseteq (A \cup B)'$, it now follows that $(A \cup B)' = A' \cap B'$. □

   b. PROVE: $(A \cap B)' = A' \cup B'$

   PROOF: Let $x \in (A \cap B)'$. Then $x \notin A$ or $x \notin B$ (since, otherwise, we would have $x \in (A \cap B)$). Thus, if $x \notin A$ or $x \notin B$, then $x \in A'$ or $x \in B'$ which implies that $x \in A' \cup B'$. Therefore, $(A \cap B)' \subseteq A' \cup B'$.

Now suppose that $x \in A' \cup B'$. Then $x \in A'$ or $x \in B'$ which implies that $x \notin A$ or $x \notin B$, and, hence, $x \notin A \cap B$. Therefore, $x \in (A \cap B)'$, and, thus, $A' \cup B' \subseteq (A \cap B)'$. Furthermore, since $(A \cap B)' \subseteq A' \cup B'$ and $A' \cup B' \subseteq (A \cap B)'$, it now follows that $(A \cap B)' = A' \cup B'$. □

2. PROVE: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$

PROOF: Suppose $x \in (A \cup B) \cap C$. Then $x \in A \cup B$ and $x \in C$. However, if $x \in A \cup B$, then $x \in A$ or $x \in B$. If $x \in A$, then $x \in A \cap C$, and if $x \in B$, then $x \in B \cap C$. Thus, one way or another, we have that $x \in (A \cap C) \cup (B \cap C)$ and, hence, $(A \cup B) \cap C \subseteq (A \cap C) \cup (B \cap C)$.

Now suppose that $x \in (A \cap C) \cup (B \cap C)$. Then $x \in A \cap C$ or $x \in B \cap C$, and, hence, it follows that $x \in C$. It also follows that $x \in A$ or $x \in B$, and thus, $x \in A \cup B$. Consequently, $x \in (A \cup B) \cap C$, and, hence, $(A \cap C) \cup (B \cap C) \subseteq (A \cup B) \cap C$.

$\therefore (A \cup B) \cap C = (A \cap C) \cup (B \cap C)$. □

3. Explain why the *null set* is a subset of every set.

By definition, the *null set* is the set containing no objects. Also, $A \subseteq B$ means that if

$x \in A$, then $x \in B$. If we replace A by $\varnothing$, then we can rewrite this condition as $\varnothing \subseteq B$ means that if $x \in \varnothing$ then $x \in B$ (or in other words, $x \in \varnothing \Rightarrow x \in B$). However, the statement $x \in \varnothing$ is always false, and recall that a false statement can imply anything. Hence, it is true that $x \in \varnothing \Rightarrow x \in B$, and, therefore, $\varnothing \subseteq B$ for any set B.

4. List the subsets of the following sets: $\varnothing, \{a\}, \{a,b\}, \{a,b,c\}$. Do you see a pattern with respect to the number of subsets?

We'll denote the set of all subset of a set A by $P(A)$. Consequently,

$P(\varnothing) = \{\varnothing\}$

$P(\{a\}) = \{\varnothing, \{a\}\}$

$P(\{a,b\}) = \{\varnothing, \{a\}, \{b\}, \{a,b\}\}$

$P(\{a,b,c\}) = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\}$

If we now denote the number of elements in a set A by $|A|$, then

$|P(\varnothing)| = 2^0 = 1$

$|P(\{a\})| = 2^1 = 2$

$|P(\{a,b\})| = 2^2 = 4$

$|P(\{a,b,c\})| = 2^3 = 8$

These results correctly suggest that if a finite set A has n elements, then $|P(A)| = 2^n$.

5. Who was Georg Cantor? How did he live? How did he die?

George Cantor was born in 1845 and died in 1918, and he is best remembered as the creator of set theory. While today we consider set theory to be the foundation for all of mathematics, in his day Cantor and his work were the subject of several vicious attacks from not only his former teacher, Leopold Kronecker, but also other leading mathematicians of the day. This is because Cantor's work included the existence of infinities of different sizes as well as various other items that had not been a part of mathematics up to that point. For example, prior to Cantor, infinity was a subject to be eschewed in mathematics as is evidenced by the following quote from mathematician Carl Friedrich Gauss:

"I protest against the use of infinite magnitude as something completed, which is never permissible in mathematics. Infinity is merely a way of speaking, the true meaning being a limit which certain ratios approach indefinitely close, while others are permitted to increase without restriction."

Because his ideas were so radical at the time, his former professor, Kronecker, actively worked to block the publication of his papers and appointments to more prestigious institutions. These attacks on Cantor appear to have eventually taken their toll, and Cantor suffered his first hospitalization for chronic depression in 1884. He

was to be hospitalized many more times in his life for this mental disorder including the last year of his life which he spent in a sanatorium in Halle, Germany. In spite of the opposition generated during Cantor's lifetime, his theories eventually gained wide acceptance throughout mathematics, and in the first half of the twentieth century, the great mathematician David Hilbert declared, "No one shall expel us from the Paradise that Cantor has created."

6. Explain *Russell's Paradox*. What does it tell us about set theory? How do mathematicians "weasel out" of this paradox?

At its beginning, the word *set* seemed to be just another word for *collection*, and surely we can talk about the collection of anything we wish. Or so it seemed. This point of view that we naturally understand what a set is without needing any formal axiomatic structure is now referred to as *naive set theory*, and its limitations were discovered in 1901 when Bertrand Russell formulated what we now call *Russell's Paradox*. The heart of the paradox is that, intuitively, a set *R* can either be a member or element of itself or, alternatively, not an element of itself. Most of time we experience the latter. For example, if $A = \{1, 2\}$, then $1 \in A$ and $2 \in A$, but $A \notin A$ (even though $A \subseteq A$). However, on the other hand, if I say, "Let *A* be the set of all sets that I can describe with a finite number of words," then it seems like I have described the set *A* itself with a finite number of words, and, hence, $A \in A$. In this spirit, Bertrand Russell asked us to consider the set *R* defined as the set of all sets that do not contain

themselves as members. In other words, $R = \{A \mid A \notin A\}$. We now ask ourselves the question, "Is $R \in R$?" If $R \in R$, then by definition, $R \notin R$, and if $R \notin R$, then again, by definition, $R \in R$. Thus, we arrive at the following paradox, $R \in R \Leftrightarrow R \notin R$.

Russell's Paradox showed us that, if we wanted to avoid contradictions, we had to be much more careful about what we did and what we didn't call a *set*. The ultimate result, for most mathematicians, was the creation of an axiomatic version of set theory known as *Zermelo-Frankel*. This version of set theory allows infinite sets of all sorts to exist, but, at the same time, it doesn't allow us to talk about things that are so large or so self-referential that they lead us into paradoxes. Furthermore, in this axiomatic formulation, the term *set* is left undefined, and those collections that we don't believe as deserving to be called sets within the framework of *Zermelo-Frankel* we simply label as *classes*. Thus, the collection *R* defined above is simply referred to as a *class* instead of a *set*, and the paradox is avoided. Nonetheless, in my opinion, we haven't really resolved the paradox. We've simply decided to ignore an inconvenient truth and, instead, sweep it under the rug. Furthermore, we should note that paradoxes, in general, are quite interesting because they show us where our logical framework for reality tends to break down. More paradoxes to come!

Lesson 3

Set Theory Continued

In set theory, the size or number of elements in a set is called its *cardinality*. There are various symbols that are used for *cardinality*, but my favorite is to simply enclose the set inside a pair of absolute value signs. Thus, if $A=\{a,b,c\}$, then $|A|=3$ since the set has three elements. We can show that two sets have the same *cardinality* by finding a function that establishes a one-to-one correspondence between the elements of the sets. A function $f:A\to B$ is *one-to-one* if $\forall x,y\in A,\ x\neq y\Rightarrow f(x)\neq f(y)$. We also call a *one-to-one* function an *injection* or *injective function*. A function $f:A\to B$ is *onto* if $\forall y\in B,\ \exists x\in A$ such that $f(x)=y$. We also call an *onto* function a *surjection* or *surjective function*. We can now define a *one-to-one correspondence* between two sets *A* and *B* as a function $f:A\to B$ that is both *one-to-one* and *onto*. We also call a *one-to-one and onto* function a *bijection* or *bijective function*.

When we are dealing with *cardinality* or size of sets, everything behaves as we expect when the sets are finite. However, if our sets are infinite, then strange things can happen. For example, one set can be a proper subset of another, and yet the two sets can be the same size ($A\subset B$ & $|A|=|B|$). Additionally, some infinite sets can have more elements in them than other infinite sets (As you'll prove below, if $\mathbb{N}=$ counting or natural numbers and $\mathbb{R}=$ real numbers, then $|\mathbb{N}|<|\mathbb{R}|$.). Note that if there is an *injective function* $f:A\to B$, but no *surjective function* $f:A\to B$, then we'll say that $|A|<|B|$.

1. If a set *A* has the same cardinality as the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$, then we say that *A* is *countable*. Prove that the set of even natural numbers, $2\mathbb{N} = \{2, 4, 6, \ldots\}$, is *countable* by finding a *bijective function* $f : \mathbb{N} \to 2\mathbb{N}$. Conclude that there are just as many even natural numbers as natural numbers.

2. Prove that there are just as many numbers in the interval $(0,1)$ as there are real numbers by finding a *bijective function* $f : (0,1) \to \mathbb{R}$. (HINT: You can find a *bijection* by modifying a well-known trigonometric function.)

3. Study *Cantor's Diagonal Theorem* and use the argument to prove that there is no bijection $f : \mathbb{N} \to (0,1)$.

4. *Cantor's Diagonal Theorem* is an example of *proof by contradiction*. In other words, we make an assumption, prove that that assumption leads to a contradiction, and then we conclude the opposite of our assumption. However, some mathematicians don't like proofs done by this method. Review the concepts introduced in Lesson 1, and explain why some people don't like this method.

5. Conclude from 2 & 3 above that $|\mathbb{N}| < |\mathbb{R}|$.

Lesson 3 – Answers

Set Theory Continued

In set theory, the size or number of elements in a set is called its *cardinality*. There are various symbols that are used for *cardinality*, but my favorite is to simply enclose the set inside a pair of absolute value signs. Thus, if $A=\{a,b,c\}$, then $|A|=3$ since the set has three elements. We can show that two sets have the same *cardinality* by finding a function that establishes a one-to-one correspondence between the elements of the sets. A function $f:A\to B$ is *one-to-one* if $\forall x,y\in A,\ x\neq y\Rightarrow f(x)\neq f(y)$. We also call a *one-to-one* function an *injection* or *injective function*. A function $f:A\to B$ is *onto* if $\forall y\in B,\ \exists x\in A$ such that $f(x)=y$. We also call an *onto* function a *surjection* or *surjective function*. We can now define a *one-to-one correspondence* between two sets *A* and *B* as a function $f:A\to B$ that is both *one-to-one* and *onto*. We also call a *one-to-one and onto* function a *bijection* or *bijective function*.

When we are dealing with *cardinality* or size of sets, everything behaves as we expect when the sets are finite. However, if our sets are infinite, then strange things can happen. For example, one set can be a proper subset of another, and yet the two sets can be the same size ($A\subset B$ & $|A|=|B|$). Additionally, some infinite sets can have more elements in them than other infinite sets (As you'll prove below, if $\mathbb{N}=$ counting or natural numbers and $\mathbb{R}=$ real numbers, then $|\mathbb{N}|<|\mathbb{R}|$.). Note that if there is an *injective function* $f:A\to B$, but no *surjective function* $f:A\to B$, then we'll say that $|A|<|B|$.

1. If a set *A* has the same cardinality as the natural numbers $\mathbb{N} = \{1, 2, 3, \ldots\}$, then we say that *A* is *countable*. Prove that the set of even natural numbers, $2\mathbb{N} = \{2, 4, 6, \ldots\}$, is *countable* by finding a *bijective function* $f : \mathbb{N} \to 2\mathbb{N}$. Conclude that there are just as many even natural numbers as natural numbers.

   Proof: Define $f : \mathbb{N} \to 2\mathbb{N}$ by $f(x) = 2x$. Then clearly $f : \mathbb{N} \to 2\mathbb{N}$ is one-to-one since $x_1 \neq x_2 \Rightarrow 2x_1 \neq 2x_2$. Also, $f : \mathbb{N} \to 2\mathbb{N}$ is onto since if $y \in 2\mathbb{N}$, then *y* is even which means that $y/2 \in \mathbb{N}$, and, hence, $f(y/2) = 2 \cdot y/2 = y$. Therefore, $f : \mathbb{N} \to 2\mathbb{N}$ is a bijection, and $|\mathbb{N}| = |2\mathbb{N}|$. □

2. Prove that there are just as many numbers in the interval $(0,1)$ as there are real numbers by finding a *bijective function* $f : (0,1) \to \mathbb{R}$. (HINT: You can find a *bijection* by modifying a well-known trigonometric function.)

   Proof: Clearly, $g : (-\pi/2, \pi/2) \to \mathbb{R}$ defined by $g(x) = \tan x$ is a bijection. Thus, all we need to do is transform this function into a related function with domain $(0,1)$. If we replace $\tan x$ by $\tan \pi x$, then this will shift our interval $(-\pi/2, \pi/2)$ to $(-1/2, 1/2)$. Next, if we replace *x* by $x - 1/2$, then that will shift $(-1/2, 1/2)$ to $(0,1)$. Thus, $f : (0,1) \to \mathbb{R}$ defined by $f(x) = \tan \pi(x - 1/2)$ is a bijective function from $(0,1)$ to $\mathbb{R}$. Therefore, $|(0,1)| = |\mathbb{R}|$. □

3. Study *Cantor's Diagonal Theorem* and use the argument to prove that there is no bijection $f:\mathbb{N}\to(0,1)$.

Prove: There is no bijection $f:\mathbb{N}\to(0,1)$.

Proof: By way of contradiction, suppose $f:\mathbb{N}\to(0,1)$ is a bijective function. Then the correspondence between elements of $\mathbb{N}$ and elements of $(0,1)$ could be written as a list like the following.

$$\begin{aligned}
1 &\to 0.1451939\ldots\\
2 &\to 0.5876624\ldots\\
3 &\to 0.9942146\ldots\\
4 &\to 0.3621722\ldots\\
&\vdots
\end{aligned}$$

Now, highlight the numbers along the diagonal from upper left on down.

$$\begin{aligned}
1 &\to 0.\mathbf{1}451939\ldots\\
2 &\to 0.5\mathbf{8}76624\ldots\\
3 &\to 0.99\mathbf{4}2146\ldots\\
4 &\to 0.362\mathbf{1}722\ldots\\
&\vdots \qquad\qquad \ddots
\end{aligned}$$

If we now generate a new number by changing each odd digit in the highlighted number to 2 and each even digit to 1, then our new number will belong to the interval $(0,1)$, but our function will have no natural number assigned to it since the *nth* digit

of the decimal expansion will always differ from that of the number in our list that has been assigned to the natural number *n*. In other words, using the list above we can begin to generate the number $0.2112\ldots$. This number, however, doesn't correspond to 1 because the first digit is different from what's in our list. Similarly, it doesn't correspond to 2, since the second digit is different from what's in our list, and so on and so on. Thus, given any table or list for our function, we can always construct a real number in the interval $(0,1)$ that has no natural number assigned to it. Therefore, there is no bijection $f:\mathbb{N}\to(0,1)$. □

4. *Cantor's Diagonal Theorem* is an example of *proof by contradiction*. In other words, we make an assumption, prove that that assumption leads to a contradiction, and then we conclude the opposite of our assumption. However, some mathematicians don't like proofs done by this method. Review the concepts introduced in Lesson 1, and explain why some people don't like this method.

   When we do a proof by contradiction, we are assuming that either *A* is true or *not-A* is true. In other words, we are assuming the Law of the Excluded Middle, and this makes some mathematicians uneasy. Hence, direct proofs that do not require this assumption are generally considered superior to proofs by contradiction.

5. Conclude from 2 & 3 above that $|\mathbb{N}|<|\mathbb{R}|$.

From 3 above, we know that $|\mathbb{N}| < |(0,1)|$, and from 2 above we know that $|(0,1)| = |\mathbb{R}|$.

Therefore, $|\mathbb{N}| < |\mathbb{R}|$.

Lesson 4

Mathematical Induction

Mathematical induction is a standard proof technique for showing that some proposition *P* about natural numbers holds true for all $n \in \mathbb{N}$.

Mathematical Induction: If *P* is a proposition about natural numbers $n \in \mathbb{N}$, then *P* is true for all $n \in \mathbb{N}$ if,

1. *P* is true for $n = 1$, and
2. *P* true for $n \in \mathbb{N} \Rightarrow$ *P* is true for $n + 1 \in \mathbb{N}$.

There are several variations we could do of this basic principle. For example, if we began by showing that *P* is true for $n = 0$, then we could possibly prove that *P* is true for all whole numbers. Similarly, if we started our argument by showing that *P* is true for $n = 10$, then a successful induction argument could show that *P* is true for all natural numbers greater than or equal to 10. Another variant form of mathematical induction is shown below.

The Second Principle of Mathematical Induction: If *P* is a proposition about natural numbers $n \in \mathbb{N}$, then *P* is true for all $n \in \mathbb{N}$ if,

3. *P* is true for $n = 1$, and
4. *P* true for all natural numbers less than $n \in \mathbb{N} \Rightarrow$ *P* is true for $n \in \mathbb{N}$.

1. Use mathematical induction to prove that $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

2. Use mathematical induction to prove that $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

3. Find the flaw in the following inductive argument that all horses are the same color.

   By way of induction, suppose that you have a set containing $n=1$ horses. Then clearly all the horses in that set are the same color. Now assume that it is true that in any set of n horses, all the horses have the same color (our induction hypothesis). At this point we want to argue that it is also true that any set of $n+1$ horses will also all be the same color. Thus, suppose we are given a set containing $n+1$ horses. If we remove one horse, then by our inductive hypothesis the remaining n horses will all be the same color. Now return the horse we originally removed and remove a different horse. Then once again our inductive hypothesis states that the resulting set of n horses all have the same color. From this it follows that the two horses we successively removed have the same color, and therefore, all of the horses in our set of $n+1$ horses have the same color. It now follows by mathematical induction that for any set of n horses, $n \in \mathbb{N}$, all the horses have the same color.

4. If *A* is a set, then the set of all subsets of *A*, denoted by $P(A)$, is called the power set of *A*. For example, if $A = \varnothing$, then $P(A) = \{\varnothing\}$. If $A = \{a\}$, then $P(A) = \{\varnothing, \{a\}\}$. And if $A = \{a, b\}$, then $P(A) = \{\varnothing, \{a\}, \{b\}, \{a, b\}\}$. Use mathematical induction to show that

if $|A| = n \in \mathbb{N}$, then $|P(A)| = 2^n = 2^{|A|}$.

5. The result from the previous problem not only shows why we call $P(A)$ the power set of *A*, but also that for any finite set *A*, $|P(A)| > |A|$. This last result can be extended to infinite sets as well, and this provides a technique for constructing an infinite number of infinite sets of different sizes. In other words, for any infinite set, the cardinality of its power set will be greater than the cardinality of the original set. The smallest infinite set is represented by the set of counting or natural numbers, and we denote the size of this set by $\aleph_0$ (aleph null). Any set of size $\aleph_0$ is called countable or countably infinite. Larger infinite cardinal numbers are denoted by $\aleph_1, \aleph_2, \aleph_3 \ldots$ and so on. Below is a sloppy proof of mine that for any set *A*, there is no bijection from *A* to $P(A)$. This shows that the two sets have different cardinalities. However, since we can easily find an injective function from *A* to $P(A)$, (for instance, if $a \in A$, then pair *a* with $\{a\} \in P(A)$), it immediately follows that $|P(A)| > |A|$ for any set *A*. Clean up this proof.

Theorem: Let *A* be a set and let $P(A)$ be the set of all subsets of *A*. Then there is no bijective function from *A* to $P(A)$, and hence, $|P(A)| > |A|$.

Sloppy Proof: Let *A* be a set and let $P(A)$ be the set of all subsets of *A*. Since the result is obvious when *A* is empty, assume *A* is non-empty. Now assume that *f* is a

bijection from A to $P(A)$, and let T be the set of all elements x in A such that x is not an element of $f(x)$. Since f is a bijection, there exists an element t in A such that $f(t) = T$. Now ponder the question is t an element of T? Bummer. Therefore, no bijection exists from A to $P(A)$, and thus, $|P(A)| > |A|$. □

6. Georg Cantor contemplated the set of all sets, but his discovery of the theorem presented in exercise 5 led to a contradiction known as *Cantor's Paradox*. Give an informal discussion of why if U is the set of all sets, then we can reach both the conclusion that $|P(U)| > |U|$ and $|P(U)| \leq |U|$.

Lesson 4 – Answers

Mathematical Induction

Mathematical induction is a standard proof technique for showing that some proposition P about natural numbers holds true for all $n \in \mathbb{N}$.

Mathematical Induction: If P is a proposition about natural numbers $n \in \mathbb{N}$, then P is true for all $n \in \mathbb{N}$ if,

1. P is true for $n = 1$, and
2. P true for $n \in \mathbb{N} \Rightarrow P$ is true for $n + 1 \in \mathbb{N}$.

There are several variations we could do of this basic principle. For example, if we began by showing that P is true for $n = 0$, then we could possibly prove that P is true for all whole numbers. Similarly, if we started our argument by showing that P is true for $n = 10$, then a successful induction argument could show that P is true for all natural numbers greater than or equal to 10. Another variant form of mathematical induction is shown below.

The Second Principle of Mathematical Induction: If P is a proposition about natural numbers $n \in \mathbb{N}$, then P is true for all $n \in \mathbb{N}$ if,

3. P is true for $n = 1$, and
4. P true for all natural numbers less than $n \in \mathbb{N} \Rightarrow P$ is true for $n \in \mathbb{N}$.

1. Use mathematical induction to prove that $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

Proof: Let $n=1$. Then $\frac{1(1+1)}{2} = \frac{2}{2} = 1 = \sum_{k=1}^{1} k$. Hence, the statement is true for $n=1$.

Assume now that the statement is true for some natural number *n*, and consider if it is true for $n+1$. Clearly,

$$\sum_{k=1}^{n+1} k = \left(\sum_{k=1}^{n} k\right) + n + 1 = \frac{n(n+1)}{2} + n + 1 = \frac{n(n+1)+2(n+1)}{2} = \frac{(n+1)(n+2)}{2} = \frac{(n+1)\left[(n+1)+1\right]}{2}.$$

Hence, if the formula is true for *n*, then it is also true for $n+1$. Therefore, by mathematical induction, $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ for all natural numbers *n*. □

2. Use mathematical induction to prove that $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

Proof: Let $n=1$. Then $\frac{1(1+1)(2+1)}{6} = \frac{6}{6} = 1 = \sum_{k=1}^{1} k^2$. Hence, the statement is true for $n=1$. Assume now that the statement is true for some natural number *n*, and consider if it is true for $n+1$. Clearly,

$$\sum_{k=1}^{n+1} k^2 = \left(\sum_{k=1}^{n} k^2\right) + (n+1)^2 = \frac{n(n+1)(2n+1)}{6} + \frac{6(n+1)^2}{6} = \frac{(n+1)\left[n(2n+1)+6(n+1)\right]}{6}$$

$$= \frac{(n+1)\left[2n^2+7n+6\right]}{6} = \frac{(n+1)(n+2)(2n+3)}{6} = \frac{(n+1)\left[(n+1)+1\right]\left[2(n+1)+1\right]}{6}.$$

Hence, if the formula is true for *n*, then it is also true for $n+1$. Therefore, by

mathematical induction, $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$ for all natural numbers *n*. □

3. Find the flaw in the following inductive argument that all horses are the same color.

By way of induction, suppose that you have a set containing $n=1$ *horses. Then clearly all the horses in that set are the same color. Now assume that it is true that in any set of n horses, all the horses have the same color (our induction hypothesis). At this point we want to argue that it is also true that any set of* $n+1$ *horses will also all be the same color. Thus, suppose we are given a set containing* $n+1$ *horses. If we remove one horse, then by our inductive hypothesis the remaining n horses will all be the same color. Now return the horse we originally removed and remove a different horse. Then once again our inductive hypothesis states that the resulting set of n horses all have the same color. From this it follows that the two horses we successively removed have the same color, and therefore, all of the horses in our set of* $n+1$ *horses have the same color. It now follows by mathematical induction that for any set of n horses,* $n \in \mathbb{N}$*, all the horses have the same color.*

In the reading of the above argument, one often imagines a case where we might have, for example, 10 horses. We remove one horse, and then our induction hypothesis says that the remaining 9 horses are all the same color. We then replace our first horse, remove another horse, and again our induction hypothesis says that the remaining 9 horses are all the same color. And then finally, we conclude that because of the overlap of the two situations that all 10 horses are the same color. It is, indeed,

clear that the induction argument works for the case of $n=10$. However, where the argument breaks down is for $n=2$. When we have 2 horses, then we can remove either one, but the resulting singleton sets this time have no intersection or overlap, and thus, we can't conclude that the two horses have to be of the same color. This is the one break in the chain of the induction argument that at first glance would appear to prove the assertion true for all natural numbers *n*.

4. If *A* is a set, then the set of all subsets of *A*, denoted by $P(A)$, is called the power set of *A*. For example, if $A=\varnothing$, then $P(A)=\{\varnothing\}$. If $A=\{a\}$, then $P(A)=\{\varnothing,\{a\}\}$. And if $A=\{a,b\}$, then $P(A)=\{\varnothing,\{a\},\{b\},\{a,b\}\}$. Use mathematical induction to show that if $|A|=n\in\mathbb{N}$, then $|P(A)|=2^n=2^{|A|}$.

Proof: Our assertion is actually true for all whole numbers since as we see above, $|\varnothing|=0$ and $P(\varnothing)=\{\varnothing\}\Rightarrow|P(\varnothing)|=1=2^0$. Similarly, if $A=\{a\}$, then $P(A)=\{\varnothing,\{a\}\}$ and $|P(A)|=2^1=2$. Thus, let's assume that for any set with cardinality *n* that the cardinality of its power set is 2^n, and let's suppose that we have a set *A* such that $|A|=n+1\geq 2$. Then there exists $a\in A$ such that if we remove *a* from *A*, then $|A-\{a\}|=n$, and hence, by our induction hypothesis, $|P(A-\{a\})|=2^n$. Now consider the structure of $P(A)$. Clearly, every set in $P(A-\{a\})$ also belongs to $P(A)$. Furthermore, we can divide the sets in $P(A)$ into two categories, those that contain *a* and those that don't. A moment's reflection should convince us that *A* should have just as many subsets that contain *a* as don't. For example, we can take any subset that

doesn't contain *a* and create one that contains *a* just by adding *a* to it. Similarly, we could start with any subset containing *a* and delete *a* from this subset to obtain one that is lacking *a*. Thus, in $P(A)$ there is a one-to-one correspondence between those subsets that contain *a* and those that don't, and from this it follows that $|P(A)| = 2 \cdot |P(A) - \{a\}| = 2 \cdot 2^n = 2^{n+1} = 2^{|A|}$. Therefore, the assertion is true for all natural numbers, and, indeed, all whole numbers. □

5. The result from the previous problem not only shows why we call $P(A)$ the power set of *A*, but also that for any finite set *A*, $|P(A)| > |A|$. This last result can be extended to infinite sets as well, and this provides a technique for constructing an infinite number of infinite sets of different sizes. In other words, for any infinite set, the cardinality of its power set will be greater than the cardinality of the original set. The smallest infinite set is represented by the set of counting or natural numbers, and we denote the size of this set by $\aleph_0$ (aleph null). Any set of size $\aleph_0$ is called countable or countably infinite. Larger infinite cardinal numbers are denoted by $\aleph_1, \aleph_2, \aleph_3 \ldots$ and so on. Below is a sloppy proof of mine that for any set *A*, there is no bijection from *A* to $P(A)$. This shows that the two sets have different cardinalities. However, since we can easily find an injective function from *A* to $P(A)$, (for instance, if $a \in A$, then pair *a* with $\{a\} \in P(A)$), it immediately follows that $|P(A)| > |A|$ for any set *A*. Clean up this proof.

Theorem: Let A be a set and let $P(A)$ be the set of all subsets of A. Then there is no bijective function from A to $P(A)$, and hence, $|P(A)| > |A|$.

Sloppy Proof: Let A be a set and let $P(A)$ be the set of all subsets of A. Since the result is obvious when A is empty, assume A is non-empty. Now assume that f is a bijection from A to $P(A)$, and let T be the set of all elements x in A such that x is not an element of $f(x)$. Since f is a bijection, there exists an element t in A such that $f(t) = T$. Now ponder the question is t an element of T? Bummer. Therefore, no bijection exists from A to $P(A)$, and thus, $|P(A)| > |A|$. □

Proof: Let A be a set and let $P(A)$ be the set of all subsets of A. Since the result is obvious when A is empty, assume A is non-empty. Now assume that f is a bijection from A to $P(A)$, and let T be the set of all elements x in A such that x is not an element of $f(x)$. Since f is a bijection, there exists an element t in A such that $f(t) = T$. Now ponder the question is t an element of T? If $t \in T$, then since T is defined as the set of all elements x in A such that x is not an element of $f(x)$, it follows that $t \notin T$. But on the other hand, if $t \notin T$, then it follows from the definition of T that $t \in T$. Either way we go, we arrive at a contradiction, and the source of these contradictions is the assumption that we have a bijective function f from A to $P(A)$. Hence, no such bijection can exist, and so $|P(A)| \neq |A|$. On the other hand, the function $f : A \to P(A)$ defined by $f(a) = \{a\}$ is clearly an injection, and therefore, $|P(A)| > |A|$. □

6. Georg Cantor contemplated the set of all sets, but his discovery of the theorem presented in exercise 5 led to a contradiction known as *Cantor's Paradox*. Give an informal discussion of why if U is the set of all sets, then we can reach both the conclusion that $|P(U)| > |U|$ and $|P(U)| \leq |U|$.

   By the theorem proved in exercise 5, we know that $|P(U)| > |U|$. But on the other hand, if U is the set of all sets, then it must contain every element in $P(U)$ and that means that there is an obvious injection from $P(U) \to U$. Hence, it must also be true that $|P(U)| \leq |U|$.

   Naively, we think of a set as a collection of objects, but paradoxes such as Russell's Paradox and Cantor's Paradox show us that we have to be more careful about what we are allowed to call a set. This has resulted in axiomatic versions of set theory with the most popular one being known as *Zermelo-Frankel-Axiom of Choice* or *ZFC*. A fairly non-technical version of the axioms for set theory is given below:

   a. Two sets are identical if they have the same elements.
   b. The empty set exists.
   c. If A and B are sets, then $\{A, B\}$ is a set.
   d. The union of a set of sets is a set.
   e. Infinite sets exist.
   f. A property that can be formalized in the language of the theory can be used to define a set.
   g. The power set of a set is a set.

h. If a set has an element in it, then we can "choose" that element. (Axiom of Choice)

i. If A is a set, then $A \notin A$.

The above axioms help us avoid the kinds of paradoxes and conundrums that appeared early on in set theory by limiting what we can now call a set, and any collection that is not a *set* is now called a *class*. However, don't think that philosophical problems don't remain. After all, in a sense all we have done is to simply say that it's forbidden to talk about something like "the set of all sets." Aside from the paradoxes that arise, we still haven't adequately explained why we can't talk about the set of everything.

"Do I contradict myself? Very well, then I contradict myself, I am large, I contain multitudes."

-*Walt Whitman, Leaves of Grass*

Lesson 5

Epsilon-Delta Proofs

Recall the epsilon-delta definition of a limit:

"$\lim_{x\to a} f(x) = L$ if and only if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that if $0 < |x - a| < \delta$, then $|f(x) - L| < \varepsilon$."

In a more symbolic form we write it this way:

$$\lim_{x\to a} f(x) = L \Leftrightarrow \forall \varepsilon > 0, \exists \delta > 0 \ni 0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon .$$

Ponder what this means until you have a deep understanding of it.

In doing a proof that involves an $\varepsilon - \delta$ argument, the proof usually starts with the phrase "Let $\varepsilon > 0$," and then we have to find a suitable δ as a function of ε. For example, some proof involving a limit might begin with "Let $\varepsilon > 0$ and set $\delta = \frac{\varepsilon}{5}$." Below are a few proofs for you to do as exercises, most of which involve $\varepsilon - \delta$ arguments. I should probably give you an example or two at this point, but (a) I'm feeling very lazy today, and (b) you need to realize that you are not helpless and that you can take responsibility for your own learning. Lots and lots of information is already available to you for free online, so give it your best shot! ☺

1. Prove: $\lim_{x\to 1}(3x+2)=5$

2. If $\lim_{x\to a} f(x)=L$ and $c\in\mathbb{R}$, then $\lim_{x\to a} c\cdot f(x)=c\cdot L$.

3. A *lemma* is a theorem that is used to prove another theorem. A *corollary* is a theorem which is an immediate consequence of another theorem. Below is a *lemma* that is quite useful for proving other theorems about limits. Prove the *lemma*.

   Lemma: If $\lim_{x\to a} f(x)=L$ and $\lim_{x\to a} g(x)=M$, then for every $\varepsilon>0$, there exists a common $\delta>0$ such that if $0<|x-a|<\delta$, then both $|f(x)-L|<\varepsilon$ and $|g(x)-M|<\varepsilon$.

4. Certain inequalities are useful for doing lots of proofs. One of the most famous of such inequalities is *the triangle inequality*. Prove the following theorem, and also figure out why it is called *the triangle inequality*.

   The Triangle Inequality: If $a,b\in\mathbb{R}$, then $|a+b|\le|a|+|b|$.

5. Use *the triangle inequality* to prove the following limit theorem,

   Prove: If $\lim_{x\to a} f(x)=L$ and $\lim_{x\to a} g(x)=M$, then $\lim_{x\to a}\left[f(x)+g(x)\right]=L+M$.

Lesson 5 – Answers

Epsilon-Delta Proofs

Recall the epsilon-delta definition of a limit:

"$\lim_{x\to a} f(x) = L$ if and only if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that if $0 < |x - a| < \delta$, then $|f(x) - L| < \varepsilon$."

In a more symbolic form we write it this way:

$$\lim_{x\to a} f(x) = L \Leftrightarrow \forall \varepsilon > 0, \exists \delta > 0 \ni 0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon .$$

Ponder what this means until you have a deep understanding of it.

In doing a proof that involves an $\varepsilon - \delta$ argument, the proof usually starts with the phrase "Let $\varepsilon > 0$," and then we have to find a suitable δ as a function of ε. For example, some proof involving a limit might begin with "Let $\varepsilon > 0$ and set $\delta = \frac{\varepsilon}{5}$." Below are a few proofs for you to do as exercises, most of which involve $\varepsilon - \delta$ arguments. I should probably give you an example or two at this point, but (a) I'm feeling very lazy today, and (b) you need to realize that you are not helpless and that you can take responsibility for your own learning. Lots and lots of information is already available to you for free online, so give it your best shot! ☺

1. Prove: $\lim_{x\to 1}(3x+2)=5$

   Proof: Let $\varepsilon>0$ and let $\delta=\frac{\varepsilon}{3}$. Then

   $0<|x-1|<\delta \Rightarrow |x-1|<\frac{\varepsilon}{3} \Rightarrow 3|x-1|<\varepsilon \Rightarrow |3x-3|<\varepsilon \Rightarrow |(3x+2)-5|<\varepsilon$. Therefore,

   $\lim_{x\to 1}(3x+2)=5$. □

2. If $\lim_{x\to a} f(x)=L$ and $c\in\mathbb{R}$, then $\lim_{x\to a} c\cdot f(x)=c\cdot L$.

   Proof: Suppose $\lim_{x\to a} f(x)=L$. Then for every $\varepsilon>0$, there exists $\delta>0$ such that if $0<|x-a|<\delta$, then $|f(x)-L|<\varepsilon$. In particular, we can find a $\delta>0$ that corresponds to some $\frac{\varepsilon}{|c|}>0$ where $c\in\mathbb{R}$. Thus, let $\frac{\varepsilon}{|c|}$ and δ be so given. Then

   $0<|x-a|<\delta \Rightarrow |f(x)-L|<\frac{\varepsilon}{|c|} \Rightarrow |c|\cdot|f(x)-L|<|c|\cdot\frac{\varepsilon}{|c|} \Rightarrow |c\cdot f(x)-c\cdot L|<\varepsilon$. Therefore,

   $\lim_{x\to a} c\cdot f(x)=c\cdot L$. □

3. A *lemma* is a theorem that is used to prove another theorem. A *corollary* is a theorem which is an immediate consequence of another theorem. Below is a *lemma* that is quite useful for proving other theorems about limits. Prove the *lemma*.

   Lemma: If $\lim_{x\to a} f(x)=L$ and $\lim_{x\to a} g(x)=M$, then for every $\varepsilon>0$, there exists a

common $\delta > 0$ such that if $0 < |x - a| < \delta$, then both $|f(x) - L| < \varepsilon$ and $|g(x) - M| < \varepsilon$.

Proof: Let $\varepsilon > 0$. Then there exists $\delta_1 > 0$ such that if $0 < |x - a| < \delta_1$, then $|f(x) - L| < \varepsilon$. Similarly, there exists $\delta_2 > 0$ such that if $0 < |x - a| < \delta_2$, then $|g(x) - M| < \varepsilon$. Now let δ equal the lesser of δ_1 and δ_2. Then $0 < |x - a| < \delta$ implies both $|f(x) - L| < \varepsilon$ and $|g(x) - M| < \varepsilon$. □

4. Certain inequalities are useful for doing lots of proofs. One of the most famous of such inequalities is *the triangle inequality*. Prove the following theorem, and also figure out why it is called *the triangle inequality*. NOTE: This inequality is used in a lot of proofs in calculus and analysis, but unlike others, it doesn't require an epsilon-delta argument.

The Triangle Inequality: If $a, b \in \mathbb{R}$, then $|a + b| \leq |a| + |b|$.

Proof: We'll consider several separate cases.

(Case 1: *a* and *b* both positive) Suppose both $a, b > 0$. Then $|a + b| = a + b = |a| + |b|$.

(Case 2: *a* and *b* both negative) Suppose both $a, b < 0$. Then $|a + b| = -(a + b) = -a + (-b) = |a| + |b|$.

(Case 3: *a* and *b* different signs with $|a| > |b|$) In this case, $|a+b| < |a| < |a| + |b|$.

(Case 4: *a* and *b* different signs with $|a| < |b|$) In this case, $|a+b| < |b| < |a| + |b|$.

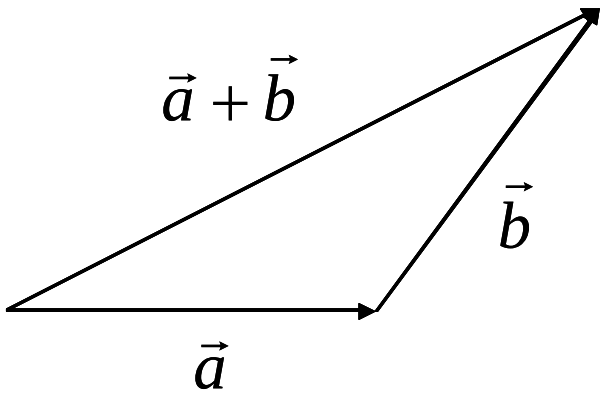
(Case 5: $a = 0$) In this case, $|a+b| = |b| = |a| + |b|$.

(Case 6: $b = 0$) In this case, $|a+b| = |a| = |a| + |b|$.

Therefore, $|a+b| \le |a| + |b|$. □

The proof I've giving above is far from the shortest proof that is available. Nonetheless, it is a very straightforward proof even though one of my professors used to refer to such straightforward, and perhaps inelegant, proofs as examples of "brute force and ignorance."

The *triangle inequality* derives its name from the observation that the length of one side of a triangle is always less than or equal to the sum of the lengths of the other two sides. If we represent the lengths of our sides by vectors, then we just arrive at another version of the *triangle inequality*.



$$\left\|\vec{a}+\vec{b}\right\| \le \left\|\vec{a}\right\| + \left\|\vec{b}\right\|$$

5. Use *the triangle inequality* to prove the following limit theorem,

Prove: If $\lim_{x\to a} f(x) = L$ and $\lim_{x\to a} g(x) = M$, then $\lim_{x\to a}\left[f(x)+g(x)\right] = L+M$.

Proof: Let $\varepsilon > 0$. Then by previous proof, there exists a common $\delta > 0$ such that if $0 < |x-a| < \delta$, then $|f(x)-L| < \frac{\varepsilon}{2}$ and $|g(x)-M| < \frac{\varepsilon}{2}$. Hence, by the *triangle inequality*,

$$\left|\left(f(x)+g(x)\right)-(L+M)\right| = \left|(f(x)-L)+(g(x)-M)\right| \le |f(x)-L| + |g(x)-M| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2} = \varepsilon.$$

Therefore, $\lim_{x\to a}\left[f(x)+g(x)\right] = L+M$. □

Lesson 6

Axioms for the Real Numbers

One of the great achievements of modern mathematics was the establishment of axiomatic systems for all its major branches. As a result, proofs are now simply logical consequences of the established definitions and axioms. In many respects, we can think of these axioms as specifying the rules of the game that may be used to construct proofs. Below is one presentation of the axioms for the real number system, $\mathbb{R}$.

Definition: If *A* is a set, then * is a *binary operation* on *A* if for every $a,b \in A$, we have that $a*b \in A$.

The Algebraic Axioms:

1. There exists a binary operation "+" on $\mathbb{R}$ such that for every $a,b \in \mathbb{R}$, $a+b \in \mathbb{R}$. (Closure under Addition)
2. For every $a,b,c \in \mathbb{R}$, $a+(b+c)=(a+b)+c$. (Associative Law of Addition)
3. There exists an element 0 in $\mathbb{R}$ such that for every $a \in \mathbb{R}$, $a+0=a$. (Existence of an Additive Identity)
4. For every $a \in \mathbb{R}$, there exists $-a \in \mathbb{R}$ such that $a+(-a)=0$. (Existence of Additive Inverses)
5. For every $a,b \in \mathbb{R}$, $a+b=b+a$. (Commutative Law of Addition)

6. There exists a binary operation "$\cdot$" on $\mathbb{R}$ such that for every $a,b \in \mathbb{R}$, $a \cdot b \in \mathbb{R}$. (Closure under Multiplication)
7. For every $a,b,c \in \mathbb{R}$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. (Associative Law of Multiplication)
8. There exists an element 1 in $\mathbb{R}$ that is not equal to 0 such that for every $a \in \mathbb{R}$, $a \cdot 1 = a$. (Existence of a Multiplicative Identity)
9. For every nonzero $a \in \mathbb{R}$, there exists $a^{-1} \in \mathbb{R}$, (called either the reciprocal of *a* or *a*-inverse) such that $a \cdot a^{-1} = 1$. (Existence of Multiplicative Inverses)
10. For every $a,b \in \mathbb{R}$, $a \cdot b = b \cdot a$. (Multiplicative Law of Addition)
11. For every $a,b,c \in \mathbb{R}$, $a \cdot (b + c) = a \cdot b + a \cdot c$. (The Distributive Law)

Notation: For convenience, we will often write $a \cdot b$ as ab.

The Order Axioms:

There exists a subset $P \subset \mathbb{R}$ (called the set of positive real numbers) satisfying the following:

1. The set *P* is closed under addition.
2. The set *P* is closed under multiplication.
3. For every $a \in \mathbb{R}$, exactly one of the following holds: $a = 0$ or $a \in P$ or $-a \in P$. (the Law of Trichotomy)

Definition:

1. $a > b$ if and only if $a - b \in P$.
2. $a \geq b$ means that either $a > b$ or $a = b$.

3. $a < b$ if and only if $b - a \in P$.

4. $a \leq b$ means that either $a < b$ or $a = b$.

Definition: If $A \subseteq \mathbb{R}$ and if $b \in \mathbb{R}$ such that for every $a \in A$, we have that $b \geq a$, then *b* is an *upper bound* for the set *A*.

Definition: If $A \subseteq \mathbb{R}$, then $b \in \mathbb{R}$ is a *least upper bound* for *A* if *b* is an upper bound and if for every other upper bound *c* of *A*, we have that $b \leq c$.

Completeness Axiom: Every nonempty set $A \subseteq \mathbb{R}$ with an upper bound *b* has a least upper bound.

It can be shown that the only set which satisfies all three sets of axioms (the algebraic axioms, the order axioms, and the completeness axiom) is the familiar set of real numbers. Below, you will use the axioms for the real number system to prove several familiar properties. As you do so, keep in mind that the axioms are essentially the rules for a game, and they tell us what moves are legal. Thus, we can only perform a move if it is permitted by the axioms or by some theorem that we have already derived from the axioms. For example, at this point all that $-a$ means is "the additive inverse of *a*," and something that we generally take for granted such as $-(-a) = a$ is now going to be a theorem that we have prove as a consequence of our axioms.

1. Prove: $0 + a = a$

2. Prove: $1 \cdot a = a$

3. Prove: The additive identity element 0 is unique.

4. Prove: The multiplicative identity element 1 is unique.

5. Prove: Additive inverses are unique.

6. Prove: Multiplicative inverses are unique.

7. Prove: For every $a \in \mathbb{R}$, $a \cdot 0 = 0 = 0 \cdot a$.

8. Prove: For every $a \in \mathbb{R}$, $-(-a) = a$.

9. Prove: For every $a \in \mathbb{R}$, $(-a) = (-1) \cdot a$.

10. Prove: For every $a, b \in \mathbb{R}$, $(-a)b = -(ab)$.

Lesson 6 – Answers

Axioms for the Real Numbers

One of the great achievements of modern mathematics was the establishment of axiomatic systems for all its major branches. As a result, proofs are now simply logical consequences of the established definitions and axioms. In many respects, we can think of these axioms as specifying the rules of the game that may be used to construct proofs. Below is one presentation of the axioms for the real number system, $\mathbb{R}$.

Definition: If *A* is a set, then * is a *binary operation* on *A* if for every $a,b \in A$, we have that $a*b \in A$.

The Algebraic Axioms:

1. There exists a binary operation "+" on $\mathbb{R}$ such that for every $a,b \in \mathbb{R}$, $a+b \in \mathbb{R}$. (Closure under Addition)
2. For every $a,b,c \in \mathbb{R}$, $a+(b+c)=(a+b)+c$. (Associative Law of Addition)
3. There exists an element 0 in $\mathbb{R}$ such that for every $a \in \mathbb{R}$, $a+0=a$. (Existence of an Additive Identity)
4. For every $a \in \mathbb{R}$, there exists $-a \in \mathbb{R}$ such that $a+(-a)=0$. (Existence of Additive Inverses)
5. For every $a,b \in \mathbb{R}$, $a+b=b+a$. (Commutative Law of Addition)

6. There exists a binary operation "$\cdot$" on $\mathbb{R}$ such that for every $a,b \in \mathbb{R}$, $a \cdot b \in \mathbb{R}$. (Closure under Multiplication)
7. For every $a,b,c \in \mathbb{R}$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. (Associative Law of Multiplication)
8. There exists an element 1 in $\mathbb{R}$ that is not equal to 0 such that for every $a \in \mathbb{R}$, $a \cdot 1 = a$. (Existence of a Multiplicative Identity)
9. For every nonzero $a \in \mathbb{R}$, there exists $a^{-1} \in \mathbb{R}$, (called either the reciprocal of *a* or *a*-inverse) such that $a \cdot a^{-1} = 1$. (Existence of Multiplicative Inverses)
10. For every $a,b \in \mathbb{R}$, $a \cdot b = b \cdot a$. (Multiplicative Law of Addition)
11. For every $a,b,c \in \mathbb{R}$, $a \cdot (b + c) = a \cdot b + a \cdot c$. (The Distributive Law)

Notation: For convenience, we will often write $a \cdot b$ as ab.

The Order Axioms:

There exists a subset $P \subset \mathbb{R}$ (called the set of positive real numbers) satisfying the following:

1. The set *P* is closed under addition.
2. The set *P* is closed under multiplication.
3. For every $a \in \mathbb{R}$, exactly one of the following holds: $a = 0$ or $a \in P$ or $-a \in P$. (the Law of Trichotomy)

Definition:

1. $a > b$ if and only if $a - b \in P$.
2. $a \geq b$ means that either $a > b$ or $a = b$.

3. $a < b$ if and only if $b - a \in P$.

4. $a \leq b$ means that either $a < b$ or $a = b$.

Definition: If $A \subseteq \mathbb{R}$ and if $b \in \mathbb{R}$ such that for every $a \in A$, we have that $b \geq a$, then *b* is an *upper bound* for the set *A*.

Definition: If $A \subseteq \mathbb{R}$, then $b \in \mathbb{R}$ is a *least upper bound* for *A* if *b* is an upper bound and if for every other upper bound *c* of *A*, we have that $b \leq c$.

Completeness Axiom: Every nonempty set $A \subseteq \mathbb{R}$ with an upper bound *b* has a least upper bound.

It can be shown that the only set which satisfies all three sets of axioms (the algebraic axioms, the order axioms, and the completeness axiom) is the familiar set of real numbers. Below, you will use the axioms for the real number system to prove several familiar properties. As you do so, keep in mind that the axioms are essentially the rules for a game, and they tell us what moves are legal. Thus, we can only perform a move if it is permitted by the axioms or by some theorem that we have already derived from the axioms. For example, at this point all that $-a$ means is "the additive inverse of *a*," and something that we generally take for granted such as $-(-a) = a$ is now going to be a theorem that we have prove as a consequence of our axioms.

1. Prove: $0 + a = a$

   Proof: By our axioms, $a + 0 = a$ and $a + b = b + a$. Hence, $0 + a = a + 0 = a$. □

2. Prove: $1 \cdot a = a$

   Proof: By our axioms, $a \cdot 1 = a$ and $a \cdot b = b \cdot a$. Hence, $a \cdot 1 = 1 \cdot a = a$. □

3. Prove: The additive identity element 0 is unique.

   Proof: Suppose $0'$ is also an additive identity element. Then it follows that $0' = 0' + 0 = 0$. Therefore, the additive identity element is unique. □

4. Prove: The multiplicative identity element 1 is unique.

   Proof: Suppose $1'$ is also a multiplicative identity element. Then it follows that $1' = 1' \cdot 1 = 1$. Therefore, the multiplicative identity element is unique. □

5. Prove: Additive inverses are unique.

   Proof: Suppose that $-a$ and $-a'$ are both additive inverses of a. Then

   $$-a + a = 0 = -a' + a \Rightarrow -a + [a + (-a)] = -a' + [a + (-a)] \Rightarrow -a + 0 = -a' + 0 \Rightarrow -a = -a'.$$

Therefore, additive inverses are unique. □

6. Prove: Multiplicative inverses are unique.

Proof: Suppose a^{-1} and $a^{-1'}$ are both multiplicative inverses of a. Then $a^{-1}a = 1 = a^{-1'}a \Rightarrow a^{-1}(aa^{-1}) = a^{-1'}(aa^{-1}) \Rightarrow a^{-1} \cdot 1 = a^{-1'} \cdot 1 \Rightarrow a^{-1} = a^{-1'}$. Therefore, multiplicative inverses are unique. □

7. Prove: For every $a \in \mathbb{R}$, $a \cdot 0 = 0 = 0 \cdot a$.

Proof: Let $a \in \mathbb{R}$. Then $a \cdot 0 = a(0+0) = a \cdot 0 + a \cdot 0 \Rightarrow 0 = a \cdot 0 = 0 \cdot a$. □

8. Prove: For every $a \in \mathbb{R}$, $-(-a) = a$.

Proof: Let $a \in \mathbb{R}$. Then $-(-a) + (-a) = 0 \Rightarrow -(-a)$ is an additive inverse of $a \Rightarrow -(-a) = a$ since additive inverses are unique. □

9. Prove: For every $a \in \mathbb{R}$, $(-a) = (-1) \cdot a$.

Proof: Let $a \in \mathbb{R}$. Then $(-1) \cdot a + a = (-1) \cdot a + 1 \cdot a = (-1+1) \cdot a = 0 \cdot a = 0 \Rightarrow (-1) \cdot a = -a$ since additive inverses are unique. □

10. Prove: For every $a, b \in \mathbb{R}$, $(-a)b = -(ab)$.

Proof: Let $a, b \in \mathbb{R}$. Then $(-a)b + ab = (-a + a)b = 0 \cdot b = 0 \Rightarrow (-a)b = -(ab)$ since additive inverses are unique. □

Lesson 7

Everything Else that I Forgot to Mention

If A and B are sets, then the *Cartesian product* of A and B is defined as $A \times B = \{(a,b) \mid a \in A \text{ and } b \in B\}$. In other words, the Cartesian product of the two sets is the set of all ordered pairs that can be formed by pairing elements of the first set with elements of the second set. For example, using this definition, we can think of the coordinate plane as just the Cartesian product of the real numbers with the real numbers, $\mathbb{R} \times \mathbb{R}$. Additionally, this construction is going to allow us to give a more abstract definition of things we are already familiar with such as *functions*.

Anytime you see something in mathematics called *Cartesian*, you know it is being named after René Descartes (1596 – 1650). During his lifetime, he was famous both as a philosopher and as a mathematician. He also worked as a mercenary soldier. It was a time when mathematicians were more like Rambo. Furthermore, he died at age 53 as a result of getting up too early in the morning. Hey! I don't make these things up!

A *relation* in $A \times B$ is any subset of $A \times B$. For example, in $\mathbb{R} \times \mathbb{R}$ the subset $L = \{(a,b) \mid a \in \mathbb{R} \text{ and } b \in \mathbb{R} \text{ and } a < b\}$ shows how we can represent the familiar relation "less than" in terms of a Cartesian product.

A *function* in $A \times B$ is a relation *F* such that if (x, y_1) and (x, y_2) belong to *F*, then $y_1 = y_2$. This is probably a very fancy way of saying something that you already know. Namely, that a function can't take a single element from its domain and pair it with more than one element in its range.

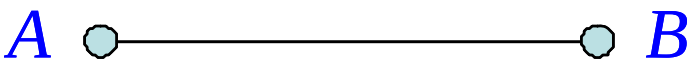
The equal sign was invented by the Welsh mathematician Robert Recorde in 1557, and since it consists of two parallel lines of equal length $(=)$, Recorde felt that there was no symbol better suited for denoting equality. Later mathematicians noted that equality has certain properties that we now call the *reflexive property* $(a = a)$, the *symmetric property* $(\text{If } a = b, \text{ then } b = a)$, and the *transitive property* $(\text{If } a = b \text{ and } b = c, \text{ then } a = c)$, and from these properties they abstracted to define an *equivalence relation*.

If *A* is a set and *E* is a relation in $A \times A$, then *E* is called an *equivalence relation* if the following conditions are met:

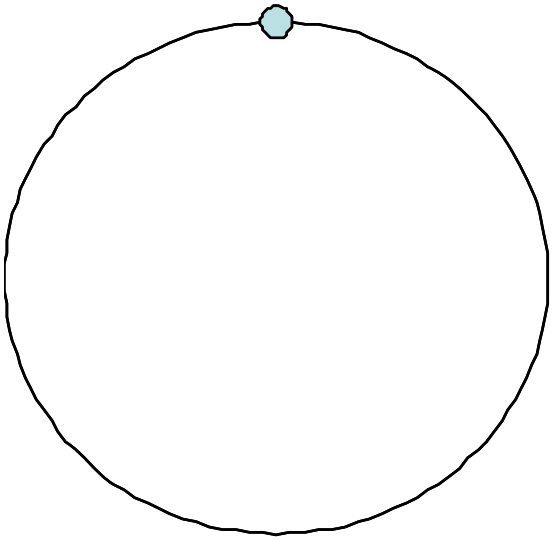
1. $\forall x \in A, (x, x) \in E$. (reflexive)
2. $\forall x, y \in A, \text{ if } (x, y) \in E, \text{ then } (y, x) \in E$. (symmetric)
3. $\forall x, y, z \in A, \text{ if } (x, y) \in E \text{ and } (y, z) \in E, \text{ then } (x, z) \in E$. (transitive)

If we are given for some set *A* an element *a* and an equivalence relation *E* in $A \times A$, then the set of all elements of *A* that are equivalent to *a* is called the *equivalence class of a*. I won't go into detail here, but suffice it to say that the concept of an equivalence class is one of the deepest and furthest reaching in all of mathematics. We can literally change

our reality through what things we see as equivalent with one another. For example, we don't see *even numbers* until we see integers that are divisible by 2 as being "equivalent." Similarly, we don't see the forest until we see what makes one tree equivalent to another, and by making the two ends of a string equivalent to one another, we can turn a line segment into a circle.



$A = B$



If G is a set and if we have a function that goes from $G \times G \to G$, then we call this type of function a *binary operation*. When we have such a function, we don't normally use the $f(a,b)$ type of notation. Instead, we use maybe a symbol such as * for this operation and

write something like $a*b$. For example, we can interpret "+" as a binary operation from $\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, and we commonly write things as $2+3=5$ rather than $f(2,3)=5$.

Using our definition of a binary operation and just a few of the algebraic properties for the real numbers, we can now define the concept of a *group* which is one of the most important algebraic structures of higher mathematics. Because there are so many things in the world of mathematics that are groups, a single theorem about groups applies to many different situations.

Let *G* be a non-empty set and let * be a binary operation defined on $G\times G$. Then *G* is a *group* if the following axioms are satisfied.

1. For every $a,b\in G$, $a*b\in G$. (closure)
2. For every $a,b,c\in G$, $a*(b*c)=(a*b)*c$. (associativity)
3. There exists an element $e\in G$ such that for every $a\in G$, $e*a=a=a*e$. (identity)
4. For every $a\in G$ there exists $a^{-1}\in G$ such that $a*a^{-1}=e=a^{-1}*a$. (inverses)

If the following additional property holds, then we call *G* a *commutative* or *abelian group*.

5. For every $a,b\in G$, $a*b=b*a$. (commutativity)

Again, the concept of a group is far reaching, and as you'll see in the exercises, groups are also associated with permutations and symmetry.

1. If $A=\{1,2,3\}$ and $B=\{4,5\}$, find $A\times B$, $B\times A$, and $B\times B$.

2. *Seven Brides for Seven Brothers*

   Let *B* be the set of brothers in this popular movie musical, and let $E=\{(a,b)\,|\,a,b\in B \text{ and } a \& b \text{ are brothers}\}$. Is *E* a relation in $B\times B$? If so, then is *E* an equivalence relation? Why or why not?

3. Read the article in the Wikipedia on René Descartes. Everyone should know something about Descartes.

4. To learn more about groups, go to www.docbenton.com and read as much as you can of *Doc Benton's Fantastic Guide to Group Theory, Rubik's Cube, Permutations, Symmetry, and All That Is!*. If you want, you can skip the chapters on Rubik's cube and focus only on those on group theory. This should give you a pretty good introduction to the subject. Enjoy!

Lesson 7 – Answers

Everything Else that I Forgot to Mention

If A and B are sets, then the *Cartesian product* of A and B is defined as $A \times B = \{(a,b) \mid a \in A \text{ and } b \in B\}$. In other words, the Cartesian product of the two sets is the set of all ordered pairs that can be formed by pairing elements of the first set with elements of the second set. For example, using this definition, we can think of the coordinate plane as just the Cartesian product of the real numbers with the real numbers, $\mathbb{R} \times \mathbb{R}$. Additionally, this construction is going to allow us to give a more abstract definition of things we are already familiar with such as *functions*.

Anytime you see something in mathematics called *Cartesian*, you know it is being named after René Descartes (1596 – 1650). During his lifetime, he was famous both as a philosopher and as a mathematician. He also worked as a mercenary soldier. It was a time when mathematicians were more like Rambo. Furthermore, he died at age 53 as a result of getting up too early in the morning. Hey! I don't make these things up!

A *relation* in $A \times B$ is any subset of $A \times B$. For example, in $\mathbb{R} \times \mathbb{R}$ the subset $L = \{(a,b) \mid a \in \mathbb{R} \text{ and } b \in \mathbb{R} \text{ and } a < b\}$ shows how we can represent the familiar relation "less than" in terms of a Cartesian product.

A *function* in $A \times B$ is a relation *F* such that if (x, y_1) and (x, y_2) belong to *F*, then $y_1 = y_2$. This is probably a very fancy way of saying something that you already know. Namely, that a function can't take a single element from its domain and pair it with more than one element in its range.

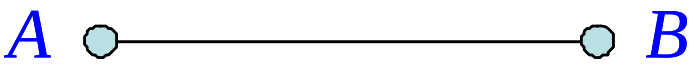
The equal sign was invented by the Welsh mathematician Robert Recorde in 1557, and since it consists of two parallel lines of equal length $(=)$, Recorde felt that there was no symbol better suited for denoting equality. Later mathematicians noted that equality has certain properties that we now call the *reflexive property* $(a = a)$, the *symmetric property* $(\text{If } a = b, \text{ then } b = a)$, and the *transitive property* $(\text{If } a = b \text{ and } b = c, \text{ then } a = c)$, and from these properties they abstracted to define an *equivalence relation*.

If *A* is a set and *E* is a relation in $A \times A$, then *E* is called an *equivalence relation* if the following conditions are met:

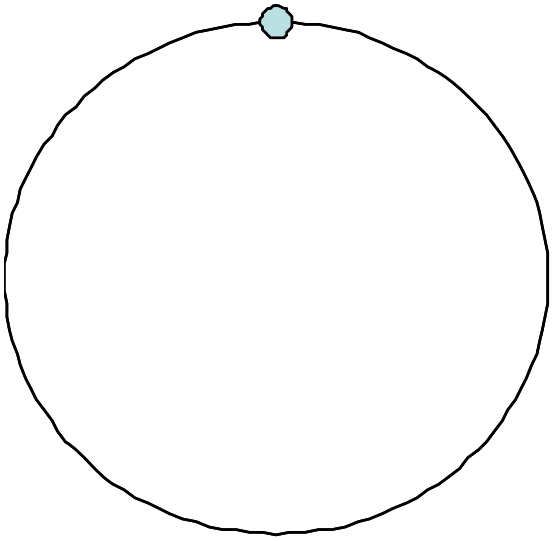
1. $\forall x \in A, (x, x) \in E$. (reflexive)
2. $\forall x, y \in A, \text{ if } (x, y) \in E, \text{ then } (y, x) \in E$. (symmetric)
3. $\forall x, y, z \in A, \text{ if } (x, y) \in E \text{ and } (y, z) \in E, \text{ then } (x, z) \in E$. (transitive)

If we are given for some set *A* an element *a* and an equivalence relation *E* in $A \times A$, then the set of all elements of *A* that are equivalent to *a* is called the *equivalence class of a*. I won't go into detail here, but suffice it to say that the concept of an equivalence class is one of the deepest and furthest reaching in all of mathematics. We can literally change

our reality through what things we see as equivalent with one another. For example, we don't see *even numbers* until we see integers that are divisible by 2 as being "equivalent." Similarly, we don't see the forest until we see what makes one tree equivalent to another, and by making the two ends of a string equivalent to one another, we can turn a line segment into a circle.



$A = B$



If G is a set and if we have a function that goes from $G \times G \to G$, then we call this type of function a *binary operation*. When we have such a function, we don't normally use the $f(a,b)$ type of notation. Instead, we use maybe a symbol such as * for this operation and

write something like $a*b$. For example, we can interpret "+" as a binary operation from $\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, and we commonly write things as $2+3=5$ rather than $f(2,3)=5$.

Using our definition of a binary operation and just a few of the algebraic properties for the real numbers, we can now define the concept of a *group* which is one of the most important algebraic structures of higher mathematics. Because there are so many things in the world of mathematics that are groups, a single theorem about groups applies to many different situations.

Let *G* be a non-empty set and let * be a binary operation defined on $G\times G$. Then *G* is a *group* if the following axioms are satisfied.

1. For every $a,b\in G$, $a*b\in G$. (closure)
2. For every $a,b,c\in G$, $a*(b*c)=(a*b)*c$. (associativity)
3. There exists an element $e\in G$ such that for every $a\in G$, $e*a=a=a*e$. (identity)
4. For every $a\in G$ there exists $a^{-1}\in G$ such that $a*a^{-1}=e=a^{-1}*a$. (inverses)

If the following additional property holds, then we call *G* a *commutative* or *abelian group*.

5. For every $a,b\in G$, $a*b=b*a$. (commutativity)

Again, the concept of a group is far reaching, and as you'll see in the exercises, groups are also associated with permutations and symmetry.

1. If $A=\{1,2,3\}$ and $B=\{4,5\}$, find $A\times B$, $B\times A$, and $B\times B$.

   $A\times B=\{(1,4),(1,5),(2,4),(2,5),(3,4),(3,5)\}$

   $B\times A=\{(4,1),(5,1),(4,2),(5,2),(4,3),(5,3)\}$

   $B\times B=\{(4,4),(4,5),(5,4),(5,5)\}$

2. *Seven Brides for Seven Brothers*

   Let *B* be the set of brothers in this popular movie musical, and let $E=\{(a,b)\,|\,a,b\in B \text{ and } a \& b \text{ are brothers}\}$. Is *E* a relation in $B\times B$? If so, then is *E* an equivalence relation? Why or why not?

   The seven brothers from the musical are *Adam, Benjamin, Gideon, Frank, Daniel, Caleb,* and *Ephraim*. The set *E* is a relation in $B\times B$, but it's not an equivalence relation. It is symmetric since if Adam is a brother of Benjamin, then Benjamin is a brother of Adam. It is also transitive since Adam is a brother of Benjamin and Benjamin is a brother of Gideon implies that Adam is a brother of Gideon. However, it is not reflexive since Adam is not a brother of Adam.

3. Read the article in the Wikipedia on René Descartes. Everyone should know something about Descartes.

   I read it. It's great!

4. To learn more about groups, go to www.docbenton.com and read as much as you can of *Doc Benton's Fantastic Guide to Group Theory, Rubik's Cube, Permutations, Symmetry, and All That Is!*. If you want, you can skip the chapters on Rubik's cube and focus only on those on group theory. This should give you a pretty good introduction to the subject. Enjoy!

   I not only read it, I wrote it!