Notes on Mathematical Implication

by David Easdown, February 2012

Implication is used extensively throughout mathematics, and most noticeably in proofs and statements which involve the words

"if ... then" and "if and only if"

and corresponding symbols

$$\Longrightarrow \qquad \text{and} \qquad \Longleftrightarrow \ .$$

It is the purpose of these notes to explain precisely what we mean by *mathematical implication* and provide the student with an opportunity to develop some fluency in using it.

In mathematics there are two *truth values*, abbreviated T for *TRUE* and F for *FALSE*. A *proposition* is a statement which can meaningfully be assigned a truth value. It is a basic premise of mathematical logic that

> Propositions are either *TRUE* or *FALSE*,
> but never both simultaneously.

For example, the following is a true proposition:

Sydney is in Australia.

The following is false:

Sydney is not in Australia.

Of course, the words are subject to interpretation, and by "Sydney" we mean the city of Sydney, not some particular person whose name happens to be Sydney, who may or may not happen to be in Australia. The following are grammatical statements in English, but they are not propositions because they cannot be assigned truth values. The first is a question, and the second is the basis of the so-called *liar paradox*, violating our basic premise above by trying to be both true and false simultaneously:

Why are we here?

This sentence is false.

Here are two familiar mathematical propositions, the first of which is true:

$$2+2 \; = \; 4$$

and the second false:

$$2+2 \; \neq \; 4$$

The only difference is that "equals" in the first is replaced by "not equals" in the second, using the device of a stroke through the equality symbol. This is very common in mathematics: a stroke through a symbol can flip the truth value by representing the word "not". This leads to the first and simplest logical connective:

Negation: If P is a proposition then $\sim P$ (read "not P") is the *negation* of P and is

$$\begin{cases} \mathit{TRUE} & \text{when } P \text{ is } \mathit{FALSE} \\ \mathit{FALSE} & \text{when } P \text{ is } \mathit{TRUE}. \end{cases}$$

Thus $2+2 \neq 4$ is the negation of $2+2=4$, and the statement "Sydney is not in Australia" is the negation of "Sydney is in Australia".

The next two connectives combine two propositions into one:

Conjunction: If P and Q are propositions then $P \wedge Q$ (read "P and Q") is the *conjunction* of P and Q and is

$$\begin{cases} \mathit{TRUE} & \text{when } P \text{ and } Q \text{ are both } \mathit{TRUE} \\ \mathit{FALSE} & \text{when } \begin{cases} P \text{ is } \mathit{TRUE} \text{ and } Q \text{ is } \mathit{FALSE}, \text{ or} \\ P \text{ is } \mathit{FALSE} \text{ and } Q \text{ is } \mathit{TRUE}, \text{ or} \\ P \text{ is } \mathit{FALSE} \text{ and } Q \text{ is } \mathit{FALSE}. \end{cases} \end{cases}$$

Thus a conjunction can only be true if each of the component propositions is true, and is false otherwise.

Disjunction: If P and Q are propositions then $P \vee Q$ (read "P or Q") is the *disjunction* of P and Q and is

$$\begin{cases} \textit{TRUE} \quad \text{when} \begin{cases} P \text{ is } \textit{TRUE} \text{ and } Q \text{ is } \textit{TRUE}, \text{ or} \\ P \text{ is } \textit{TRUE} \text{ and } Q \text{ is } \textit{FALSE}, \text{ or} \\ P \text{ is } \textit{FALSE} \text{ and } Q \text{ is } \textit{TRUE} \end{cases} \\ \textit{FALSE} \quad \text{when } P \text{ and } Q \text{ are both } \textit{FALSE}. \end{cases}$$

Thus a disjunction can only be false if each of the component propositions is false. A disjunction will be true provided at least one of the component propositions is true.

(There is another connective called "exclusive or" which is only true if *exactly one* of the components is true. "Exclusive or" is rarely used in mathematics arguments without explicit mention. It is, however, common in ordinary language. For example, if you tell a child to choose an apple *or* an orange, the child will almost certainly interpret your instruction to take one or the other of the pieces of fruit but not both.)

A truth table displays the truth values for compound propositions given truth values for components. Here is the truth table for negation:

P	$\sim P$
T	F
F	T

and a combined truth table for both conjunction and disjunction:

P	Q	$P \wedge Q$	$P \vee Q$
T	T	T	T
T	F	F	T
F	T	F	T
F	F	F	F

Combinations of connectives can be built up in steps. For example, the compound proposition

$$\sim P \ \vee \ Q$$

can be interpreted as asserting either P is false or Q is true (or both). The following truth table lists its truth values, after an intermediate step:

P	Q	$\sim P$	$\sim P \ \vee \ Q$
T	T	F	T
T	F	F	F
F	T	T	T
F	F	T	T

Thus the only way $\sim P \vee Q$ can be false is if P is true and Q is false. In a sense we can make precise, the truth of P "entails" the truth of Q. Shortly we will discuss *implication* and the expression $P \implies Q$ and explain why this is naturally defined as an abbreviation for $\sim P \vee Q$.

> We say that (compound) propositions P and Q are *logically equivalent*, and write
>
> $$P \ \equiv \ Q$$
>
> if they produce the same truth values, that is, if they have identical columns in a truth table.

For example, for any proposition P we have

$$P \equiv \sim (\sim P)$$

because P is true precisely when $\sim P$ is false, which occurs precisely when $\sim (\sim P)$ is true:

P	$\sim P$	$\sim (\sim P)$
T	F	T
F	T	F

A more elaborate example is one of *De Morgan's Laws*:

$$\boxed{\sim (P \vee Q) \equiv \sim P \wedge \sim Q}$$

which follows from the fact that the third and last columns of truth values in the next table are identical:

P	Q	$P \vee Q$	$\sim (P \vee Q)$	$\sim P$	$\sim Q$	$\sim P \wedge \sim Q$
T	T	T	F	F	F	F
T	F	T	F	F	T	F
F	T	T	F	T	F	F
F	F	F	T	T	T	T

The other *De Morgan's Law*

$$\boxed{\sim (P \wedge Q) \equiv \sim P \vee \sim Q}$$

can be verified similarly.

Let's think now informally about the meaning of the proposition

$$\boxed{\textsf{if } P \textsf{ then } Q\,,}$$

which can be expressed symbolically by

$$\boxed{P \implies Q\,.}$$

In the above we call P the *hypothesis*, the arrow symbol *implies* and Q the *conclusion*. The hypothesis being true should somehow "force" the conclusion to be true. We would like the following

Rule of inference: If P is $TRUE$ and $P \implies Q$ is $TRUE$ then Q is $TRUE$.

Logicians call this rule of inference *Modus Ponens*, and it can be displayed as follows (though this is rather old-fashioned):

$$\begin{array}{rl} & P \\ & P \implies Q \\ \hline \therefore & Q \end{array}$$

This diagram is called a *syllogism* and the three dots to the left are an abbreviation of "therefore". Probably the most famous example is attributed to Aristotle:

$$\begin{array}{rl} & \text{Socrates is a man} \\ & \text{All men are mortal} \\ \hline \therefore & \text{Socrates is mortal} \end{array}$$

We can express this symbolically by putting

$$P(X) \equiv X \text{ is a man}$$

$$Q(X) \equiv X \text{ is mortal}$$

and letting the upside-down symbol $\forall$ represent "for all":

$$\begin{array}{rl} & P(\text{Socrates}) \\ & (\forall X)\ \ P(X) \implies Q(X) \\ \hline \therefore & Q(\text{Socrates}) \end{array}$$

This is a variation of Modus Ponens involving the "universal quantifier" $\forall$. If we remove it by taking the special case $X \equiv \text{Socrates}$ at the second line, then this becomes the usual form of Modus Ponens:

$$\begin{array}{rl} & P(\text{Socrates}) \\ & P(\text{Socrates}) \implies Q(\text{Socrates}) \\ \hline \therefore & Q(\text{Socrates}) \end{array}$$

Now, the expression $P \Longrightarrow Q$ should be some kind of compound proposition built from P and Q. The connective $\Longrightarrow$ must have a truth table. We will find this table in a couple of steps, by thinking about properties implication should have. For example, the following implication is evidently true:

$$\textit{student is listening attentively} \quad \Longrightarrow \quad \textit{student is awake}$$

However, a related implication, where we interchange the hypothesis and conclusion, need not be true:

$$\textit{student is awake} \quad \Longrightarrow \quad \textit{student is listening attentively}$$

> The *converse* of the implication $P \Longrightarrow Q$ is the implication $Q \Longrightarrow P$ obtained by interchanging the hypothesis and conclusion.

The previous example relating the attentiveness and wakefulness of a student demonstrates that an implication and its converse should not be logically equivalent. Thus

$$\boxed{P \Longrightarrow Q \quad \not\equiv \quad Q \Longrightarrow P}$$

and so

The truth tables for $P \Longrightarrow Q$ and $Q \Longrightarrow P$ **must** be different!

We have yet another variation for transforming an implication:

The *contrapositive* of the implication $P \Longrightarrow Q$ is the implication $\sim Q \Longrightarrow \sim P$ obtained by negating and then interchanging the hypothesis and the conclusion.

If you think about the contrapositive in familiar settings you quickly become convinced that it should be logically equivalent to the original implication. For example, expressed as an "if ... then" statement, the following implication is evidently true:

If I am in Sydney then I am in Australia.

The contrapositive is also evidently true and carries the same information:

If I am not in Australia then I am not in Sydney.

If you agree that our logic requires

$$\boxed{P \Longrightarrow Q \;\equiv\; \sim Q \Longrightarrow \sim P}$$

then you are forced to conclude the following:

The truth tables for $P \Longrightarrow Q$ and $\sim Q \Longrightarrow \sim P$ **must** be the same!

We now set about constructing the truth table for $P \Longrightarrow Q$. There are two cases where there should be no controversy about the truth values. If P and Q are both true then $P \Longrightarrow Q$ should be true. If P is true whilst Q is false then $P \Longrightarrow Q$ should be false, for otherwise we have not captured the idea of the truth of P entailing the truth of Q.

The less intuitively obvious truth values for $P \Longrightarrow Q$ occur when P is false. In the table below we have called them X, for when Q is true, and Y, for when Q is false:

P	Q	$P \Longrightarrow Q$
T	T	T
T	F	F
F	T	X
F	F	Y

As a first step, these values now imply the following table for the contrapositive $\sim Q \Longrightarrow \sim P$, expressed in terms of X and Y:

P	Q	$\sim Q$	$\sim P$	$\sim Q \Longrightarrow \sim P$
T	T	F	F	Y
T	F	T	F	F
F	T	F	T	X
F	F	T	T	T

The last column of truth values has to coincide with the values for the table for $P \Longrightarrow Q$, which forces

$$Y = T\,.$$

We can revise our truth table further and add a column for the converse $Q \Longrightarrow P$:

P	Q	$P \Longrightarrow Q$	$Q \Longrightarrow P$
T	T	T	T
T	F	F	X
F	T	X	F
F	F	T	T

If $X = F$ then the last two columns have identical truth values, which will contradict that implication and its converse are not logically equivalent. Hence, we are forced to conclude

$$X = T.$$

Here then is the final truth table for implication:

P	Q	$P \Longrightarrow Q$
T	T	T
T	F	F
F	T	T
F	F	T

Notice then, from the truth table of an earlier example, that

$$\boxed{P \Longrightarrow Q \;\equiv\; \sim P \vee Q\,.}$$

If you are ever in doubt, simply think of an implication as asserting either that the hypothesis is false or the conclusion is true (or both).

Intuitively, it may seem strange that $P \Longrightarrow Q$ has a truth value at all when P is false, since there seems to be no "connection" with the truth value of Q. However, mathematical logic is coherent and requires assigning truth values to every proposition. The above analysis demonstrates that the truth values are forced, even in these less intuitive cases.

When P is false, mathematicians say that the implication $P \Longrightarrow Q$ is *trivially true* or *trivially satisfied*.

For example, both of the following implications are trivially true, regardless of the conclusion, since the hypothesis is obviously false:

$$0 = 1 \quad \Longrightarrow \quad \textit{the author of these notes is Superman}$$

$$0 = 1 \quad \Longrightarrow \quad \textit{the author of these notes is not Superman}$$

Typical examples in mathematics use a *quantifier*, such as the following:

$$(\forall x \in \mathbb{R}) \quad x^2 = 1 \quad \Longrightarrow \quad x = \pm 1$$

For a particular $x \in \mathbb{R}$ the implication holds, either trivially, when $x \neq \pm 1$, or by actual properties of the hypothesis, when $x = \pm 1$. By altering the hypothesis slightly, so that it is always false, we can make the statement trivially true with any conclusion we like:

$$(\forall x \in \mathbb{R}) \quad x^2 = -1 \quad \Longrightarrow \quad x \text{ is a cow jumping over the moon}$$

Notice in the previous two examples that the quantifier is linked to membership of the set $\mathbb{R}$ of real numbers. The overriding membership class is called the *universe of discourse*. If we change the universe of discourse to the set $\mathbb{C}$ of complex numbers, then the previous statement becomes false:

$$(\forall x \in \mathbb{C}) \quad x^2 = -1 \quad \Longrightarrow \quad x \text{ is a cow jumping over the moon}$$

We can make it true by adjusting the conclusion:

$$(\forall x \in \mathbb{C}) \quad x^2 = -1 \quad \Longrightarrow \quad x = \pm i$$

There is one final matter to discuss:

Double implication: Define $P \iff Q$ to be an abbreviation for the conjunction $(P \Longrightarrow Q) \wedge (Q \Longrightarrow P)$ and we say or write

"P if and only if Q",

which in turn may get abbreviated to

"P iff Q".

The words "only if" take a little getting used to. To say

$$P \text{ only if } Q$$

means

$$P \Longrightarrow Q\,.$$

Removing the word "only" changes the meaning dramatically. To say

$$P \text{ if } Q$$

means

$$Q \implies P\,,$$

which may also be written

$$P \impliedby Q\,.$$

Putting the two arrows $\implies$ and $\impliedby$ together gives the double arrow or double implication $\iff$. Using the truth table for conjunction, we can work out its truth table in steps:

P	Q	$P \implies Q$	$Q \implies P$	$(P \implies Q) \wedge (Q \implies P)$
T	T	T	T	T
T	F	F	T	F
F	T	T	F	F
F	F	T	T	T

More compactly, we have:

P	Q	$P \iff Q$
T	T	T
T	F	F
F	T	F
F	F	T

This can be expressed in words:

> The double implication $P \iff Q$ is true precisely when P and Q have identical truth values, that is, P and Q are logically equivalent propositions.

Mathematicians also commonly refer to "necessary" and/or "sufficient" conditions when analysing mathematical relationships. The idea is that for

$$P \Longrightarrow Q$$

to be true, the truth of P should be *sufficient* to guarantee the truth of Q, and we say

P is a sufficient condition for Q.

But here, also, the truth of Q is *necessary* given that we know the truth of P, and we say

Q is a necessary condition for P.

From the point of view of mathematical logic, both of the preceding wordings carry the same information. The choice of words is often a question of emphasis in a given problem.

An ideal situation sought by mathematicians is when one phenomenon is *characterised* in terms of another. If the phenomena are P and Q, then to say P *is characterised by* Q means precisely that P is a necessary and sufficient condition for Q (and automatically vice versa). In terms of the above discussion, this happens when the double implication

$$P \iff Q$$

holds. Striking examples are ubiquitous in mathematics, e.g. the following beautiful characterisation of invertibility of a square matrix M:

M is invertible if and only if $\det M \neq 0$,

or, symbolically,

$$M^{-1} \textit{ exists} \iff \det M \neq 0 .$$

If one says that Q *is a necessary but not sufficient condition for* P, then technically this means that $P \Longrightarrow Q$ holds, but $Q \Longrightarrow P$ fails. An example of this would be

P is the property that a square matrix M has a row of zeros

and

Q is the property that a square matrix M has zero deteminant.

Here, the implication $Q \implies P$ can be made to fail by choosing, say,

$$M = \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}.$$

Then M has zero determinant but does not possess a row of zeros. We call M a *counterexample* to the implication $Q \implies P$. It provides an instance where Q is true whilst P is false. We can also say that P *is a sufficient but not necessary condition for* Q.

As a final illustration, consider the following conditions:

(1) Students who study Logic and Foundations have a truly perceptive view of the world.

(2) No person leading an unfulfilling life can be happy.

(3) People who are unhappy are unable to grasp a truly perceptive view of the world.

(4) Students who study Logic and Foundations lead fulfilling lives.

We will explain why condition (4) is necessary for conditions (1), (2) and (3). Therefore, if you agree that (1), (2) and (3) are plausible, and you have studied Logic and Foundations, then you are guaranteed to lead a fulfilling life!

To analyse the morass of information in these statements using our knowledge of implication, we introduce the following symbols:

$$\begin{aligned} S(X) &\equiv X \text{ is a student studying Logic and Foundations} \\ P(X) &\equiv X \text{ has a truly perceptive view of the world} \\ F(X) &\equiv X \text{ leads a fulfilling life} \\ H(X) &\equiv X \text{ is happy} \end{aligned}$$

Earlier we introduced the *universal quantifier* $\forall$ which means "for all". We also now need the *existential quantifier* $\exists$ which means "for some" or "there exists". Universally quantified statements are like gigantic conjunctions. Existentially quantified statements are like disjunctions. Analogues of De Morgan's Laws discussed earlier apply where $Q(X)$ is a proposition about X:

$$\boxed{\sim \big[(\forall X)\ \ Q(X)\big] \ \equiv\ (\exists X)\ \ \sim Q(X)}$$

and similarly

$$\boxed{\sim \big[(\exists X)\ \ Q(X)\big] \ \equiv\ (\forall X)\ \ \sim Q(X)}$$

The earlier statements can now be translated into symbols, ready then for processing:

(1)	$(\forall X)\ \ S(X) \implies P(X)$
(2)	$\sim \big[(\exists X)\ \ \sim F(X) \wedge H(X)\big]$
(3)	$(\forall X)\ \ \sim H(X) \implies \sim P(X)$
(4)	$(\forall X)\ \ S(X) \implies F(X)$

Proof that (4) follows from (1), (2) and (3): By applying De Morgan's Laws to (2), we get

$$(\forall X)\ \ \sim(\sim F(X)) \vee \ \sim H(X)\,,$$

which simplifies to

$$(\forall X)\ \ \sim H(X) \ \vee\ F(X)\,,$$

which in turn can be reformulated using implication as

(5) $\quad (\forall X)\ \ H(X) \implies F(X)\,.$

Using the contrapositive, we can reformulate (3) as

(6) $\quad (\forall X)\ \ P(X) \implies H(X)\,.$

Now let X be any given person. Suppose X is a student studying Logic and Foundations, that is, $S(X)$ holds. By (1), (5) and (6) we have the following chain of implications:

$$S(X) \implies P(X) \implies H(X) \implies F(X)$$

By three applications of Modus Ponens, we conclude that $F(X)$ holds. Thus the truth of $S(X)$ entails the truth of $F(X)$, so that we can conclude

$$(\forall X)\ \ S(X) \implies F(X)\,.$$

Thus (4) holds and the proof is complete. May students of Logic and Foundations indeed have fulfilling lives!