# Are werewolves vegan?

Oliver Snellman
oliver.snellman@helsinki.fi

January 13, 2022

**Abstract**

I first introduce the concept of implication in propositional logic, then present a puzzling consequence of it called Hempel's paradox and provide an intuition for why it actually does make sense. Lastly, I point attention to situations where the Hempel's "paradox" becomes paradoxical again, due to poorly defined propositions.

## What does implication mean?

Material implication is a rule of inference in propositional logic, which allows for the concept of implication between propositions p and q to be expressed as a truth functional operator of the arguments p and q. This means, that whether p implies q or not, depends on the truth values of the propositions p and q. I will refer to material implication as just implication from now on.

Consider two unspecified propositions p and q (a practical example in the next section), called the antecedent and consequent, respectively. They both have binary and mutually exclusionary truth values true (T) or false (F), meaning that a proposition must be either true or false. There are four possible relations between propositions p and q, based on the combinations of their truth values. Implication is defined as an operator, which form new propositions with truth values T or F, on the basis of its arguments p and q, in the following way.

| p | q | Relation | Implication $\neg p \vee q$ |
|---|---|---|---|
| T | T | $p \wedge q$ | T |
| F | F | $\neg p \wedge \neg q$ | T |
| F | T | $\neg p \wedge q$ | T |
| T | F | $p \wedge \neg q$ | F |

From the above truth table we can see, that the implication operator $[p \implies q]$ is only evaluated as False, when $\neg[p \wedge \neg q]$[1]. Hence, in all of the other cases it holds that p implies q, that is $\Big[[p \wedge q] \vee [\neg p \wedge q] \vee [\neg p \wedge \neg q]\Big]$ which is the same as $[\neg p \vee q]$.

[1]The symbol $\neg$ means "not", $\wedge$ means "and" and $\vee$ means "or".

In other words, the conditional operator "implies", $p \implies q$, is interpreted as stating, that it is not possible that the consequent proposition q is true when the antecedent proposition p is false $\neg[p \wedge \neg q]$.

$$[p \implies q] \text{ means } [\neg p \vee q] \qquad \text{Implication}$$

## All ravens are black

Let's specify the propositions as p=Ravens=:r and q=Black=:b. The set of all existing entities is denoted with H, where each entity has a combination of two truth values corresponding to propositions r and b, i.e. a black cat would belong to $\neg r \wedge b$ and a blue raincoat would be in $\neg r \wedge \neg b$.

Now the hypothesis "All ravens are black" can be expressed as an implication $r \implies b$, which reads "if something is a raven, then it's also black". By inductive reasoning, observing ravens which are black constitute evidence for this statement.

But this statement can also be expressed in logically equivalent way with contraposition as $\neg b \implies \neg r$ meaning "if something is not black, then it's not a raven" (notice, that this still expresses exactly the same idea). Now observing things that are not black, which are also not ravens, such as red apples, would constitute evidence for this statement. But because this is logically still the same statement about ravens being black, then observing a red apple is also evidence in favor of all ravens being black, hence, giving rise to the paradox.

Think of the situation where p is false, cases $[\neg p \wedge q]$ and $[\neg p \wedge \neg q]$ in the truth table, which both satisfy the given definition of implication. Let's say that you observe $\neg p$ and $q$. Would you find it reasonable to then state "this is evidence in favor of the hypothesis that p implies q"?

This question was presented in a thought experiment called Hempel's raven, which is sometimes characterized as a paradox. According to this "paradox", we can increase our likelihood of a hypothesis "All ravens are black", by observing a non-black non-raven, for example a red apple.

To make sense of this, let's draw a Venn-diagram for the situation at hand, Figure 1. Additional to black ravens $r \wedge b$, the world H also (logically) consists of black non-ravens $[\neg r \wedge b]$, non-black non-ravens $[\neg r \wedge \neg b]$ and non-black ravens $[r \wedge \neg b]$. The only situation what the hypothesis $r \implies b$ claims to be impossible, is observing a raven which is not black $r \wedge \neg b$.

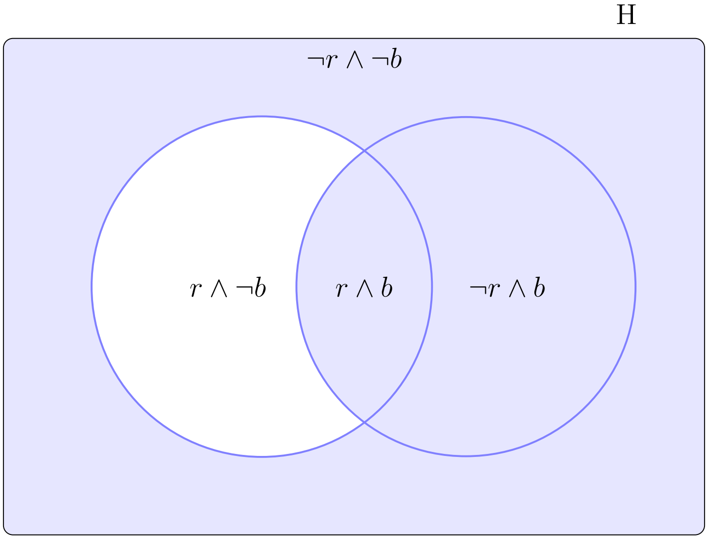


Figure 1: A venn diagram of propositions r and b, with blue color indicating the areas where implication $r \implies b$ holds and white where it does not. The only situation where the statement "All ravens are black" turns out to be false is when observing a raven which is not black $r \wedge \neg b$.

Observing even one raven, which is not black, would falsify the proposed implication. Therefore, in order for the implication to hold, the set of non-black ravens have to be empty in the world H, $r \wedge \neg b = \emptyset$.

Let's assume first, that the set of all existing entities is finite, $|H| =: n < \infty$. We could then check that the implication holds, by going through every existing entity and confirming that there indeed are no non-black ravens, $r \wedge \neg b = \emptyset$. This can be thought of as randomly sampling from H without replacement for n draws, each time finding out the truth values of r and b for the new draw.

Before we engaged in this laborious process, we didn't know whether the proposed implication would hold or not. But after we observed everything that exists without finding any non-black ravens, we can be maximally confident that all ravens indeed are black. Therefore, every time we observed something, that could have wrecked the proposed implication, but didn't, our confidence legitimately grew by a little. After each observation there was one less thing in the world that could prove the implication wrong.

Hence, each new observation, even a non-black non-raven like a red apple, can be considered as evidence for the implication $r \implies b$.

By how much should we increase our confidence about the proposition after each new observation? Proportionally to the size n of the world H. The more there are entities in H, the smaller the increment per observation.

Visually this process of observing (sampling) the entities in H and gradually increasing confidence in the proposed implication $r \implies b$, can be thought of as changing the relative areas of each of the logical relations in the initial Venn diagram. Each new observation, which is not a non-black raven, reduces the relative size of the white subset $r \wedge \neg b$, as can be seen in Figure 2.

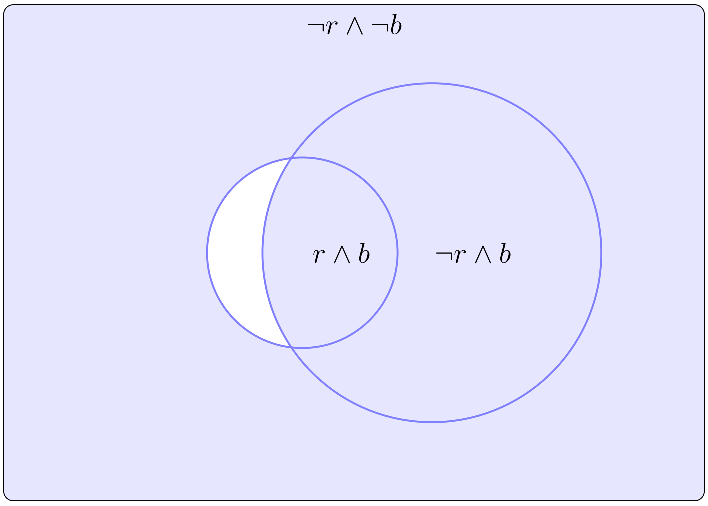


Figure 2: Sampling outside the black raven population and not finding any non-black ravens shrinks the $r \wedge \neg b$ option by small increments, indicated by white in the figure. However, when cardinality of the whole set of ravens and non-ravens approach infinity, the shrinking factor goes to zero.

When the set H grows towards infinity, the growth of confidence caused by new observations diminishes towards zero, but only becomes zero if the set H indeed is actually infinite.[2]

Are all observations equal in value? Should we try to find black ravens $r \wedge b$, look around randomly, or try to find the ravens which are not black $r \wedge \neg b$?

Remember, that it is enough even in the infinite world H to observe just one non-black raven to falsify the proposition. The larger the world H is, the more impactful is the falsifying observation relative to the confirming observation, which only have a minuscule increase in our confidence. Even if we searched only for the observations confirming the proposed implication and managed to find all of them (this is only even possible in the finite world), it still wouldn't prove the implication, if we had not ruled out the existence of the dis-confirming cases.

The same idea generalizes from simple implications to scientific theories. Karl Popper suggested, that our aim should be in falsifying the existing theories, instead of searching for evidence confirming them. This way we could learn quicker if and how our theories are false, and proceed by proposing better ones.

In an infinite world we can never arrive at full certainty regarding any implication or theory. But we can still give more weight to those theories, which have survived longer despite being subjected to falsification attempts. If a theory has been around for a long while, it means that many people have sampled multiple times from H without finding any instances from the falsifying white subset.

[2]Actually we could still update our confidence by increasing the number of observations sufficiently, so that the probability measure of the sample becomes strictly positive. This would require an infinite number of observations, though.

## All werewolves are vegan

Hopefully you know have some intuition, for why it's reasonable to observe for example an instance of $\neg p \wedge q$, and regard that as evidence for a hypothesis, that p implies q. Now let's break the intuition down again.

What if it was impossible for p and q to both be true at the same time, meaning that we could never observe $p \wedge q$. Would it still make sense to regard observing for example $\neg p \wedge q$ as evidence for the implication $p \implies q$? How would such situation even be possible?

Suppose, that the set denoted by the antecedent proposition p is empty, meaning that no members of that population exists; $p =$ "werewolves" $=: w = \emptyset$. Now it is impossible to falsify the hypothesis "All werewolves are vegan", $w \implies m$, as we can't find any werewolves who eat animal products, $w \wedge \neg v$. But just as before, observing non-werewolves of any kind gives us more confidence in the statement, that all werewolves are vegan.

But we can also propose another implication which is contradictory to the first one, that all werewolves eat meat, $w \implies m$. Likewise in the previous case, we can't falsify this claim either, as can be seen in Figure 3!

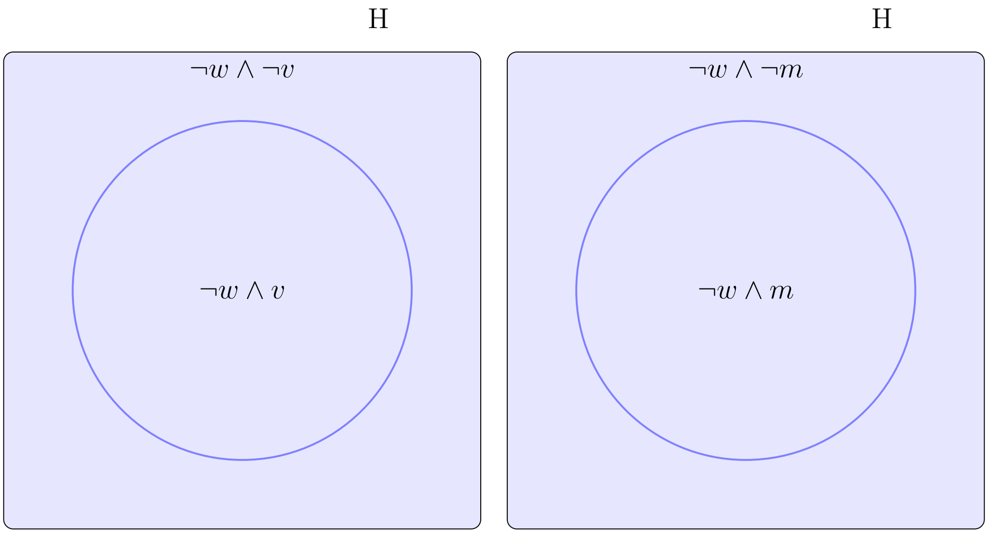


Figure 3: The world H on the left consists of entities that are vegan and those who aren't, based on the implication "all werewolves are vegan". On the right panel the same world H is similarly divided into those who eat meat and those who don't.

If we observed all of the possible (finite) cases in H, we could not find any werewolves who violated either of the proposed implications, because there are no werewolves at all. Because we can show that werewolves are meat eating vegans, meaning that both the proposition q and its negation ¬q can be true at the same time, we violate the law of non-contradiction.

Another similarly strange situation is, when both propositions do denote something in existence, but they can't co-exist. Let p=square and q=triangle. Now $p \wedge q = \emptyset$ by definition. In this case there surprisingly is no problem with the hempel's paradox, see Figure 4, as we can easily falsify this silly implication by observing any square which is not a triangle (naturally we also know a priori, that there can't be triangular squares). Thereby observing non-squares do shrink the set of leftover observations with the potential to falsify the claim.

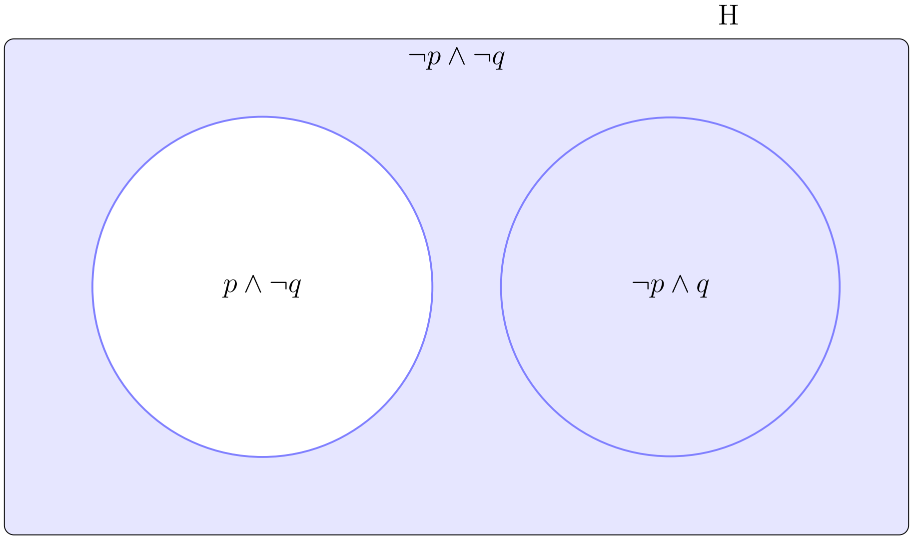


Figure 4: Existing but mutually exclusionary propositions.

The situation is also not paradoxical, if the consequent proposition q denotes for an empty set instead of the antecedent. Let's postulate an implication "All economists are hobbits", with $h = \emptyset$. Now the proposed implication can be rejected by finding any economist and confirming that she is not a hobbit, Figure 5.

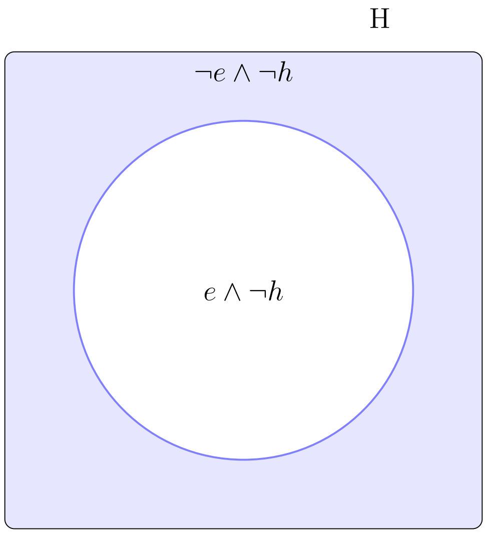


Figure 5: Implication $e \longrightarrow h$ with the consequent proposition h being an empty set.

Finally, if both of the propositions denote an empty set, then again we are back in trouble. Consider "Unicorns are faster than light", or $u \implies f$ with $u = \emptyset$ and

$f = \emptyset$, Figure 6. Now again there are no existing observation that could falsify the hypothesis, bringing back the paradox.

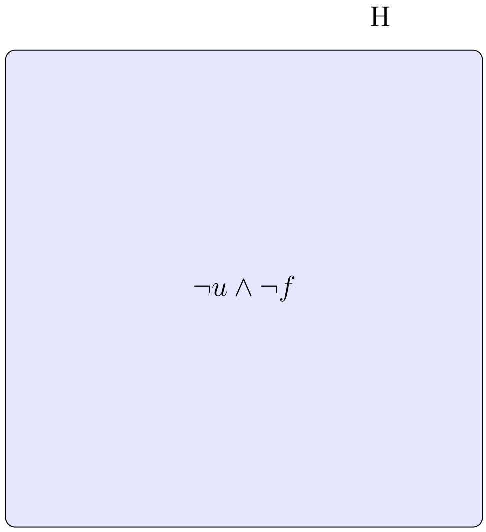


Figure 6: Implication $u \implies f$ with $u, f = \emptyset$

In the first and last of these strange examples we resurrected a new paradox to take the place of the one we expelled earlier. What is paradoxical here, is that the definition of implication allows us to learn about senseless statements by making virtually non-related observations.

Let this be a cautionary tale about using poorly-defined propositions, which prevents the falsification of the proposed implication even in principle. The troubles caused by these cartoonish example also illustrate, why previously mentioned Karl Popper suggested, that we should require all scientific propositions to be falsifiable.

## Afterword

We know, that infinities can cause troubles, as witnessed by a string of great minds from Zeno to Cantor. The moral of this article is, that also the opposite of infinity, an empty set or non-existence, can be puzzling. Therefore, we should imagine Sisyphus unhappy, as he is burdened with a combination of both: A never ending task that accomplishes nothing.

More seriously, the concept of (material) implication is well defined and Hempel's "paradox" is not really a paradox, as long as the propositions p and q are well defined.