

Math 240 Tutorial Proof Techniques

1 Logical Operators and Truth Tables

Sentential Calculus

In the sentential calculus, there are 16 distinct binary logical operations. The principle operators we will need are

- conjunction $\&$,
- disjunction \vee ,
- implication/conditional \Rightarrow , and
- biconditional \Leftrightarrow .

We also have the unary operator negation \neg .

The truth tables for these operators are as follows.

P	$\neg P$
f	t
t	f

P	Q	$P \vee Q$	$P \& Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
f	f	f	f	t	t
f	t	t	f	t	f
t	f	t	f	f	f
t	t	t	t	t	t

If we have $P \Leftrightarrow Q$, then we say that P and Q are logically equivalent. We will need to use logically equivalent forms of formal statements in this course. Two sentential atoms are equivalent if their truth tables coincide. Consider the following examples.

P	Q	$\neg(P \vee Q)$	$\neg P \& \neg Q$
f	f	t	t
f	t	f	f
t	f	f	f
t	t	f	f

This shows that $\neg(P \vee Q)$ and $\neg P \& \neg Q$ are logically equivalent, that is, $\neg(P \vee Q) \Leftrightarrow (\neg P \& \neg Q)$.

Useful logical equivalencies are the following.

- $\neg(P \vee Q) \Leftrightarrow (\neg P \& \neg Q)$,
- $\neg(P \& Q) \Leftrightarrow \neg P \vee \neg Q$,
- $(P \Leftrightarrow Q) \Leftrightarrow ((P \Rightarrow Q) \& (Q \Rightarrow P))$,
- $\neg(P \Rightarrow Q) \Leftrightarrow (P \& \neg Q)$,
- $(P \Rightarrow Q) \Leftrightarrow (\neg Q \Rightarrow \neg P)$,
- $((P \& Q) \Rightarrow Q) \Leftrightarrow (P \Rightarrow (R \Rightarrow Q))$,
- $(P \vee (Q \& R)) \Leftrightarrow ((P \vee Q) \& (P \vee R))$, and
- $(P \& (Q \vee R)) \Leftrightarrow ((P \& Q) \vee (P \& R))$.

Predicate Calculus

In the sentential calculus, we are not interested in the form of the particular sentences P, Q, R , etc. In the predicate calculus, we are. The predicate calculus uses two quantifiers.

- For all \forall , and
- there exists \exists .

For the predicate calculus, we have some implied domain of definition, say D . For example, D might be a vector space, or all vector spaces, etc. A property \mathcal{P} is then a function transforming collections of objects from D to the logical sentences. Given the property \mathcal{P} and $x \in D$, the statement $\mathcal{P}x$ means “ x has the property \mathcal{P} .” The sentence $\forall x \mathcal{P}x$ means “every element $x \in D$ has the property \mathcal{P} .” This is true precisely in the case that $\mathcal{P}x$ is true for every $x \in D$. The sentence $\exists x \mathcal{P}x$ means “there is at least one element $x \in D$ that has the property \mathcal{P} .” This is true whenever we can find or construct at least one object $x \in D$ that has the property \mathcal{P} .

The variable x inside $\forall x \mathcal{P}x$ is said to be in the scope of or bound by the quantifier \forall . In the formula $\forall x \mathcal{P}x \& \mathcal{Q}y$, the variable x is bound by \forall , but the variable y is not bound by any quantifier. In such a case, the variable y is said to be free. In the expression $(\forall x \mathcal{P}x) \vee \mathcal{Q}x$, only the first occurrence of x is bounded; the second occurrence is free. Technically, these are different variables, and a better equivalent way to write it is as $(\forall x \mathcal{P}x) \vee \mathcal{Q}y$. Note that a proper logical formula can never have a variable that is bounded by multiple quantifiers.

We also need to distinguish between free variables and names. A free variable is allowed to range over all of D . A name, however, is a fixed element of D . The previous statement $(\forall x \mathcal{P}x) \vee \mathcal{Q}y$ has a different meaning if we take y to a name, i.e., a fixed element of D . Similar comments hold when we replace \forall with \exists in the above discussion.

An expression involving free variables is in general not a proper sentence unless we assume they are implicitly bound with \forall . For example, take D to be \mathbf{R} , let x and y be variables, and let c be a name. Consider

$$(x = y) \Rightarrow (x + c = y + c).$$

In this valid expression, x and y are free. By implicitly assuming they are bound by \forall , this is equivalent to

$$\forall x \forall y \left((x = y) \Rightarrow (x + c = y + c) \right).$$

We need to be very careful in identifying the free variables, bound variables, and the names.

In general, one cannot use the truth table method to validate logical expressions using quantifiers. However, there are some useful equivalencies that aren't too difficult to intuit.

- $\neg \forall x \mathcal{P}x \Leftrightarrow \exists x \neg \mathcal{P}x$ (it is not true that every element $x \in D$ has \mathcal{P} is the same as there is at least one element $x \in D$ which does not have \mathcal{P}),
- $\neg \exists x \mathcal{P}x \Leftrightarrow \forall x \neg \mathcal{P}x$ (there does not exist an element $x \in D$ with \mathcal{P} is the same as every element $x \in D$ does not have \mathcal{P}),
- $\forall x \forall y \mathcal{P}xy \Leftrightarrow \forall y \forall x \mathcal{P}xy$ (note that in general it cannot be assumed that $\mathcal{P}xy$ and $\mathcal{P}yx$ are the same),
- $\exists x \exists y \mathcal{P}xy \Leftrightarrow \exists y \exists x \mathcal{P}xy$,
- $\exists x; \mathcal{P}xx \Rightarrow \exists x \exists y \mathcal{P}xy$,
- $\forall x \mathcal{P}x \Rightarrow \exists x \mathcal{P}x$,
- $\exists x \forall y \mathcal{P}xy \Rightarrow \forall y \exists x \mathcal{P}xy$,
- $\left((\forall x \mathcal{P}x) \& (\forall y \mathcal{Q}y) \right) \Leftrightarrow \forall x \mathcal{P}x \& \mathcal{Q}x$,

- $\left((\exists x \mathcal{P}x) \vee (\exists y \mathcal{Q}y) \right) \Leftrightarrow \exists x \mathcal{P}x \vee \mathcal{Q}x$,
- $(\exists x \mathcal{P}x \& \mathcal{Q}x) \Rightarrow \left((\exists x \mathcal{P}x) \& (\exists y \mathcal{Q}y) \right)$, and
- $(\forall x \mathcal{P}x \vee \mathcal{Q}x) \Rightarrow \left((\forall x \mathcal{P}x) \vee (\forall y \mathcal{Q}y) \right)$.

As a final note, observe that $\forall x \forall y \left(\mathcal{P}x \Rightarrow (\mathcal{P}y \Rightarrow x = y) \right)$ means if any two elements of the domain D both have the property \mathcal{P} , then they are equal. This is another way of saying “there is a unique element $x \in D$ which has the property \mathcal{P} .” This is often simply denoted as $\exists!x \mathcal{P}x$.

2 Methods of Proof

We will need to be able to provide deductive proofs of logical sentences of the form $P \Rightarrow Q$, $P \Leftrightarrow Q$, $\forall x \mathcal{P}x$, and $\exists x \mathcal{P}x$.

Proof of Implications $P \Rightarrow Q$

There are three ways to prove an implication $P \Rightarrow Q$: direct proof, proof by contradiction, and proof by contraposition. We handle these each in turn.

- To prove $P \Rightarrow Q$ directly, means to assume P and use this to infer Q . This is usually the most natural method of proof.
- To prove $P \Rightarrow Q$ by contradiction means to assume its negation and then derive a contradiction. That is, we assume $\neg(P \Rightarrow Q)$, which is equivalent to $P \& \neg Q$, from which we derive a fallacy such as $1 = 0$. From this, our original assumption that $P \Rightarrow Q$ is false is incorrect, i.e., $P \Rightarrow Q$ is true.
- To prove $P \Rightarrow Q$ by contraposition simply means to provide a direct proof of $\neg Q \Rightarrow \neg P$.

Proof of Universal Statements $\forall x \mathcal{P}x$

To prove the universal statement $\forall x \mathcal{P}x$, we have to show that every element $x \in D$ has the property \mathcal{P} . To do this, we let c be an arbitrary name. We then show $\mathcal{P}c$ is true. Because c was an arbitrary fixed element of D , and because it was shown to have \mathcal{P} , it must be that every element of D has \mathcal{P} , that is, $\forall x \mathcal{P}x$ is true.

We CANNOT assume anything more than c is an arbitrary fixed element of D in showing that c has \mathcal{P} . If over the course of showing $\mathcal{P}c$ we also need to assume c has the property \mathcal{Q} , then we will NOT have shown $\forall x \mathcal{P}x$; rather, we will have shown $\forall x (\mathcal{Q}x \Rightarrow \mathcal{P}x)$.

The method of proof we have described is usually called universal generalization. It can take special forms such as the various proofs by induction in the case that D is the natural numbers $\mathbf{N} = \{0, 1, 2, \dots\}$. We will cover induction and strong induction in turn. Throughout, remember that we are assuming $D = \mathbf{N}$.

- To prove $\forall x \mathcal{P}x$ by induction, we must prove the following statements:
 - Base case: $\mathcal{P}0$ (show 0 has \mathcal{P}).
 - Inductive step: $\forall x (\mathcal{P}x \Rightarrow \mathcal{P}(x + 1))$ (if x has \mathcal{P} , then so does $x + 1$).

Having shown these, we can infer $\forall x \mathcal{P}x$.

- To prove $\forall x \mathcal{P}x$ by strong induction, we must prove the following statements:
 - Base case: $\mathcal{P}0$ (show 0 has \mathcal{P}).

- Inductive step: $\forall x \forall y (y < x \Rightarrow (\mathcal{P}y \Rightarrow \mathcal{P}x))$ (if every y smaller than x has \mathcal{P} , then so does x).

Having shown these, we can infer $\forall x \mathcal{P}x$.

Proof of Existence Statements $\exists x \mathcal{P}x$

To show $\exists x \mathcal{P}x$, we let c a name, i.e., an element of the domain of discourse D . We then show that c has \mathcal{P} to infer $\exists x \mathcal{P}x$. If over the course of showing $\mathcal{P}c$ we also assume that c has the property \mathcal{Q} , we will still be able to infer $\exists x \mathcal{P}x$ by the following string of implications

$$(\exists x \mathcal{P}x \& \mathcal{Q}x) \Rightarrow ((\exists x \mathcal{P}x) \& (\exists y \mathcal{Q}y)) \Rightarrow \exists x \mathcal{P}x.$$

3 An Example

We will work through an example to illustrate some the proof techniques outlined above.

Question

Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbf{R}^n$ be such that if $a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_m\vec{v}_m = 0$, for any collection a_1, a_2, \dots, a_m of scalars, then $a_1 = a_2 = \dots = a_m = 0$. Show that every vector in $\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$ has a unique representation as a linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$.

Logical Form

The domain of discourse is \mathbf{R}^n , the space of column vectors of length n with entries from \mathbf{R} . We will also be quantifying over elements from the subspace \mathbf{R} . So, for the sake of clarity we will include whatever domain we are quantifying over in the quantification statements, that is, we will write $\forall \vec{v}_i \in \mathbf{R}^n$ instead of simply $\forall \vec{v}_i$. We will also shorten statements like $\forall \vec{v}_1 \in \mathbf{R}^n \forall \vec{v}_2 \in \mathbf{R}^n \dots \forall \vec{v}_m \in \mathbf{R}^n$ to simply $\forall \vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbf{R}^n$. Denoting $V = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$, the logical form of the proposition we want to prove is then given by

$$\begin{aligned} & \left((\forall a_1, \dots, a_m \in \mathbf{R}) \mathcal{P}(\vec{v}_1, \dots, \vec{v}_m, a_1, \dots, a_m) \right) \\ & \Rightarrow \left((\forall \vec{u} \in V) (\exists! b_1, \dots, b_m \in \mathbf{R}) \mathcal{Q}(\vec{u}, \vec{v}_1, \dots, \vec{v}_m, b_1, \dots, b_m) \right) \end{aligned}$$

where we taken advantage of free variables to simplify the notation. We noe the following

- $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ are free variables ranging over \mathbf{R}^n ;
- $\mathcal{P}(\vec{v}_1, \dots, \vec{v}_m, a_1, \dots, a_m)$ means “if $a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_m\vec{v}_m = 0$, then $a_1 = a_2 = \dots = a_m = 0$ ”; and
- $\mathcal{Q}(\vec{u}, \vec{v}_1, \dots, \vec{v}_m, b_1, \dots, b_m)$ means “ \vec{u} is equal to $b_1\vec{v}_1 + b_2\vec{v}_2 + \dots + b_m\vec{v}_m$.”

Direct Proof

We assume $(\forall a_1, \dots, a_m \in \mathbf{R}) \mathcal{P}(\vec{v}_1, \dots, \vec{v}_m, a_1, \dots, a_m)$ holds, and we show $(\forall \vec{u} \in V) (\exists! b_1, \dots, b_m \in \mathbf{R}) \mathcal{Q}(\vec{u}, \vec{v}_1, \dots, \vec{v}_m, b_1, \dots, b_m)$ directly. Here is what is expected for a written proof.

We will prove the result directly. Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbf{R}^n$, and let $V = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$. Assume that if $a_1\vec{v}_1 + a_2\vec{v}_2 + \dots + a_m\vec{v}_m = 0$, for any collection a_1, a_2, \dots, a_m of scalars, then $a_1 = a_2 = \dots = a_m = 0$. We will show that this implies that every $\vec{u} \in V$ has a unique representation as a linear combination of the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$.

Indeed, suppose there are two collections a_1, a_2, \dots, a_m and b_1, b_2, \dots, b_m of scalars such that

$$u = a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_m \vec{v}_m = b_1 \vec{v}_1 + b_2 \vec{v}_2 + \dots + b_m \vec{v}_m.$$

Then

$$0 = \vec{u} - \vec{u} = (a_1 - b_1) \vec{v}_1 + (a_2 - b_2) \vec{v}_2 + \dots + (a_m - b_m) \vec{v}_m.$$

Our assumption then implies that

$$a_1 - b_1 = a_2 - b_2 = \dots = a_m - b_m = 0.$$

Equivalently,

$$a_1 = b_1, \quad a_2 = b_2, \quad \dots, \quad a_m = b_m.$$

But this is what it means for the representation of \vec{u} as a linear combination of $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ to be unique. We have therefore shown what we have set out to prove. QED

There are several components of the proof to identify:

Introduction 1. We start by saying what method of proof we are using. In this case, we are using a direct proof. This is where we also introduce the objects we are working with, namely, the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ in \mathbf{R}^n .

Assumptions 2. We must make explicit the assumptions (if any) we are using. For this problem, we are assuming the truth of $(\forall a_1, \dots, a_m \in \mathbf{R}) \mathcal{P}(\vec{v}_1, \dots, \vec{v}_m, a_1, \dots, a_m)$.

Derivation 3. This is the real derivation included in the proof of the statement. Here we used our assumption to show directly that $(\forall \vec{u} \in V)(\exists! b_1, \dots, b_m \in \mathbf{R}) \mathcal{Q}(\vec{u}, \vec{v}_1, \dots, \vec{v}_m, b_1, \dots, b_m)$ holds.

Conclusion 4. Once we have concluded the derivation, we must state what we have actually shown. You don't have to be needlessly explicit here; we just have to make clear what we've done in some way.

Proof by Contradiction

We assume the negation of what we want to show. In this case, this amounts to assuming $(\forall a_1, \dots, a_m \in \mathbf{R}) \mathcal{P}(\vec{v}_1, \dots, \vec{v}_m, a_1, \dots, a_m)$ AND $\neg(\forall \vec{u} \in V)(\exists! b_1, \dots, b_m \in \mathbf{R}) \mathcal{Q}(\vec{u}, \vec{v}_1, \dots, \vec{v}_m, b_1, \dots, b_m)$ both hold. We then use this assumption to derive a contradiction (think $1 = 0$). Having derived a contradiction, we can infer that our original assumption was incorrect, and the proposition holds true. We now show what is required in writing this proof.

Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbf{R}^n$, and let $V = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$. We will use proof by contradiction. Assume that if $a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_m \vec{v}_m = 0$, for any collection a_1, a_2, \dots, a_m of scalars, then $a_1 = a_2 = \dots = a_m = 0$, and assume that there is a vector $\vec{u} \in V$ such that \vec{u} has two distinct representations as a linear combination of the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$. We now derive a contradiction.

Let two distinct representations of \vec{u} be given by

$$u = a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_m \vec{v}_m = b_1 \vec{v}_1 + b_2 \vec{v}_2 + \dots + b_m \vec{v}_m,$$

for scalars a_1, a_2, \dots, a_m and b_1, b_2, \dots, b_m in \mathbf{R} , where it is required that there be at least one index $i \in \{1, 2, \dots, m\}$ for which $a_i \neq b_i$. It follows that

$$0 = \vec{u} - \vec{u} = (a_1 - b_1) \vec{v}_1 + (a_2 - b_2) \vec{v}_2 + \dots + (a_m - b_m) \vec{v}_m.$$

From our first assumption, we then have that

$$a_1 - b_1 = a_2 - b_2 = \dots = a_m - b_m = 0.$$

Equivalently,

$$a_1 = b_1, \quad a_2 = b_2, \quad \dots, \quad a_m = b_m.$$

This means there is no index $i \in \{1, 2, \dots, m\}$ for which $a_i \neq b_i$, contrary to what we have observed.

This contradiction implies that our original assumption was incorrect, namely, every vector in V has a unique representation as a linear combination of the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$. This is what we wanted to show. QED

Introduction 1. We start by saying what method of proof we are using. In this case, we are using a direct proof. This is where we also introduce the objects we are working with, namely, the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ in \mathbf{R}^n .

Assumptions 2. Since we are using proof by contradiction, we are assuming the negation of what we want to prove. Namely, we are assuming that $(\forall a_1, \dots, a_m \in \mathbf{R}) \mathcal{P}(\vec{v}_1, \dots, \vec{v}_m, a_1, \dots, a_m)$ is true but $(\forall \vec{u} \in V)(\exists! b_1, \dots, b_m \in \mathbf{R}) \mathcal{Q}(\vec{u}, \vec{v}_1, \dots, \vec{v}_m, b_1, \dots, b_m)$ is false.

Derivation 3. Having assumed the negation of what we wanted to prove, we now want to derive a contradiction. We do this by showing there both is and isn't an index for which the coefficients in the representations of some vector \vec{u} disagree.

Conclusion 4. Once we have concluded the derivation, we must state what we have actually shown. You don't have to be needlessly explicit here; we just have to make clear what we've done in some way.

For this example, the direct proof and the proof by contradiction “look” quite similar. One subtle difference is for the direct proof, we did not assume the two given presentations were unique; while in the proof by contradiction, we must assume the two representations are distinct. In general and for more involved derivations, these forms of proof look very different.

Proof by Contraposition

For this technique, we give a direct proof of the contrapositive of what we want to show, that is, we assume $(\forall \vec{u} \in V)(\exists! b_1, \dots, b_m \in \mathbf{R}) \mathcal{Q}(\vec{u}, \vec{v}_1, \dots, \vec{v}_m, b_1, \dots, b_m)$ is false and use this to show directly that $(\forall a_1, \dots, a_m \in \mathbf{R}) \mathcal{P}(\vec{v}_1, \dots, \vec{v}_m, a_1, \dots, a_m)$ must also be false.

Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m \in \mathbf{R}^n$, and let $V = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$. We will prove our result by showing the contrapositive. To this end, we assume that there is a vector $\vec{u} \in V$ which has two distinct representations as a linear combination of the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$. We will show that this implies there exists some linear combination of the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ which vanishes and for which not all the coefficients are zero.

Let two distinct representations of \vec{u} be given by

$$u = a_1 \vec{v}_1 + a_2 \vec{v}_2 + \dots + a_m \vec{v}_m = b_1 \vec{v}_1 + b_2 \vec{v}_2 + \dots + b_m \vec{v}_m,$$

for scalars a_1, a_2, \dots, a_m and b_1, b_2, \dots, b_m in \mathbf{R} , where it is required that there be at least one index $i \in \{1, 2, \dots, m\}$ for which $a_i \neq b_i$. It follows that

$$0 = \vec{u} - \vec{u} = (a_1 - b_1) \vec{v}_1 + (a_2 - b_2) \vec{v}_2 + \dots + (a_m - b_m) \vec{v}_m.$$

Since $a_i \neq b_i$, we have $a_i - b_i \neq 0$.

We have shown that if there is a vector in V with two distinct representations, then there is a vanishing linear combination of the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ where not all the coefficients are zero. This what we wanted to show. QED

Introduction 1. We start by saying what method of proof we are using. In this case, we are using a direct proof. This is where we also introduce the objects we are working with, namely, the vectors $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m$ in \mathbf{R}^n .

Assumptions 2. Since we are using contraposition, we assuming the negation of $(\forall \vec{u} \in V)(\exists! b_1, \dots, b_m \in \mathbf{R}) \mathcal{Q}(\vec{u}, \vec{v}_1, \dots, \vec{v}_m, b_1, \dots, b_m)$. and then use this to show the negation of $(\forall a_1, \dots, a_m \in \mathbf{R}) \mathcal{P}(\vec{v}_1, \dots, \vec{v}_m, a_1, \dots, a_m)$ must therefore hold.

Derivation 3. Having assumed the negation of $(\forall \vec{u} \in V)(\exists! b_1, \dots, b_m \in \mathbf{R}) \mathcal{Q}(\vec{u}, \vec{v}_1, \dots, \vec{v}_m, b_1, \dots, b_m)$, we use this to show the negation of $(\forall a_1, \dots, a_m \in \mathbf{R}) \mathcal{P}(\vec{v}_1, \dots, \vec{v}_m, a_1, \dots, a_m)$ must also hold.

Conclusion 4. Once we have concluded the derivation, we must state what we have actually shown. You don't have to be needlessly explicit here; we just have to make clear what we've done in some way.