The Substitution Theorem

How do we prove an equivalence principle such as: for any \( \varphi \) and \( \psi \):

\[
\neg(\varphi \lor \psi) \Leftrightarrow \neg\varphi \land \neg\psi?
\]

Easy: Take any \( \varphi \), \( \psi \), and truth-assignment \( h \):

\[
h(\neg(\varphi \lor \psi)) = \text{True iff (semantics } \neg) \]
\[
h(\varphi \lor \psi) = \text{False iff (semantics } \lor) \]
\[
h(\varphi) = \text{False and } h(\psi) = \text{False iff (semantics } \neg) \]
\[
h(\neg\varphi) = \text{True and } h(\neg\psi) = \text{True iff (semantics } \land) \]
\[
h(\neg\varphi \land \neg\psi) = \text{True} \]

So, \( \neg(\varphi \lor \psi) \Leftrightarrow \neg\varphi \land \neg\psi \)

Of course, in Introduction to Logic we made a quick combined truth-table to demonstrate the equivalence, but notice how that really isn’t quite the same: since this table would have 2 reference columns, we would (despite the fact that we would label these as \( \varphi \) and \( \psi \)) effectively be proving the equivalence only for the case where \( \varphi \) and \( \psi \) are both atomic sentences. Moreover, it is not immediately clear that if the principle is true for atomic sentences, it would also hold true for compound sentences: \( \varphi \) and \( \psi \) may be sharing certain atomic sentences, which would result in some kind of complex interaction of truth-conditions, and it is not clear that this interaction would result in the same overall truth-conditions on either side or the supposed equivalence. To illustrate this, suppose I were to make a truth-table for ‘\( \varphi \)’: I would immediately find that \( \varphi \) could be either True or False, and thus might be inclined to conclude that \( \varphi \) is a contingency. In other words, I would have ‘proven’ that every sentence is a contingency, which is clearly not the case: if \( \varphi \) were to be the statement \( A \lor \neg A \), \( \varphi \) is a tautology, not a contingency!

So there is a real danger of trying to say things about sentences in general through the use of a truth-table,

Fortunately, it turns out that once we have established an equivalence such as \( \neg(A \lor B) \Leftrightarrow \neg A \land \neg B \), we can substitute any other sentences for the atomic sentences \( A \) and \( B \) (compound or not), and retain the equivalence, thus establishing that for any \( \varphi \) and \( \psi \):

\[
\neg(\varphi \lor \psi) \Leftrightarrow \neg\varphi \land \neg\psi.
\]

This result is called the Substitution Theorem. So, one can therefore prove all the equivalences we saw before by first proving them for the case where the variables are atomic sentences, and then
using the Substitution Theorem to show that the atomic sentences can be used as sentence variables. The proof of the Substitution Theorem is far from trivial though, and we need a couple of Lemma’s to do this.

Remember that \( S(\varphi) \) refers to any sentence that contains zero or more instances of \( \varphi \) as a component sentence. Also, \( S(\psi/\varphi) \) is the result of substituting \( \psi \) for every \( \varphi \) in \( S(\varphi) \). In the context of \( S(\varphi) \), we’ll often just write this simply as \( S(\psi) \).

Substitution Lemma 1: For any \( \varphi \), \( S(\varphi) \), and \( h \):

1. If \( h(\varphi) = \text{True} \) then \( h(S(\varphi)) = h(S(\top)) \)
2. If \( h(\varphi) = \text{False} \) then \( h(S(\varphi)) = h(S(\bot)) \)

Proof:

The Lemma is true for any \( \varphi \) and \( S(\varphi) \) where \( S(\varphi) = \varphi \). For this means that \( S(\top) = \top \), and that \( S(\bot) = \bot \). So then for any \( h \):

1. If \( h(\varphi) = \text{True} \), then \( h(S(\varphi)) = h(\varphi) = \text{True} = h(\top) = h(S(\top)). \)
2. If \( h(\varphi) = \text{False} \), then \( h(S(\varphi)) = h(\varphi) = \text{False} = h(\bot) = h(S(\bot)).\)

Now let’s prove the result for any \( S(\varphi) \) that does contain at least one instance of \( \varphi \). We’ll do this by induction on structure of \( S \).

Base: \( S(\varphi) \) is atomic, say \( S(\varphi) = A \). Since \( S \) contains at least one instance of \( \varphi \), then that must mean that \( \varphi = A \), i.e. that \( S(\varphi) = \varphi \). We already proved the desired result for this case.

Step: Again, if \( S(\varphi) = \varphi \), then we have our desired result. So, let’s assume that any instances of \( \varphi \) in \( S(\varphi) \) occur as strict components of \( S(\varphi) \). Now, let’s just prove that it is true for any n-place operator \(*\), i.e. \( S(\varphi) = *(S_1(\varphi), S_2(\varphi), \ldots, S_n(\varphi)) \), for which we have some formal semantics defined, such that that there exists some truth-function \( f^* \) such that for any \( h \): \( h(*(S_1(\varphi), S_2(\varphi), \ldots, S_n(\varphi))) = f^*(h(S_1(\varphi)), h(S_2(\varphi)), \ldots, h(S_n(\varphi))) \). A little thought shows that for any substitution \( \psi \) for \( \varphi \), where any instances of \( \varphi \) in \( S(\varphi) \) occur as strict components of \( S(\varphi) \): \( S(\psi) = *(S_1(\psi), \ldots, S_n(\psi)) \).
S₂(ψ), …, Sₙ(ψ)). Now, the inductive hypothesis is that the Lemma is true for all Sᵢ(ϕ). So, take any h:

1. If h(ϕ) = True, then h(S(ϕ)) = h*(S₁(ϕ), S₂(ϕ), …, Sₙ(ϕ)) = f*(h(S₁(ϕ)), h(S₂(ϕ)), …, h(Sₙ(ϕ))) = (Inductive hypothesis) f*(h(S₁(⊤)), h(S₂(⊤)), …, h(Sₙ(⊤))) = h*(S₁(⊤), S₂(⊤), …, Sₙ(⊤)) = h(S(⊤))

2. If h(ϕ) = False, then h(S(ϕ)) = h*(S₁(ϕ), S₂(ϕ), …, Sₙ(ϕ)) = f*(h(S₁(ϕ)), h(S₂(ϕ)), …, h(Sₙ(ϕ))) = (Inductive hypothesis) f*(h(S₁(⊥)), h(S₂(⊥)), …, h(Sₙ(⊥))) = h*(S₁(⊥), S₂(⊥), …, Sₙ(⊥)) = h(S(⊥))

We can now prove the Substitution Lemma 2: For any atomic sentence A, S₁(A), and S₂(A):

If S₁(A) ⇔ S₂(A), then:

1. S₁(⊤) ⇔ S₂(⊤)
2. S₁(⊥) ⇔ S₂(⊥)

Notice that in substituting ⊤ or ⊥ for every A, there are no longer any instances of A. In other words, S₁(⊤), S₁(⊥), S₂(⊤), and S₂(⊥) do not contain any instances of A. So, take any h defined for S₁(⊤), S₁(⊥), S₂(⊤), and S₂(⊥), which is therefore not defined for A.

Since h is not defined for A, we can define a truth-assignment h’ for S₁(⊤), S₁(⊥), S₂(⊤), S₂(⊥), and A, such that h’(A) = True, and h’(P) = h(P) for every P occurring in S₁(⊤), S₁(⊥), S₂(⊤), S₂(⊥). Using Substitution Lemma 1, we thus have that h’(S₁(A)) = h’(S₁(⊤)), and that h’(S₂(A)) = h’(S₂(⊤)). But since S₁(A) ⇔ S₂(A), we know that for every truth-assignment h for S₁(A) and S₂(A): h(S₁(A)) = h(S₂(A)). Since h’ is a truth-assignment for S₁(A) and S₂(A), we thus have that h’(S₁(A)) = h’(S₂(A)). Hence, we have that h’(S₁(⊤)) = h’(S₂(⊤)). Since h’ extends h, and h is defined for S₁(⊤) and S₂(⊤), we thus have that h(S₁(⊤)) = h(S₂(⊤)). Since h was arbitrary, this
means that $S_1(\top) \iff S_2(\top)$. Similarly (by extending $h$ into $h'$ such that $h'(a) = \text{False}$), we can show that $S_1(\bot) \iff S_2(\bot)$.

Finally, let’s prove the Substitution Theorem: For any $A$, $\varphi$, $S_1(A)$, and $S_2(A)$:

If $S_1(A) \iff S_2(A)$, then $S_1(\varphi) \iff S_2(\varphi)$.

Proof: Assume $S_1(A) \iff S_2(A)$. By Substitution Lemma 2, we thus have that $S_1(\top) \iff S_2(\top)$ and $S_1(\bot) \iff S_2(\bot)$. Now take any $h$ for $S_1(\varphi)$ and $S_2(\varphi)$. This is automatically defined for $S_1(\top)$, $S_1(\bot)$, $S_2(\top)$, and $S_2(\bot)$. It is also defined for $\varphi$, so $h(\varphi) = \text{True}$ or $h(\varphi) = \text{False}$. If $h(\varphi) = \text{True}$, then (by Substitution Lemma 1) $h(S_1(\varphi)) = h(S_1(\top))$, and $h(S_2(\varphi)) = h(S_2(\top))$. But since $S_1(\top) \iff S_2(\top)$, and $h$ is defined for $S_1(\top)$ and $S_2(\top)$, we have that $h(S_1(\top)) = h(S_2(\top))$. Hence, $h(S_1(\varphi)) = h(S_2(\varphi))$. Similarly, if $h(\varphi) = \text{False}$, then $h(S_1(\varphi)) = h(S_1(\bot)) = h(S_2(\bot)) = h(S_2(\varphi))$. So, in either case, $h(S_1(\varphi)) = h(S_2(\varphi))$. Since $h$ was arbitrary, we thus have that $S_1(\varphi) \iff S_2(\varphi)$.

Notice that the Substitution Theorem does not state that for any $\varphi$, $\psi$, $S_1(\varphi)$, and $S_2(\varphi)$: If $S_1(\varphi) \iff S_2(\varphi)$, then $S_1(\psi) \iff S_2(\psi)$. Indeed, it better not, because this is simply not true (e.g. if we substitute $C$ for the compound sentence $A \lor B$ in the equivalence $\neg(A \lor B) \iff \neg A \land \neg B$, we get $\neg C \iff \neg A \land \neg B$, which is patently false). Indeed, the proof no longer works for compound sentences: while the substitution of $\top$ or $\bot$ for every atomic sentence $A$ gets rid of all the $A$’s (and hence we could define an extended truth-assignment $h'$ with the desired properties), substituting $\top$ or $\bot$ for a compound sentence such as $A \lor B$ does not necessarily get rid of the $A$’s nor the $B$’s, and thus the existence of any extended truth-assignments is no longer guaranteed.

So, the Substitution Theorem can be used to prove general equivalence principles. However, this is going about it in a rather round-about way since, as we saw, using formal semantics, these equivalences can be established much more directly and quickly. Still, the Substitution Theorem has another important application, and that is that it can be used to prove the Duality Theorem.
The Duality Theorem

You’ll have noticed that there is a kind of symmetry in the equivalence results regarding \( \land \) and \( \lor \). For example, both \( \land \) and \( \lor \) are commutative and associative. The Idempotence laws hold for both of them, and we don’t only have that \( \land \) distributes over \( \lor \), but also that \( \lor \) distributes over \( \land \) (as opposed to, say, \( + \) and \( * \) for numbers: \( * \) distributes over \( + \), but \( + \) does not distribute over \( * \). So, \( + \) and \( * \) do not have the kind of symmetry that \( \land \) and \( \lor \) have). The DeMorgan’s Laws, Absorption Laws, and Reduction Laws, etc. reveal this kind of symmetry even more clear: whenever we have some equivalence principle for \( \land \), there is a corresponding one for \( \lor \).

In fact, what exactly is this ‘corresponding’ or ‘dual’ principle? Well, it seems that if we have some equivalence principle that states that \( \varphi \iff \psi \), then if we systematically replace all \( \land \)’s with \( \lor \)’s, and all \( \lor \)’s with \( \land \)’s in the formulation of that equivalence principle, we obtain a new equivalence principle. For example, with one DeMorgan’s law stating that \( \neg(\varphi \land \psi) \iff \neg\varphi \lor \neg\psi \), we get as its dual principle the other DeMorgan’s law stating that \( \neg(\varphi \lor \psi) \iff \neg\varphi \land \neg\psi \).

As you might suspect, having these dual principles for DeMorgan, Distribution, Absorption, etc. is not a coincidence, as they all follow from a general Duality Theorem. To state this theorem, we first define what it means for two (binary) operators to be dual operators:

**Def:** Two operators \( \times \) and \( \times^D \) are dual operators iff for any \( \varphi \) and \( \psi \): \( \neg(\varphi \times \psi) \iff \neg\varphi \times^D \neg\psi \)

Notice that if for any \( \varphi \) and \( \psi \): \( \neg(\varphi \times \psi) \iff \neg\varphi \times^D \neg\psi \), then it is also true that for any \( \varphi \) and \( \psi \): \( \neg(\varphi \times^D \psi) \iff \neg\varphi \times \neg\psi \). This is because if for any \( \varphi \) and \( \psi \): \( \neg(\varphi \times \psi) \iff \neg\varphi \times^D \neg\psi \), then we also have for any \( \varphi \) and \( \psi \) that \( \neg(\neg\varphi \times \neg\psi) \iff \neg\varphi \times^D \neg\psi \). Hence, we have for any \( \varphi \) and \( \psi \) that: \( \neg(\varphi \times^D \psi) \iff \neg(\neg\varphi \times \neg\psi) \iff \neg\varphi \times \neg\psi \).

It is clear that \( \land \) and \( \lor \) are dual operators: the DeMorgan principles exactly fit the definition. However, there are other pairs of dual operators as well: nand and nor, iff and xor, ‘only if’ and ‘nif’, ‘if’ and ‘nonly if’, and the binary operators \( \top \) (i.e. the one that always returns True) and \( \bot \). The duals of id1, id2, nid1, and nid2 are themselves.
The fact that if \(-(\varphi \times \psi) \iff -\varphi \times^D -\psi\), then \(-(\varphi \times^D \psi) \iff -\varphi \times -\psi\) is one way in which the duality, or symmetry, of dual operators plays out, but the Duality Theorem is more general than that. To state this general principle, we need to provide a general definition of the dual sentence \(\varphi^D\) of any sentence \(\varphi\). Specifically, with \(\times\) and \(\times^D\) being dual operators, and where \(\varphi\) is any sentence that has \(-, \times,\) and \(\times^D\) as its only operators, we define \(\varphi^D\) to be the dual sentence of \(\varphi\) if it is the result of systematically switching all \(\times\) and \(\times^D\) operators in \(\varphi\). As a recursive definition:

1. \(\varphi^D = \varphi\) if \(\varphi\) is an atomic sentence
2. \((\varphi \times \psi)^D = (\varphi^D \times^D \psi^D)\)
3. \((\varphi \times^D \psi)^D = (\varphi^D \times \psi^D)\)
4. \((-\varphi)^D = -\varphi^D\)

Intuitively, if we take the dual sentence of \(\varphi^D\) we would get back \(\varphi\), since all switches between operators \(\times\) and \(\times^D\) when going from \(\varphi\) to \(\varphi^D\) will be switched back when taking the dual of \(\varphi^D\). More formally, we have as a theorem that for any sentence: \((\varphi^D)^D = \varphi\), and this can be proven by induction on the formation of \(\varphi\) (this is left to the reader as an exercise).

Having defined dual operators, and dual sentences based on dual operators, we can now state the:

Duality Theorem: For any \(\varphi\) and \(\psi\): if \(\varphi \iff \psi\), then \(\varphi^D \iff \psi^D\).

How can prove the Duality Theorem? We can’t use direct mathematical induction here, since the problem is that we are dealing with 2 sentences, \(\varphi\) and \(\psi\), rather than a single sentence \(\varphi\). So, we’ll have to try something different instead.

First, we’ll define the sentence \(\varphi'\) to be the complement of any sentence \(\varphi\) to be the sentence that one obtains by putting a negation in front of every atomic sentence occurring in \(\varphi\). Formally:

1. \(\varphi' = -\varphi\) if \(\varphi\) is an atomic sentence
2. \((-\varphi)' = -\varphi'\)
3. \((\varphi \times \psi)' = (\varphi' \times \psi')\) where \(\times\) is any binary operator
We can now prove the following Duality-Negation-Complement Lemma:

For every sentence $\varphi$ that has $\neg$, $\times$, and $\times^D$ as its only operators: $\varphi^D \iff \neg \varphi'$

Proof: By induction on the structure of $\varphi$

Base: $\varphi$ is atomic, say $\varphi = A$. Then $\varphi^D = A$ and $\varphi' = \neg A$, so indeed $\varphi^D = A \iff \neg \neg A = \neg \varphi'$

Step: Since $\varphi$ has $\neg$, $\times$, and $\times^D$ as its only operators, we (only) have to consider the following 3 corresponding cases, where the inductive hypothesis is that the Lemma holds for its components:

1. $\varphi = \neg \psi$. Then $\varphi^D = (\neg \psi)^D = \neg \psi^D$. Moreover, $\varphi' = (\neg \psi)' = \neg \psi'$. By inductive hypothesis, $\psi^D \iff \neg \psi'$, and therefore $\neg \psi^D \iff \neg \neg \psi'$. So, $\varphi^D = \neg \psi^D \iff \neg \varphi'$

2. $\varphi = \psi_1 \times \psi_2$. Then $\varphi^D = (\psi_1 \times \psi_2)^D = (\psi_1^D \times^D \psi_2^D)$. Also, $\varphi' = (\psi_1 \times \psi_2)' = (\psi_1' \times \psi_2')$, so $\neg \varphi' = -(\psi_1' \times \psi_2')$. By inductive hypothesis, $\psi_1^D \iff \neg \psi_1'$, and $\psi_2^D \iff \neg \psi_2'$, so $\varphi^D = (\psi_1^D \times^D \psi_2^D) \iff (\neg \psi_1' \times^D \neg \psi_2') \iff (\times$ and $\times^D$ are duals) $\neg (\psi_1' \times \psi_2') = \neg \varphi'$

3. $\varphi = \psi_1 \times^D \psi_2$. Then $\varphi^D = (\psi_1 \times^D \psi_2)^D = (\psi_1^D \times \psi_2^D) \iff (\text{inductive hypothesis}) (\neg \psi_1' \times \neg \psi_2') \iff -(\psi_1' \times^D \psi_2') = \neg \varphi'$

Now we use the Substitution Theorem to prove the Complement Lemma:

For every sentence $\varphi$ and $\psi$: If $\varphi \iff \psi$ then $\varphi' \iff \psi'$

Proof: For any atomic sentence occurring in $\varphi$ or $\psi$, we can say that $\varphi = S_1(A)$ and $\psi = S_2(A)$. By the Substitution Theorem, it follows that $S_1(\neg A) \iff S_2(\neg A)$. Repeating this for all atomic sentences in $\varphi$ or $\psi$, we thus not only obtain $\varphi'$ and $\psi'$, but also that $\varphi' \iff \psi'$

The Duality Theorem immediately follows from the Duality-Negation-Complement Lemma and the Complement Lemma: Take any $\varphi$ and $\psi$, and assume $\varphi \iff \psi$. By the Duality-Negation-Complement Lemma: $\varphi^D \iff \neg \varphi'$ and $\psi^D \iff \neg \psi'$. By the Complement Lemma: $\varphi' \iff \psi'$, and therefore $\neg \varphi' \iff \neg \psi'$. So, $\varphi^D \iff \neg \varphi' \iff \neg \psi' \iff \psi^D$. 
It will come as no surprise that the Duality Theorem itself works in the opposite way as well. That is, we don’t just have that for any \( \phi \) and \( \psi \): if \( \phi \iff \psi \), then \( \phi^D \iff \psi^D \), but we can also show that for any \( \phi \) and \( \psi \): if \( \phi^D \iff \psi^D \), then \( \phi \iff \psi \). To see this, notice that from the Duality Theorem it immediately follows that for any \( \phi \) and \( \psi \): if \( \phi^D \iff \psi^D \), then \( (\phi^D)^D \iff (\psi^D)^D \). Hence, since for any \( \phi \): \( (\phi^D)^D = \phi \), we have that for any \( \phi \) and \( \psi \): if \( \phi^D \iff \psi^D \), then \( \phi = (\phi^D)^D \iff (\psi^D)^D = \psi \).

So, are we there yet? Is the duality of the DeMorgan’s Laws a straightforward instantiation of the Duality Theorem? Well, not quite. With one DeMorgan’s law stating that \( \neg(\phi \land \psi) \iff \neg\phi \lor \neg\psi \), the Duality Theorem gives us \( (\neg(\phi \land \psi))^D \iff (\neg\phi \lor \neg\psi)^D \), from which we get that for any \( \phi \) and \( \psi \): \( \neg(\phi^D \lor \psi^D) \iff \neg\phi \land \neg\psi^D \), but this is not the same as the other DeMorgan’s stating that for any \( \phi \) and \( \psi \): \( \neg(\phi \lor \psi) \iff \neg\phi \land \neg\psi \). However, from the first DeMorgan’s principle we can infer that for any \( \phi \) and \( \psi \): \( \neg(\phi^D \land \psi^D) \iff \neg\phi^D \lor \neg\psi^D \), so by the Duality Theorem we get that for any \( \phi \) and \( \psi \): \( (\neg(\phi^D \land \psi^D))^D \iff (\neg\phi^D \lor \neg\psi^D)^D \). Hence, for any \( \phi \) and \( \psi \): \( \neg(\phi \lor \psi) = \neg((\phi^D \lor \psi^D)^D) = (\neg(\phi^D \land \psi^D))^D \iff (\neg\phi^D \lor \neg\psi^D)^D = (\neg(\phi^D \lor \neg\psi^D))^D = (\neg(\phi^D)^D \land \neg(\psi^D)^D) = \neg\phi \land \neg\psi \). So, the duality of DeMorgan does quickly follow from the Duality Theorem, and the same is true for all other pairs of Boolean equivalence laws, which we can accordingly call each other’s ‘dual’.

Moreover, as stated before, \( \land \) and \( \lor \) are not the only pair of dual operators. Hence, lots of other logical principles come in pairs. For example, since \( \leftrightarrow \) and XOR are duals, we have that since \( \leftrightarrow \) is associative \( (\phi \leftrightarrow \psi) \leftrightarrow \lambda \iff \phi \leftrightarrow (\psi \leftrightarrow \lambda) \), the XOR must be (and indeed is) as well.

Also, we can define generalized definitions of dual operators and sentences for any associative operators, and still obtain the Duality Theorem. For example, all dual principles will still hold for conjunctions or disjunctions with any number of terms. In fact, in the case of \( \top \) and \( \bot \), we can define duals and complements \( (\top^D = \bot, \bot^D = \top, \top' = \top, \bot' = \bot) \), and hence recognize a principle like \( P \land \top \iff P \) as the ‘dual’ of \( P \lor \bot \iff P \), and so on.

Finally, the Duality Theorem can be made stronger so it is no longer just about equivalences, but about logical implication in general. That is, one can prove that for any \( \phi \) and \( \psi \): if \( \phi \Rightarrow \psi \), then \( \psi^D \Rightarrow \phi^D \). Hence, the ‘dual’ of \( \phi \land \psi \Rightarrow \phi \) is \( \phi \Rightarrow \phi \lor \psi \), the dual of \( \bot \Rightarrow \phi \) is \( \phi \Rightarrow \top \), etc.