Basic logic for first year analysis students

Peter Haggstrom
www.gotohaggstrom.com
mathsatbondibeach@gmail.com

October 18, 2013

1 Background

Because analysis is so highly proof driven students need to develop a facility with fundamental logical operations that underpin the proofs. In essence this boils down to understanding proof by contradiction and how to negate complex definitions such as the limit concept, continuity, uniform continuity, differentiability and so on. For instance, in a proof you may be required to assume that some function is not continuous so you will need to be able to accurately negate the concept of continuity. Once the structure of the relevant definition is understood the logical manipulations are straightforward and with some practice it is easy to "rattle" off the relevant negation.

2 Proof by contradiction

Suppose you want to prove the statement 'P'.

You assume the truth of 'not-P' (ie '$\neg P$)'.

You then derive a contradiction.

Hence '$\neg P$' is false. This is because of the so-called "law of excluded middle" ie either 'P' is true or '$\neg P$' is true - there is nothing "in between". This means that P is true.

3 More complex deductions

Every mathematical proof is essentially a series of "If P then Q" statements which is symbolised thus: '$P \implies Q$'. Now '$P \implies Q$' is truth functionally equivalent to '$\neg(P \,\&\, \neg Q)$'. Just consider the following truth tables:

Table 1: P $\implies$ Q

P	$\implies$	Q
T	**T**	T
T	**F**	F
F	**T**	T
F	**T**	F

Table 2: $\neg(P\&\neg Q)$

P	Q	P & ¬ Q	¬ (P & ¬ Q)
T	T	F	**T**
T	F	T	**F**
F	T	F	**T**
F	T	F	**T**

3.1 De Morgan's Laws

You are probably already aware of these in their set-theoretic form.

$A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C)$
$A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C)$

The truth functional analogs are respectively as follows:

$p\,.\,(q \vee r) \equiv pq \vee pr$
$p \vee qr \equiv p \vee q\,.\,p \vee r$

The truth functional equivalence symbol "$\equiv$" is used to indicate that both sides of the relation are truth functionally equivalent under any assignment of truth values.

The conditional "If P then Q" is false only when the antecedent (P) is true and the consequent (Q) is false. To prove " $P \implies Q$" the approach is to assume that "P" is true and that "Q" is false ie that "$\neg Q$" is true. In combination, then, one is assuming that "$P\&\neg Q$" is true. A contradiction is then estabished. Thus it cannot be the case that "$P\&\neg Q$" is true. This means that "$(P\&\neg Q)$" must be false which is to say that "$\neg(P\&\neg Q)$" is true. In other words "$P \implies Q$" is true.

Consider now the definition of the limit of a function. A function $f(x)$ is said to have a limit L as x approaches x_0, if for every positive ϵ there is a positive δ such that $f(x)$ is within ϵ of L for every $x \neq x_0$ which is within δ of x_0. We can translate this statement in terms of logical quantifiers as follows:

$$(\forall\epsilon)\left\{\epsilon > 0 \implies (\exists\delta)[\,\delta > 0\,\&\,(\forall x)\,(x \neq x_0\,\&\,0 < |x - x_0| < \delta \implies |f(x) - L| < \epsilon)\,]\right\} \tag{1}$$

(1) makes it clear that you must be able to take any $\epsilon > 0$ and then find the $\delta > 0$ which will ensure that for x within δ of x_0 the value of $f(x)$ will be within ϵ of L.

Note the structure of (1) - it is of the form:

$$(\forall\epsilon)\,\big\{\mathbf{P} \implies (\exists\delta)\,\big[\mathbf{Q}\,\&\,(\forall x)\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)\big]\big\} \tag{2}$$

If you wanted to assume for the purposes of proof by contradiction that f does not have a limit L you would need to negate (1) and although most mathematicians will assume you can simply do this negation verbally, there is a systematic symbolic way of doing it that will get you the correct answer. Indeed, the logic software that exists simply implements a suite of transformational rules such as '$\neg(\forall x) \rightarrow (\exists x)\neg'$.

The negation of (2) is:

$$\neg(\forall\epsilon)\,\big\{\mathbf{P} \implies (\exists\delta)\,\big[\mathbf{Q}\,\&\,(\forall x)\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)\big]\big\} \tag{3}$$

If everything has a certain property then the negation means that there is at least one object that does not have the required property. This means that you can "push through" the negation sign while changing the universal quantifier '$\forall$' to '$\exists$' along the way thus:

$$(\exists\epsilon)\,\neg\big\{\mathbf{P} \implies (\exists\delta)\,\big[\mathbf{Q}\,\&\,(\forall x)\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)\big]\big\} \tag{4}$$

Now we have to negate something that is structurally of the form "$P \implies W$ " where "W" is "$(\exists\delta)\,\big[\mathbf{Q}\,\&\,(\forall x)\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)\big]$"

Now "$\neg(P \implies W)$" is equivalent to "$\neg\neg(P\,\&\,\neg W)$" which is simply" $P\,\&\,\neg W$". To negate "W" we require:

$$\neg(\exists\delta)\,\big[\mathbf{Q}\,\&\,(\forall x)\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)\big] \tag{5}$$

If there is at least one thing that has a certain property, the negation is that everything does not have that property and we can push the negation sign through again while changing the existential quantifier to a universal quantifier thus:

$$(\forall\delta)\,\neg\big[\mathbf{Q}\,\&\,(\forall x)\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)\big] \tag{6}$$

"$\big[\mathbf{Q}\,\&\,(\forall x)\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)\big]$" has the structure "$Q\,\&\,Z$" so "$\neg(Q\,\&\,Z)$" is equivalent to "$\neg Q \vee \neg Z$ "(ie either $\neg Q$ or $\neg Z$). But "$\neg Q \vee \neg Z$" is equivalent to "$Q \implies \neg Z$".

So this means that (6) becomes:

$$(\forall\delta)\,\big[\delta > 0 \implies \neg(\forall x)\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)\big] \tag{7}$$

But "$\neg\big(\forall x\big)\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)$" is equivalent to "$\big(\exists x\big)\neg\big(\mathbf{R_1}\,\&\,\mathbf{R_2} \implies \mathbf{S}\big)$" , which in turn is equivalent to "$\neg\,\neg\big((\mathbf{R_1}\,\&\,\mathbf{R_2})\,\&\,\neg\mathbf{S}\big)$" or "$\big((\mathbf{R_1}\,\&\,\mathbf{R_2})\,\&\,\neg\mathbf{S}\big)$".

Putting all the bits together from (3) onwards we get that the negation is:

$$\big(\exists\epsilon\big)\big\{\epsilon > 0\,\&\,\big(\forall\delta\big)\,\big[\delta > 0 \implies \big(\exists x\big)\big[x \neq x_0\,\&\,0 < |x - x_0| < \delta\,\&\,|f(x) - L| \geq \epsilon\big]\,\big\} \quad (8)$$

In words this says that there is a positive ϵ such that for all positive δ you can choose at least one x within δ of x_0 such that $|f(x) - L| \geq \epsilon$ holds.

4 A practical application - the Archimedean Property

The Archimedean property asserts the following:

[AP] If a and b are positive integers, then there exists a positive integer n such that $na \geq b$.

All this says is that if you give me two positive integers I can multiply one of them by a sufficiently large positive integer so that the product will exceed the other integer.

Such a property cries out for a non-constructive proof ie a proof by contradiction. Thus we have to negate [AP]. In words the negation will be:

$\neg[AP]$: There exist positive integers a and b such that $na < b$ for every positive integer n.

If you want to mechanically do the logic transformations you start first with [AP] in logical formalism:

[AP]: $\big(\forall a\big)\,\big(\forall b\big)\big[\,(a \in \mathbb{Z}^+\,\&\,b \in \mathbb{Z}^+ \implies \big(\exists n\big)\big(n \in \mathbb{Z}^+\,\&\,na \geq b\big)\,\big]$

Then $\neg[AP]$ becomes:

$$\neg\,\big(\forall a\big)\,\big(\forall b\big)\big[\,(a \in \mathbb{Z}^+\,\&\,b \in \mathbb{Z}^+ \implies \big(\exists n\big)\big(n \in \mathbb{Z}^+\,\&\,na \geq b\big)\,\big]$$

$$\big(\exists a\big)\,\big(\exists b\big)\neg\big[\,(a \in \mathbb{Z}^+\,\&\,b \in \mathbb{Z}^+ \implies \big(\exists n\big)\big(n \in \mathbb{Z}^+\,\&\,na \geq b\big)\,\big]$$

$$\big(\exists a\big)\,\big(\exists b\big)\neg\neg\big[\,(a \in \mathbb{Z}^+\,\&\,b \in \mathbb{Z}^+\,\&\,\neg\big(\exists n\big)\big(n \in \mathbb{Z}^+\,\&\,na \geq b\big)\,\big]$$

$$\big(\exists a\big)\,\big(\exists b\big)\big[\,(a \in \mathbb{Z}^+\,\&\,b \in \mathbb{Z}^+\,\&\,\big(\forall n\big)\neg\big(n \in \mathbb{Z}^+\,\&\,na \geq b\big)\,\big]$$

$\big(\exists a\big)\,\big(\exists b\big)\big[\,(a \in \mathbb{Z}^+ \,\&\, b \in \mathbb{Z}^+ \,\&\, \big(\forall n\big)\big(n \in \mathbb{Z}^+ \implies na < b\big)\,\big]$

This last line is simply the formal representation of $\neg[AP]$ given above.

Now getting back to the proof of the Archimedean Property, we consider the set $S = \{b - na : n \in \mathbb{Z}^+\}$. This set contains only positive integers since $b > na$ and all the components are positive integers. The Well Ordering Principle (every non-empty set of natural numbers has a least element) tells us that S possesses a least element which we will call $b - ma$ for some positive integer m.

Now $b-(m+1)a$ is in S. Why? Because S is defined to cover numbers of the form $b-na$ where n is **any** positive integer, which $m+1$ certainly is. But $b-(m+1)a = b-ma-a < b-ma$ so we have a member of S which is less than $b-ma$, thereby contradicting the Well Ordering Principle. This gives us our contradiction and establishes the Archimedean Property using proof by contradiction.

5 The contrapositive

If you have to prove an implication such as "$P \implies Q$" you could equivalently prove its contrapositive, namely, "$\neg Q \implies \neg P$".

6 Negating some other basic definitions

Let's take the definition of continuity:

[C] : f is continuous at $x_0 \in \mathcal{I}$ if and only if given $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(x_0)| < \epsilon$ when $x \in \mathcal{I}$ and $|x - x_0| < \delta$

Negating [C] is no different to negating the definition of limit given above. Thus the negation is:

There exists an $\epsilon > 0$ such that for every $\delta > 0$, there exists an $x \in \mathcal{I}$ with $|x - x_0| < \delta$ such that $|f(x) - f(x_0)| \geq \epsilon$

[UC]: f is uniformly continuous on an interval $\mathcal{I}$ if, given $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(x) - f(y)| < \epsilon$ whenever $x, y \in \mathcal{I}$ and $|x - y| < \delta$.

The detailed logical transcription is: $\big(\forall \epsilon\big)\big\{\epsilon > 0 \implies \big(\exists \delta\big)\big[\delta > 0 \,\&\, \big(\forall x\big)\big(\forall y\big)\big(x \in \mathcal{I} \,\&\, y \in \mathcal{I} \,\&\, |x - y| < \delta \implies |f(x) - f(y)| < \epsilon\big)\big]\big\}$

The negation is: There exists an $\epsilon > 0$ such that, for every $\delta > 0$, there exist $x, y \in \mathcal{I}$ and $|x - y| < \delta$ but $|f(x) - f(y)| \geq \epsilon$

The detailed logical transcription of the negation is: $\big(\exists\epsilon\big)\big\{\epsilon > 0 \,\&\, \big(\forall\delta\big)\big[\delta > 0 \,\&\, \big(\exists x\big)\big(\exists y\big)\big(x \in \mathcal{I} \,\&\, y \in \mathcal{I} \,\&\, |x-y| < \delta \,\&\, |f(x) - f(y)| \geq \epsilon\big)\big]\big\}$

We can apply the negation of uniform continuity in order to prove Heine's Theorem which states that if a function f on a closed, boudned interval $[a,b]$ is continuous, it is uniformly continuous. We proceed by proof by contradiction so we assume that f is not uniformly continuous (but note that the antecedent conditions of being continuous on a closed, bounded interval are assumed true).

Since f is not uniformly continuous there exists an $\epsilon > 0$ and sequences $\{x_n\}$ and $\{y_n\}$ where x_n and y_n are in $[a,b]$ such that $|x_n - y_n| < \frac{1}{n}$ for each $n \in \mathbb{N}$, but $|f(x) - f(y)| \geq \epsilon$.

[Note here the structure: our universal quantifier is δ and the existential quantifier applies to the x and y points and $\frac{1}{n}$ plays the role of δ because of the universal $n \in \mathbb{N}$ requirement ie since the statement is to hold, in particular, for any $\delta > 0$ it must hold for δ of the form $\frac{1}{n}$ for any $n \in \mathbb{N}$]

By the Compactness Property of $\mathbb{R}$, the sequence $\{x_n\}$ has a convergent sub-sequence $\{x_{n_k}\}$. Let $x = \lim x_{n_k}$, then $x \in [a,b]$.

But for sub-sequence $\{y_{n_k}\}$ of $\{y_n\}$ we have:

$|x_{n_k} - y_{n_k}| < \frac{1}{n_k}$ for each $k \in \mathbb{N}$. Now because $|y_{n_k} - x| = |y_{n_k} - x_{n_k} + x_{n_k} - x| \leq |y_{n_k} - x_{n_k}| + |x_{n_k} - x|$ it follows that $y_{n_k} \to x$ since each of the factors on the RHS of the inequality can be made arbitrarily small for sufficiently large k.

Since f is continuous at x we have that $f(x_{n_k}) \to f(x)$. But this means that for any $\epsilon^* > 0$, there exists a $\nu \in \mathbb{N}$ such that:

$|f(x_{n_k}) - f(x)| < \epsilon^*$ for all $k > \nu$.

But:

$|f(y_{n_k}) - f(x)| = |f(y_{n_k}) - f(x_{n_k}) - (f(x) - f(x_{n_k})\,)|$
$\geq \big|\,|f(y_{n_k}) - f(x_{n_k})| - |f(x) - f(x_{n_k})|\,\big|$
$\geq |f(y_{n_k}) - f(x_{n_k})| - |f(x) - f(x_{n_k})|$
$> \epsilon - \epsilon^*$ for all $k > \nu$ since $|f(x_{n_k}) - f(x)| < \epsilon^*$ for all $k > \nu$.

So although $\{y_{n_k}\}$ converges to x, $\{f(y_{n_k})\}$ does not converge to $f(x)$, contradicting the fact that f is continuous at x. This is our contradiction and hence f is uniformly continuous.

7 A final remark on getting the order of quantifiers right

Recall that in the definition of a limit of a function we have the following formulation:

$$(\forall\epsilon)\left\{\epsilon > 0 \implies (\exists\delta)[\,\delta > 0\,\&\,(\forall x)\,(x \neq x_0\,\&\,0 < |x - x_0| < \delta \implies |f(x) - L| < \epsilon)\,]\right\} \tag{9}$$

It is important to understand that ϵ is chosen first and for each such choice a suitable δ can be chosen. There is a BIG difference between (9) and this:

$$(\exists\epsilon)\left\{\epsilon > 0 \implies (\forall\delta)[\,\delta > 0\,\&\,(\forall x)\,(x \neq x_0\,\&\,0 < |x - x_0| < \delta \implies |f(x) - L| < \epsilon)\,]\right\} \tag{10}$$

To see why consider the following two structurally similar forms :

(9*) $\big(\forall x\big)\big(\exists y\big)Hxy$ (this is essentially the structure of the correct epsilon - delta definition - it says that for any x there is some y such that the condition Hxy holds)

(10*) $\big(\exists y\big)\big(\forall x\big)Hxy$ (this is essentially the structure of the incorrect definition - it says that there is a y such that for any x the condition Hxy holds)

To see why (10) is incorrect let " Hxy " stand for " x and y are the same thing ". Then what (9*) says is : For every object x that is chosen there is an object y such that x and y are the same thing. Clearly there will be such an object which is the same as x , namely the chosen object itself. So (9) is true.

On the other hand, if the number of objects in the universe under consideration is greater than one, no one object can be the same as each ie no one object y can be such that $\big(\forall x\big)$ (x and y are the same thing). So (10*) is false.

The nub of the distinction is that '$\big(\forall x\big)\big(\exists y\big)Hxy$ ' says that once you fix on any x, you can find an object y such that Hxy holds. Different choices of x may give rise to different choices of y. However, '$\big(\exists y\big)\big(\forall x\big)Hxy$ ' says that you can find some y so that for all comers x, the condition Hxy holds.

There is a big difference between saying: "Every number is such that some number exceeds it" (essentially (9*)) and "Some number exceeds every number" (essentially (10*)), the latter clearly being false (there being no greatest number and even if there were it would not exceed itself).

8 References

[1] W. V. Quine, ”Methods of Logic” Revised 3rd Edition, Routledge and Kegan Paul,1974.